Το Εγχειρίδιο της Δημοσιογραφίας Δεδομένων

Προς μια Κριτική Πρακτική Δεδομένων

Επιμέλεια: Liliana Bounegru, Jonathan Gray

Μετάφραση: Εβίτα Λύκου
Επιμέλεια ελληνικής έκδοσης: Κέλλυ Κική

Παραγωγή ελληνικής έκδοσης:

iMEdD: incubator for Media Education and Development

Το Εγχειρίδιο της Δημοσιογραφίας Δεδομένων

Προς μια Κριτική Πρακτική Δεδομένων

Επιμέλεια: Liliana Bounegru, Jonathan Gray
Μετάφραση: Εβίτα Λύκου
Συντονισμός έργου & Επιμέλεια ελληνικής έκδοσης: Κέλλυ Κική
Παραγωγή ελληνικής έκδοσης: iMEdD

Αθήνα, 2022

Η ελληνική έκδοση είναι παραγωγή του μη κερδοσκοπικού δημοσιογραφικού οργανισμού iMEdD (imedd.org) και είναι ψηφιακά διαθέσιμη από το iMEdD Lab (lab.imedd.org) εδώ: **datahandbook.lab.imedd.org**.

Η αγγλική έκδοση είναι παραγωγή των datajournalism.com, European Journalism Centre και Google News Initiative.

Εξώφυλλο: datajournalism.com

Γραφιστικός σχεδιασμός εντύπου: Ευγένιος Καλοφωλιάς

Σχεδιασμός ψηφιακής έκδοσης: Χριστίνα Δάλλα

Ανάπτυξη ψηφιακής έκδοσης: Δημήτρης Μακρής

e-ISBN (.pdf): 978-618-85968-1-8

e-ISBN (web html): 978-618-85968-2-5

ISBN: 978-618-85968-0-1

Creative Commons Άδεια Χρήσης CC BY-NC-ND 4.0
(https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode)

Το παρόν έργο διέπεται από την Creative Commons Άδεια Χρήσης με Αναφορά του Δημιουργού – για Μη Εμπορική Χρήση – Όχι Παράγωγα Έργα 4.0 (CC BY-NC-ND 4.0). Μερικά δικαιώματα διατηρούνται. Χωρίς να περιορίζονται τα δικαιώματα πνευματικής ιδιοκτησίας που διατηρούνται ανωτέρω, οποιοδήποτε μέρος του παρόντος βιβλίου μπορεί να αναπαραχθεί, να αποθηκευτεί ή να εισαχθεί σε σύστημα ανάκτησης ή να μεταδοθεί, σε οποιαδήποτε μορφή ή με οποιοδήποτε μέσο (ηλεκτρονικό, μηχανικό, φωτοαντιγραφικό, ηχογραφημένο ή άλλο). Μπορείτε να διαβάσετε την απλοποιημένη σύνοψη της άδειας χρήσης στα ελληνικά ή στα αγγλικά. Η απλοποιημένη σύνοψη της άδειας χρήσης δεν υποκαθιστά τον νομικό της κώδικα, ο οποίος είναι διαθέσιμος εδώ στα ελληνικά και εδώ στα αγγλικά.

Το προλογικό σημείωμα που περιλαμβάνεται στο παρόν έργο γράφτηκε για τους σκοπούς της ελληνικής έκδοσης. Το ίδιο ισχύει και για τις Σημειώσεις της Μεταφράστριας (ΣτΜ) και τις Σημειώσεις της Επιμελήτριας (ΣτΕ). Όπου υπάρχουν ΣτΕ του πρωτότυπου έργου, αυτές φέρουν την αντίστοιχη επισήμανση.

Περιεχόμενα

1. Πρόλογος & Εισαγωγή

1.1 Προλογικό Σημείωμα για την Ελληνική Έκδοση
Κέλλυ Κική 8

1.2 Εισαγωγή
Jonathan Gray, Liliana Bounegru 13

2. Θέματα με Δεδομένα

2.1 Από τον Καφέ στην Αποικιοκρατία: Πώς οι Φτωχοί Ταΐζουν τους Πλούσιους
Raúl Sánchez, Ximena Villagrán 26

2.2 Επαναχρησιμοποίηση Απογραφικών Δεδομένων για τον Υπολογισμό του Φυλετικού Διαχωρισμού στις ΗΠΑ
Aaron Williams 29

2.3 Πολλαπλασιάζοντας τις Αναμνήσεις με την Ανακάλυψη Δέντρων στην Μπογκοτά
Maria Isabel Magaña 32

2.4 Πίσω από τους Αριθμούς: Κατεδαφίσεις Σπιτιών στην Κατεχόμενη Ανατολική Ιερουσαλήμ
Mohammed Haddad 35

2.5 Χαρτογραφώντας Τροχαία Ατυχήματα για την Προώθηση της Οδικής Ασφάλειας στις Φιλιππίνες
Aika Rey 39

2.6 Παρακολούθηση των Εργατικών Θανάτων στην Τουρκία
Pinar Dağ 42

3. Συγκεντρώνοντας τα Δεδομένα

3.1 Δημιουργώντας το Δικό σας Σύνολο Δεδομένων: Καταγραφή των Εγκλημάτων με Μαχαίρια στο Ηνωμένο Βασίλειο
Caelainn Barr 45

3.2 Η Ιστορία ενός Αριθμού και πώς να Σταθεί Κανείς στο Πρόβλημα της Τιμής
Helen Verran 50

3.3 Κυριαρχία των Αυτοχθόνων επί των Δεδομένων τους: Επιπτώσεις για τη Δημοσιογραφία Δεδομένων
Tahu Kukutai, Maggie Walter 59

3.4 Εναλλακτικές Πρακτικές Δεδομένων στην Κίνα
Yolanda Jinxin Ma 66

3.5 Οικοδόμηση μιας Βάσης Δεδομένων για την Καταγραφή των Συγκρούσεων για τη Γη στην Ινδία
Kumar Sambhav Shrivastava, Ankur Paliwal 74

3.6 Ανασύνθεση Δημόσιων Δεδομένων στην Κούβα: Συνεργασίες όταν οι Πληροφορίες Απουσιάζουν, Είναι Παρωχημένες ή Ανεπαρκείς
Saimi Reyes Carmona, Yudivián Almeida Cruz, Ernesto Guerra 77

3.7 Δημιουργώντας Δεδομένα με τους Αναγνώστες της *La Nación*
Flor Coelho 81

3.8 Διεξαγωγή Έρευνας Κοινού στο πλαίσιο Δημοσιογραφικής Έρευνας
Crina-Gabriela Boroş 85

4. Δουλεύοντας με Δεδομένα

4.1 Δημοσιογραφία Δεδομένων: Πώς Σχετίζεται ο Φεμινισμός με την Τεχνολογία Πληροφοριών
Catherine D'Ignazio 90

4.2 Οικοδόμηση Συνεργασιών για τα *Panama* και *Paradise Papers*
Emilia Díaz-Struck, Cécile Schilis-Gallego, Pierre Romera 95

4.3 Το Κείμενο ως Δεδομένα: Εντοπίζοντας Θέματα σε Συλλογές Κειμένων
Barbara Maseda 101

4.4 Προγραμματισμός με Δεδομένα στην Αίθουσα Σύνταξης
Basile Simon 108

4.5 Εξηγώντας τις Μεθόδους: Υπολογιστικά Φύλλα, Script και Notebook Προγραμματισμού
Sam Leon 112

4.6 Δουλεύοντας Ανοιχτά στη Δημοσιογραφία Δεδομένων
Natalia Mazotte 121

4.7 Αξιοποιώντας τους Αλγορίθμους για το Ρεπορτάζ
Jonathan Stray 125

4.8 Δημοσιογραφία με Μηχανές; Από την Υπολογιστική Σκέψη στην Κατανεμημένη Νόηση
Eddy Borges-Rey 128

5. Η Εμπειρία της Εργασίας με Δεδομένα

5.1 Τρόποι Άσκησης της Δημοσιογραφίας Δεδομένων
Sarah Cohen 135

5.2 Οπτικοποιήσεις Δεδομένων: Τάσεις στις Αίθουσες Σύνταξης και Καθημερινές Αλληλεπιδράσεις
Helen Kennedy, William Allen, Martin Engebretsen, Rosemary Lucy Hill, Andy Kirk, Wibke Weber 140

5.3 Σκιτσάροντας με τα Δεδομένα
Mona Chalabi, Jonathan Gray 150

5.4 Ο Παγκόσμιος Ιστός ως Μέσο για την Οπτικοποίηση Δεδομένων
Elliot Bentley 157

5.5 Τέσσερις Πρόσφατες Εξελίξεις στα Ενημερωτικά Γραφήματα
Gregor Aisch, Lisa Charlotte Rost 166

5.6 Βάσεις Δεδομένων με Δυνατότητα Αναζήτησης ως Δημοσιογραφικό Προϊόν
Zara Rahman, Stefan Wehrmeyer 170

5.7 Η Αφήγηση Συγκρούσεων για τους Υδάτινους Πόρους Μέσω Δεδομένων και Διαδραστικών Κόμικς
Nelly Luna Amancio 178

5.8 Η Δημοσιογραφία Δεδομένων Πρέπει να Εστιάζει στους Ανθρώπους και στις Ιστορίες
Winny de Jong 182

6. Διερευνώντας τα Δεδομένα, τις Πλατφόρμες και τους Αλγορίθμους

6.1 Το Ρεπορτάζ των Αλγορίθμων: Οπτικές και Μέθοδοι Έρευνας
Nicholas Diakopoulos 188

6.2 Αφήγηση Ιστοριών με τα Μέσα Κοινωνικής Δικτύωσης
Lam Thuy Vo 197

6.3 Ψηφιακή Εγκληματολογία: Εκ νέου χρήση των ID στα Google Analytics
Richard Rogers 207

6.4 Εφαρμογές και οι Δυνατότητές τους για Έρευνες Δεδομένων
Esther Weltevrede 212

6.5 Ρίχνοντας Φως στους Αλγορίθμους: Συνεργατικές Έρευνες στο *Der Spiegel*
Christina Elmer 222

7. Οργάνωση της Δημοσιογραφίας Δεδομένων

7.1 Το Hashtag #ddj στο Twitter
Eunice Au, Marc Smith 230

7.2 Αρχειοθέτηση της Δημοσιογραφίας Δεδομένων
Meredith Broussard 236

7.3 Από τον *Guardian* στο Google News Lab: Μια Δεκαετία Εργασίας στη Δημοσιογραφία Δεδομένων
Simon Rogers 240
7.4 Οι Δεσμοί της Δημοσιογραφίας Δεδομένων με την Τεχνολογία των Πολιτών
Stefan Baack 246
7.5 Πρακτικές Ανοιχτού Κώδικα στη Δημοσιογραφία Δεδομένων
Ryan Pitts, Lindsay Muscato 250
7.6 Φεουδαρχία Δεδομένων: Πώς οι Πλατφόρμες Διαμορφώνουν Διασυνοριακά Ερευνητικά Δίκτυα
Ştefan Cândea 253
7.7 Κύρια Άρθρα Βάσει Δεδομένων; Σκέψεις για τη Χρήση Μετρήσεων Κοινού
Caitlin Petre 257

8. Μαθαίνοντας τη Δημοσιογραφία Δεδομένων από Κοινού

8.1 Δημοσιογραφία Δεδομένων, Ψηφιακή Οικουμενικότητα και Καινοτομία στην Περιφέρεια
Anita Say Chan 262
8.2 Δεδομενοποίηση της Δημοσιογραφίας: Στρατηγικές για Αφήγηση με Δεδομένα και Συνεργασία Βιομηχανίας-Ακαδημαϊκής Κοινότητας
Damian Radcliffe, Seth C. Lewis 268
8.3 Δημοσιογραφία Δεδομένων Από Περιθωριοποιημένες Κοινότητες, Σχετικά με Αυτές και για Λογαριασμό τους
Eva Constantaras 284
8.4 Διδάσκοντας Δημοσιογραφία Δεδομένων
Cheryl Phillips 290
8.5 Οργανώνοντας Έργα Δεδομένων με Γυναίκες και Μειονότητες στη Λατινική Αμερική
Eliana A. Vaca Muñoz 295

9. Χωροθέτηση της Δημοσιογραφίας Δεδομένων

9.1 Γενεαλογίες της Δημοσιογραφίας Δεδομένων
C.W. Anderson 299
9.2 Χρυσά Πρότυπα με Γνώμονα τα Δεδομένα: Τι Αποτιμά ο Κλάδος ως Δημοσιογραφία Δεδομένων που αξίζει Βραβείο και Πώς η Δημοσιογραφία Δεδομένων Εξελίσσεται μαζί με τη Δεδομενοποίηση της Κοινωνίας
Wiebke Loosen 306
9.3 Πέρα από τα Κλικ και τις Κοινοποιήσεις: Πώς και Γιατί Μετράμε τον Αντίκτυπο των Έργων Δημοσιογραφίας Δεδομένων
Lindsay Green-Barber 314
9.4 Δημοσιογραφία Δεδομένων: Προς το Συμφέρον Τίνος;
Mary Lynn Young, Candis Callison 321
9.5 Δημοσιογραφία Δεδομένων με Αντίκτυπο
Paul Bradshaw 329
9.6 Για Ποιον Λόγο Υπάρχει η Δημοσιογραφία Δεδομένων; Μετρητά, Κλικ, Δοκιμές και Σφάλματα
Nikki Usher 336
9.7 Δημοσιογραφία Δεδομένων και Ψηφιακός Φιλελευθερισμός
Dominic Boyer 343
Ευρετήριο Όρων 348
Ευρετήριο Ονομάτων 352

1.

Πρόλογος & Εισαγωγή

1.1
Προλογικό Σημείωμα για την Ελληνική Έκδοση

Κέλλυ Κική

Όταν, ως δημοσιογράφος, ξεκινάς να εργάζεσαι με δεδομένα, στο πρώτο στάδιο της εκμάθησης ακόμη, ο νέος κόσμος που σου αποκαλύπτεται είναι σαρωτικός –δηλαδή, και εντυπωσιακός και καταστροφικός: μένεις άναυδος μπροστά στις ερευνητικές δυνατότητες που ανακαλύπτεις ότι υπάρχουν και βρίσκεσαι πελαγωμένος εν μέσω αγωνιώδους προσπάθειας να αποκτήσεις τις κατάλληλες δεξιότητες, ώστε οι εν λόγω δυνατότητες να γίνουν και *δυνατότητές σου*. Στην αρχή, τα «error» πάντα είναι πολλά περισσότερα από τις μικρές επιτυχίες που έρχονται προσεχώς επί της οθόνης (του υπολογιστή σου). Κρέμεσαι από τα χείλη άλλων ανθρώπων, εμπειρότερων στον τομέα, που αδημονείς να σου εξηγήσουν, βήμα-βήμα, πώς ακριβώς το τάδε θα λυθεί και το δείνα θα δουλέψει. Τότε, όταν η απορία είναι η βασική σου κατάσταση, εκείνοι οι άλλοι άνθρωποι συχνά θα σου πουν απλώς, με την ηρεμία που διακρίνει όσους ξέρουν τι κάνουν και διογκώνει τη σύγχυση όσων δεν έχουν ιδέα πού πηγαίνουν, ότι «αυτό θα το βρεις, δεν είναι τίποτα».

Θα σου πάρει καιρό και ίσως μόνο όταν θα έχεις ήδη αποκτήσει κάποιες από τις τεχνικές δεξιότητες που απαιτεί η δημοσιογραφία δεδομένων, θα είσαι σε θέση να αντιληφθείς ότι εκείνοι οι άνθρωποι είχαν δίκιο και πως, πράγματι, δεν χρειάζεσαι a priori μια βίβλο οδηγιών για κάθε επιμέρους εργασία που δυνητικά θα εκτελέσεις ή για κάθε εργαλείο που θα χρησιμοποιήσεις. Σήμερα, περισσότερο από έξι δεκαετίες μετά τις πρώτες χρήσεις δεδομένων και υπολογιστικών μεθόδων στη δημοσιογραφία και περισσότερο από μια 12ετία διεθνούς ανάπτυξης και άνθησης της υπολογιστικής δημοσιογραφίας και της δημοσιογραφίας δεδομένων (Houston, 2021)[1], υπάρχει τέτοια πληθώρα μαθημάτων, workshop, tutorial και διαδικτυακών φόρουμ για την κάλυψη οποιουδήποτε πρακτικού «κενού» που είναι σαφές ότι το σημαντικότερο προσόν του δημοσιογράφου δεδομένων είναι να *θέλει* και να *έχει μάθει* να μαθαίνει –διαρκώς.

Παράλληλα με τη διάχυση τεχνικών, εργαλείων και καλών πρακτικών, σήμερα, που ακόμη μιλάμε για αυτό το «νέο είδος δημοσιογραφίας», εκείνο το οποίο ίσως έχει περισσότερη ανάγκη η κοινότητα είναι ένας διεθνής, εμπειρικός και εμπεριστατωμένος διάλογος για σημαντικά ζητήματα του κλάδου: Πόσο στ' αλήθεια *νέο* είναι αυτό το «νέο είδος»; Μάλιστα, μήπως ακριβώς αυτή η ακόμη υφιστάμενη αντίληψη περί «είδους» δημοσιογραφίας περιχαρακώνει υπερβολικά μια εξειδίκευση, ο ρόλος της οποίας δεν είναι άλλος από την ίδια την άσκηση της δημοσιογραφίας στην ψηφιακή εποχή μας; Στην καθημερινή πρακτική, η δημοσιογραφία δεδομένων θυμάται ότι πρώτιστος σκοπός της είναι η έρευνα ή μήπως εγκολπώνεται τη σύγχρονη γραφιστική σε βαθμό, με τρόπους και με ταχύτητες που εν τέλει συντηρείται ένας ανώριμος διαχωρισμός της από τη δημοσιογραφία εν γένει (ή, αν θέλετε, την «παραδοσιακή» δημοσιογραφία); Εν πάση περιπτώσει, τι σημαίνει πρακτικά η διεπιστημονικότητα, η ανοιχτότητα και η λογοδοσία στη δημοσιογραφία δεδομένων; Και πώς αυτή μπορεί να βρει τον δρόμο της σε

1. Βλ. επίσης το κεφάλαιο του Κρις Άντερσον στο παρόν έργο: Κεφάλαιο 9.1, «Γενεαλογίες της Δημοσιογραφίας Δεδομένων».

χώρες όπου η βιομηχανία των media, συγκριτικά, υπολείπεται πόρων, καινοτομίας, κατάρτισης και πρόσβασης σε δεδομένα;

Αυτά είναι μόνο μερικά από τα ερωτήματα που θέτει και τα ζητήματα που αναπτύσσει το παρόν έργο, *Το Εγχειρίδιο της Δημοσιογραφίας Δεδομένων: Προς μια Κριτική Πρακτική Δεδομένων*. Αυτός ο συλλογικός τόμος[2] (οι επιμελητές του πρωτότυπου έργου, Τζόναθαν Γκρέι και Λιλιάνα Μπούνεγκρου, τον αποκαλούν «συλλογικό πείραμα»[3]) απαρτίζεται από οκτώ θεματικές ενότητες και περιλαμβάνει συνολικά 55 κεφάλαια, γραμμένα από κορυφαίους δημοσιογράφους, επαγγελματίες και ερευνητές από ετερόκλητες γωνιές του πλανήτη. Το *Εγχειρίδιο* είναι ένα εγχειρίδιο ιδεών, απόψεων, πρακτικών και προβληματισμών: ανοίγει ακριβώς αυτόν τον διεθνή διάλογο, διαφωτίζοντας και καλώντας σε στοχασμό όσους ασκούν τη δημοσιογραφία δεδομένων, όσους τη διδάσκουν, όσους τη μαθαίνουν, όσους επιθυμούν να τη γνωρίσουν και όσους αποτελούν φορείς συνεργασίας ή συνέργειας.

Για τους δημοσιογράφους δεδομένων –και όχι μόνο

Οι δημοσιογράφοι που δεν είναι εξοικειωμένοι με τη συλλογή στοιχείων, την επεξεργασία τους, την ανάλυση και την οπτικοποίησή τους θα γνωρίσουν τρόπους άσκησης της δημοσιογραφίας δεδομένων ανά τον κόσμο –μεταξύ άλλων, από τις ΗΠΑ έως την Τουρκία, από την Κούβα και την Αργεντινή έως την Ινδία και την Κίνα, από τη Γερμανία έως το Πακιστάν, από το Ηνωμένο Βασίλειο έως την Κένυα. Μέσα από αφηγήσεις δημοσιογράφων για τον τρόπο επιλογής των θεμάτων, οργάνωσης των ομάδων εργασίας, ανάκτησης, ελέγχου και χρήσης των στοιχείων, συγκροτείται μια εργαλειοθήκη χρήσιμη, για περαιτέρω αναζήτηση και μελέτη, σε όσους θα ήθελαν να γνωρίσουν τον τομέα και να σχεδιάσουν τα δικά τους πρώτα βήματα.

Διαβάζοντας τις ίδιες αφηγήσεις, όσοι ασκούν τη δημοσιογραφία δεδομένων θα «συνομιλήσουν» με συναδέλφους, θα ταυτιστούν με καταστάσεις και θα αντλήσουν έμπνευση από άλλες. Για παράδειγμα, όταν η Κάελαν Μπαρ γράφει για την καταγραφή των εγκλημάτων με μαχαίρια στο Ηνωμένο Βασίλειο[4], σημειώνει ότι, ελλείψει επίσημων δημόσιων δεδομένων, «κατά το τρέχον έτος, έπρεπε να αρχίσουμε να τους καταμετρούμε καθώς συνέβαιναν [...] Αποφασίσαμε να παρακολουθούμε τα στοιχεία οι ίδιοι, μέσω αστυνομικών αναφορών, αποδελτίωσης ειδήσεων, Google Alerts, Facebook και Twitter». Σε αυτήν την περίπτωση, για εμάς στο iMEdD Lab, η ταύτιση ήταν φυσικό επόμενο –επί μήνες, η καταγραφή των θανάτων από COVID-19 στην Ελλάδα ανά γεωγραφική περιοχή, η συγκρότηση και η συντήρηση του σχετικού συνόλου δεδομένων μας δεν είχε ως πρώτη ύλη παρά την αποδελτίωση του Τύπου που γινόταν από εμάς[5]. Σε μια άλλη περίπτωση, διαβάζοντας από τη Φλορ Κοέλο για την εμπειρία της *La Nación* στις πρωτοβουλίες crowdsourcing και μαθαίνοντας ότι συνταξιούχοι, φοιτητές και άλλες ομάδες πολιτών στην Αργεντινή «επιστρατεύτηκαν» εθελοντικά στη συλλογή δεδομένων από χιλιάδες έγγραφα .pdf (στο πλαίσιο ερευνών για θέματα δημοσίου συμφέροντος, όπως τα έξοδα της Γερουσίας), χαμογελάσαμε με ειλικρινή συμπάθεια: δεν σκεφτήκαμε να οργανώσουμε κάτι αντίστοιχο στο πλαίσιο της συνεχιζόμενης έρευνάς μας για τους πλειστηριασμούς στην Ελλάδα, αλλά θα ήταν μια ιδέα δυνητικά αξιοποιήσιμη από κάθε ομάδα δεδομένων όσο οι πληροφορίες για

2. Αρχικά, κυκλοφόρησε από τις Εκδόσεις του Πανεπιστημίου του Άμστερνταμ τον Μάρτιο 2021 και, επίσης, διατίθεται στα αγγλικά στο datajournalism.com, σε παραγωγή των European Journalism Centre και Google News Initiative.

3. Βλ. την εισαγωγή των επιμελητών του πρωτότυπου έργου: Κεφάλαιο 2.1, «Εισαγωγή».

4. Κεφάλαιο 3.1, «Δημιουργώντας το Δικό σας Σύνολο Δεδομένων: Καταγραφή των Εγκλημάτων με Μαχαίρια στο Ηνωμένο Βασίλειο»

5. Στο πλαίσιο του έργου του iMEdD Lab για την πορεία της πανδημίας στην Ελλάδα και στον κόσμο. Έως την ώρα που γράφονται αυτές οι γραμμές, τον Φεβρουάριο 2022, οι αρμόδιοι φορείς δεν έχουν ανακοινώσει επισήμως τη γεωγραφική κατανομή των θανάτων από COVID-19 στη χώρα.

οποιοδήποτε θέμα δημοσίου ενδιαφέροντος εξακολουθούν να παραμένουν διαθέσιμες αλλά εσώκλειστες, διάσπαρτες και αδόμητες –συνεπώς, πραγματικά δυσπρόσιτες.

Η έκδοση, επίσης, αποτελεί πηγή αυτοκριτικής και εν μέρει αναθεώρησης στάσεων για τους δημοσιογράφους δεδομένων: Μας υπενθυμίζει ότι δεν είναι αυτοσκοπός να παρασυρόμαστε σε ένα δέκαθλο τεχνικών δεξιοτήτων, καθώς «είμαστε εδώ για να μεταδίδουμε ειδήσεις και να λέμε ιστορίες –όχι για να κάνουμε εξάσκηση», όπως γράφει η Γουίνι ντε Γιόνγκ[6]. Μας δημιουργεί δεύτερες σκέψεις για την ενδοκλαδική «γκρίνια» σχετικά με το γεγονός ότι, ελλείψει ολοκληρωμένων, μηχαναγνώσιμων συνόλων δεδομένων από τους θεσμούς και τους φορείς της πολιτείας, συχνά «κάνουμε ό,τι θα έπρεπε να κάνει το Κράτος, ενώ η δική μας δουλειά θα έπρεπε να ξεκινά από την ανάλυση» –όπως όλοι μας έχουμε βρεθεί να αναφέρουμε, με τον έναν ή με τον άλλο τρόπο, για να είμαστε ειλικρινείς. Μέχρι ποιο σημείο είναι στ' αλήθεια έτσι; Η πολιτεία πρέπει να λογοδοτεί, να είναι διαφανής, να σέβεται και να διαφυλάττει το δικαίωμα των πολιτών για πρόσβαση στην πληροφορία. Όμως, όταν εκείνη δεν το πράττει, η δημοσιογραφία δεδομένων αξίζει να θυμάται την «ακτιβιστική» διάσταση της ύπαρξής της, που ανέκαθεν υπήρχε και περιελάμβανε την ανακάλυψη, την οργάνωση και τη διάθεση ανοιχτών δεδομένων –καθιστώντας, τελικά, την πολιτεία υπόλογη για όσα δεν μας λέει, ενώ οφείλει. Δηλαδή, σε αυτή τη –θλιβερή για τη δημοκρατία μας– περίπτωση, η δημοσιογραφία δεδομένων δεν κάνει παρά ό,τι είναι η πεμπτουσία της δημοσιογραφίας συνολικά: απελευθερώνει την πληροφορία και ασκεί έλεγχο στην εξουσία, για το κοινό καλό. Και πάλι, αυτές είναι μόνο μερικές από τις σκέψεις που γεννά το βιβλίο σε εν ενεργεία δημοσιογράφους δεδομένων.

Όμως, και δημοσιογράφοι που –νομίζουν ότι– δεν χρησιμοποιούν δεδομένα στη δουλειά τους, ρεπόρτερ με εξειδίκευση σε θέματα τεχνολογίας, ακόμη και πολιτικοί ή οικονομικοί συντάκτες θα ανακαλύψουν το ρεπορτάζ των αλγορίθμων[7]: την ανάγκη, που ήδη αποτελεί δημοσιογραφική εξειδίκευση, όχι μόνο να χρησιμοποιήσουμε τη μηχανική μάθηση («machine learning») στο πλαίσιο δημοσιογραφικής έρευνας, αλλά να κάνουμε ρεπορτάζ και να διεξάγουμε έρευνα για τα αποτελέσματα των αλγοριθμικών αποφάσεων στην ιδιωτική και συλλογική ζωή, από το λεγόμενο «profiling» των ανθρώπων στα συστήματα διακυβέρνησης έως την αξιολόγηση της πιστοληπτικής μας ικανότητας –οι δημοσιογράφοι κάνουν πλέον αρκετό ρεπορτάζ *με* δεδομένα, αλλά ακόμη δεν κάνουν αρκετό ρεπορτάζ *για* τα δεδομένα, όπως είχε αναφερθεί σε μια ενδιαφέρουσα σχετική συζήτηση.

Για τα media που (δεν) έχουν επενδύσει στη δημοσιογραφία δεδομένων

Παράλληλα, το *Εγχειρίδιο* μπορεί να αποτελέσει ερέθισμα για τη γενικώς λεγόμενη «αγορά». Ενδεικτικά, ο Έντι Μπόρχες-Ρέι[8] αναφέρεται σε αγγελίες για θέσεις εργασίας, όπως Δημοσιογράφος Μηχανικής Μάθησης ή Συντάκτης Αυτοματισμού. Η Σάρα Κοέν συνοψίζει εκπληκτικά τις πολλές και διάφορες εξειδικεύσεις υπό την ομπρέλα της δημοσιογραφίας δεδομένων, γράφοντας ότι «αυτό που κάποτε ήταν μια μπάντα σε γκαράζ έχει πλέον εξελιχθεί σε ολόκληρη ορχήστρα»[9]. Με μια άλλη αναφορά στη μουσική (ενδιαφέρουσα σύμπτωση), ο Σάιμον Ρότζερς, αποτιμώντας μια δεκαετία εργασίας στη δημοσιογραφία δεδομένων, αναφέρεται στο άρθρο του, που δημοσιεύθηκε το 2012, με τίτλο *Η Δημοσιογραφία Δεδομένων είναι η Νέα Πανκ* και εξηγεί πως «το νόημα ήταν το εξής: βρισκόμασταν σε ένα σημείο όπου η δημιουργικότητα μπορούσε πραγματικά να αφεθεί ελεύθερη, αλλά και εκεί όπου το έργο θα γινόταν, τελικά,

6. Κεφάλαιο 5.8, «Η Δημοσιογραφία Δεδομένων Πρέπει να Εστιάζει στους Ανθρώπους και στις Ιστορίες»
7. Μεταξύ άλλων, βλ. το Κεφάλαιο 6.1, «Το Ρεπορτάζ των Αλγορίθμων: Οπτικές και Μέθοδοι Έρευνας», από τον Νίκολας Ντιακόπουλος και το Κεφάλαιο 6.5, «Ρίχνοντας Φως στους Αλγορίθμους: Συνεργατικές Έρευνες στο *Der Spiegel*» από την Κριστίνα Έλμερ.
8. Κεφάλαιο 4.8, «Δημοσιογραφία με Μηχανές; Από την Υπολογιστική Σκέψη στην Κατανεμημένη Νόηση»
9. Κεφάλαιο 5.1, «Τρόποι Άσκησης της Δημοσιογραφίας Δεδομένων»

mainstream»[10]. Οι παράγοντες της αγοράς, όσοι ρυθμίζουν τη διαδικασία παραγωγής σε ένα newsroom και όσοι καταρτίζουν τα οργανογράμματα των μέσων, μπορούν να βρουν ενδιαφέρον σε αυτά και άλλα κεφάλαια. Εάν δρουν σε αγορές όπου η δημοσιογραφία δεδομένων είναι αναπόσπαστο κομμάτι των εγκαθιδρυμένων μέσων, μπορούν να αναλογιστούν σχετικά με αυτό που λέει η Νίκι Άσερ, ότι «η δημοσιογραφία δεδομένων, στην καλύτερη περίπτωση, έχει γίνει βαρετή και υπερβολικά επαγγελματική και, στη χειρότερη, έχει γίνει άλλη μια στρατηγική για τη δημιουργία ψηφιακών εσόδων»[11]. Όσοι δεν αντιμετωπίζουν ακόμη τέτοια στοχαστικά προβλήματα, επειδή δρουν σε αγορές για τις οποίες η δημοσιογραφία δεδομένων παραμένει ένα άγνωστο ή αμήχανο είδος πολυτελείας για διάφορους λόγους, προκαλούνται να κάνουν τη δική τους ενδοσκόπηση και να αναλογιστούν τι σημαίνει ακριβώς αυτό, ότι το πρόβλημα που θέτει η Άσερ δεν τους αφορά, για τη δική τους δημοσιογραφική και εκδοτική σκηνή.

Για την ακαδημαϊκή και ευρύτερα ερευνητική κοινότητα

Οι δάσκαλοι των αυριανών δημοσιογράφων μπορούν να βρουν το δικό τους ειδικό ενδιαφέρον στην έκδοση, ιδίως εάν ακόμη αναζητούν τον βηματισμό τους ως προς τα επιμέρους αντικείμενα διδασκαλίας της δημοσιογραφίας δεδομένων, τη διδακτική μέθοδο, τους ανθρώπους που μπορούν να καταστήσουν κοινωνούς στην εκπαίδευση των φοιτητών τους αλλά και τους τρόπους με τους οποίους η συνεργασία ακαδημαϊκής κοινότητας και βιομηχανίας μπορεί δυνητικά να είναι αμοιβαία εποικοδομητική, ξεπερνώντας την «ιστορικά μάλλον δύσκολη σχέση» τους, όπως αναφέρεται στο κεφάλαιο των Ντέμιεν Ράντκλιφ και Σεθ Λούις[12].

Στατιστικοί, επιστήμονες των δεδομένων, επαγγελματίες της πληροφορικής, φορείς της τεχνολογίας των πολιτών (και γενικά οι «κομπιουτεράδες», με την πολύ καλώς εννοούμενη σημασία της λέξης, όπως εκείνοι τιμητικά θα έλεγαν «γραφιάδες» για τους δημοσιογράφους) μπορούν να μάθουν πόσα πολλά έχουν «δανείσει» στη δημοσιογραφία δεδομένων, πόσο εκείνη τους χρειάζεται, καθώς αποτελούν συστατικό κομμάτι της, και πως ο εναγκαλισμός των επιστημών τους και της δημοσιογραφίας έχει ήδη συντελεστεί διεθνώς.

Για τη δημοσιογραφία συνολικά

Πάντως, οι κριτικές σκέψεις για τη δημοσιογραφία δεδομένων που αναπτύσσονται σε αυτό το *Εγχειρίδιο*, αφορούν τη δημοσιογραφία εν γένει. Ενδεικτικά, ο Σαμ Λίον διαπιστώνει ότι «με την άνθηση της δημοσιογραφίας δεδομένων, οι ιδέες γύρω από το τι μπορεί να θεωρηθεί δημοσιογραφική πηγή αλλάζουν», αναδεικνύει την πρόκληση για δημοσιογραφική λογοδοσία σχετικά με τις ακολουθούμενες μεθόδους και αναφέρεται τόσο στην ανάγκη για αναπαραγωγιμότητα του δημοσιογραφικού έργου όσο και στην ήδη εφαρμοζόμενη αξιολόγησή του από ομότιμους[13] («peer review»· να μια χαρακτηριστική κληρονομιά από τον ακαδημαϊκό και επιστημονικό κόσμο). Η Νατάλια Μαζότε, που αναλύει το ζήτημα της ανοιχτότητας στη δημοσιογραφία δεδομένων, σημειώνει ότι, σε αντίθεση με το γνωστό δημοσιογραφικό απόρρητο που «σημαίνει την εξασφάλιση της ασφάλειας της πηγής», «όταν οι πληροφορίες είναι άφθονες, το να μη διαμοιραζόμαστε πώς φτάσαμε ως εκεί μπορεί να στερήσει από τον αναγνώστη τα μέσα για να κατανοήσει την ιστορία και να βγάλει νόημα από αυτήν»[14]. Ο Κρις Άντερσον εγείρει την ανα-

10. Κεφάλαιο 7.3, «Από τον Guardian στο Google News Lab: Μια Δεκαετία Εργασίας στη Δημοσιογραφία Δεδομένων»
11. Κεφάλαιο 9.6, «Για Ποιον Λόγο Υπάρχει η Δημοσιογραφία Δεδομένων; Μετρητά, Κλικ, Δοκιμές και Σφάλματα»
12. Κεφάλαιο 8.2, «Δεδομενοποίηση της Δημοσιογραφίας: Στρατηγικές για Αφήγηση με Δεδομένα και Συνεργασία Βιομηχανίας-Ακαδημαϊκής Κοινότητας»
13. Κεφάλαιο 4.5, «Εξηγώντας τις Μεθόδους: Υπολογιστικά Φύλλα, Script και Notebook Προγραμματισμού»
14. Κεφάλαιο 4.6, «Δουλεύοντας Ανοιχτά στη Δημοσιογραφία Δεδομένων»

γκαία παραδοχή της προσωρινότητας και της αβεβαιότητας, έναντι της αυτοπεποίθησης ή και της αλαζονείας, εκ μέρους των δημοσιογράφων (δεδομένων) και σημειώνει ότι «εάν αυτό γίνει σωστά, θα μπορούσε να δημιουργηθεί μια δημοσιογραφία που μοιάζει περισσότερο, και όχι λιγότερο, με επιστήμη»[15]. Και η Νίκι Άσερ κριτικά προτείνει ότι «η δημοσιογραφία δεδομένων πρέπει να επιστρέψει στις ρίζες της αλλαγής, της επανάστασης [...] του αυτοπροσδιοριζόμενου οράματος των αποστατών που κινούνται μέσα σε μια κουρασμένη και ανέμπνευστη βιομηχανία» και ότι οι δημοσιογράφοι δεδομένων πρέπει «να "χακάρουν" την αίθουσα σύνταξης, όπως είχαν υποσχεθεί κάποτε ότι θα κάνουν»[16].

Εάν όλα αυτά, και πολλά άλλα που περιλαμβάνονται σε αυτόν τον συλλογικό τόμο, συμπυκνώνουν την κριτική ματιά στη δημοσιογραφία δεδομένων, δηλαδή τη δύναμη της αναζωογόνησης του επαγγέλματος, ποιος θα μπορούσε, τελικά, να πει ότι δεν ενδιαφέρουν τη δημοσιογραφία στο σύνολό της;

Για όλους όσους αφορά, λοιπόν, αυτός ο τόμος είναι που εμείς στο iMEdD αποφασίσαμε το εγχείρημα της μετάφρασης και της παραγωγής της ελληνικής έκδοσής του, θεωρώντας ότι αξίζει αυτή να προσφερθεί στην κοινότητα και να αποτελέσει μέρος και της ελληνικής βιβλιογραφίας για το παρόν ή το μέλλον της δημοσιογραφίας δεδομένων. Καμιά φορά, είναι σημαντικό, άλλωστε, κάποιος να μιλά τη γλώσσα σου –πράγμα που επίσης σημαίνει, μην ξεχνάμε, ότι μοιράζεστε κοινές καταβολές, ανησυχίες και προσδοκίες.

Παραπομπές

Houston, B. (2021) «The history of data journalism: A historical take on every critical breakthrough from the 1950s until today», *datajournalism.com*, 13 Δεκεμβρίου. Διαθέσιμο σε: https://datajournalism.com/read/longreads/the-history-of-data-journalism (Προσπελάστηκε: 13 Φεβρουαρίου 2022).

Η συγγραφέας

Η Κέλλυ Κική είναι δημοσιογράφος δεδομένων και Project Manager στο iMEdD Lab, τον πυλώνα παραγωγής πρωτογενούς δημοσιογραφικού περιεχομένου με τα δεδομένα στον πυρήνα του, στον μη κερδοσκοπικό δημοσιογραφικό οργανισμό iMEdD. Επιμελήθηκε την ελληνική έκδοση του παρόντος έργου, *Το Εγχειρίδιο της Δημοσιογραφίας Δεδομένων: Προς Μια Κριτική Πρακτική Δεδομένων*.

15. Κεφάλαιο 9.1, «Γενεαλογίες της Δημοσιογραφίας Δεδομένων»
16. Κεφάλαιο 9.6, «Για Ποιον Λόγο Υπάρχει η Δημοσιογραφία Δεδομένων; Μετρητά, Κλικ, Δοκιμές και Σφάλματα»

1.2
Εισαγωγή

Jonathan Gray, Liliana Bounegru

Περίληψη: Μια εισαγωγή στον τόμο, ως συλλογικό πείραμα σχετικά με τις πρακτικές της δημοσιογραφίας δεδομένων σε ολόκληρο τον κόσμο, η οποία περιλαμβάνει μια επισκόπηση των ενοτήτων και κεφαλαίων του και δώδεκα προκλήσεις για μια κριτική πρακτική δεδομένων.

Λέξεις-κλειδιά: δημοσιογραφία δεδομένων, κριτική πρακτική δεδομένων, κριτικές σπουδές δεδομένων, ψηφιακές μέθοδοι, επιστημονικές και τεχνολογικές σπουδές, σπουδές Διαδικτύου

Η δημοσιογραφία δεδομένων υπό αμφισβήτηση

Τι είναι η δημοσιογραφία δεδομένων; Γιατί υπάρχει; Τι μπορεί να επιτύχει; Ποιες ευκαιρίες και ποιους περιορισμούς παρουσιάζει; Ποιοι άνθρωποι και ποιες προϋποθέσεις εμπλέκονται στη δημιουργία και στην κατανόησή της; Αυτός ο τόμος είναι ένα συλλογικό πείραμα που απαντά σε αυτές και σε άλλες ερωτήσεις σχετικά με τις πρακτικές, τις κουλτούρες, τις πολιτικές και το περιβάλλον της δημοσιογραφίας δεδομένων σε ολόκληρο τον κόσμο. Έρχεται μετά από ένα άλλο βιβλίο, με τίτλο *The Data Journalism Handbook: How Journalists Can Use Data to Improve the News* (Gray κ.α., 2012). Και τα δυο βιβλία συγκεντρώνουν πλειάδα φωνών και οπτικών σχετικά με τον εξελισσόμενο τομέα της δημοσιογραφίας δεδομένων. Η πρώτη έκδοση ξεκίνησε από ένα «συγγραφικό σπριντ»[1] στο Mozilla Festival του Λονδίνου το 2011, που συγκέντρωσε δημοσιογράφους, τεχνολόγους, ομάδες πίεσης και άλλους, με σκοπό να γράψουν κείμενα για τον τρόπο με τον οποίο λειτουργεί η δημοσιογραφία δεδομένων. Όπως σημειώσαμε στην εισαγωγή, σκοπός εκείνου του βιβλίου ήταν «να καταγραφεί το πάθος και ο ενθουσιασμός, το όραμα και η ενέργεια ενός εκκολαπτόμενου κινήματος», να προσφέρει «ιστορίες πίσω από τις ιστορίες» και να αφήσει «τις διαφορετικές φωνές και απόψεις να λάμψουν» (Gray κ.α., 2012). Η έκδοση του 2012 έχει μεταφραστεί σε περισσότερες από δέκα γλώσσες –ανάμεσα στις οποίες αραβικά, κινέζικα, τσεχικά, γαλλικά, γεωργιανά, ελληνικά, ιταλικά, μακεδονικά, πορτογαλικά, ρωσικά, ισπανικά και ουκρανικά– και χρησιμοποιείται για διδασκαλία σε πολλά κορυφαία πανεπιστήμια και κέντρα εκπαίδευσης ανά τον κόσμο, ενώ αποτελεί πηγή που εμφανίζεται σε πολλές παραπομπές για τους ερευνητές του πεδίου.

Μολονότι ο τόμος του 2012 χρησιμοποιείται ακόμη ευρέως (και ο παρών τόμος έχει ως στόχο περισσότερο να τον συμπληρώσει παρά να τον αντικαταστήσει), πολλά έχουν αλλάξει από το 2012. Από τη μία, η δημοσιογραφία δεδομένων έχει καθιερωθεί πολύ περισσότερο. Το 2011 αυτού του είδους η δημοσιογραφία δεδομένων ήταν σε μεγάλο βαθμό ένα πεδίο «εν τω γίγνεσθαι», και μόνο μια χούφτα άνθρωποι χρησιμοποιούσαν τον συγκεκριμένο όρο. Έκτοτε, έχει περάσει στην κοινωνία και έχει καθιερωθεί μέσα από σχετικούς οργανισμούς, εκπαιδευτικά προγράμματα, θέσεις εργασίας, επαγγελματικές ομάδες, βραβεία, ανθολογίες, επιστημονικά άρθρα, αναφορές, εργαλεία, διαδικτυακές κοινότητες, hashtag, συνέδρια, δίκτυα, συναντήσεις

1. ΣτΜ: «Book sprint», πρόκειται για μέθοδο συλλογικής συγγραφής βιβλίου σε σύντομο χρονικό διάστημα, συνήθως τριών έως πέντε ημερών.

δικτύωσης (τα λεγόμενα «meetup»), λίστες ηλεκτρονικού ταχυδρομείου και άλλα. Υπάρχει, επίσης, μια ευρύτερη κατανόηση του όρου μέσα από γεγονότα που καταφανώς σχετίζονται με τα δεδομένα, όπως τα Panama Papers, που ο Έντουαρντ Σνόουντεν, ως μάρτυρας δημοσίου συμφέροντος, είχε χαρακτηρίσει τότε ως «τη μεγαλύτερη διαρροή στην ιστορία της δημοσιογραφίας δεδομένων» (Snowden, 2016).

Από την άλλη πλευρά, η δημοσιογραφία δεδομένων καθίσταται αμφιλεγόμενο ζήτημα ολοένα και περισσότερο. Οι διαρροές του Σνόουντεν το 2013 βοήθησαν να επιβεβαιωθεί δημοσίως η ύπαρξη ενός διεθνούς μηχανισμού παρακολούθησης κρατών και τεχνολογικών εταιρειών ως πραγματικό γεγονός –και όχι ως εικασία. Αυτές οι διαρροές έδειξαν πώς εκτίθενται πληροφορίες που αφορούν τους πολίτες μέσω πρακτικών των μεγάλων δεδομένων («big data»), παρουσιάζοντας τις οικείες συσκευές, εφαρμογές και πλατφόρμες που παράγουν δεδομένα υπό ένα πιο σκοτεινό πρίσμα (Gray & Bounegru, 2019). Στις Ηνωμένες Πολιτείες, το 2014, η έναρξη λειτουργίας του *FiveThirtyEight* του Νέιτ Σίλβερ, που εξειδικεύεται στη δημοσιογραφία δεδομένων, έγινε δεκτή με αντιδράσεις για την υπερβολική εμπιστοσύνη που έδειχνε σε συγκεκριμένα είδη ποσοτικών μεθόδων, περιφρονώντας τη «δημοσιογραφία γνώμης» (Byers, 2014). Ενώ ο Σίλβερ ανακηρύχθηκε «άρχοντας και θεός των αλγορίθμων» από τον Τζον Στιούαρτ του *Daily Show*, για την επιτυχή πρόβλεψη του εκλογικού αποτελέσματος το 2012[2], οι στατιστικές μέθοδοι που υποστήριξε δέχθηκαν περαιτέρω κριτική και αμφισβήτηση μετά την εκλογή του Ντόναλντ Τραμπ το 2016. Αυτές οι εκλογές, μαζί με την ψηφοφορία για το Brexit στο Ηνωμένο Βασίλειο και την άνοδο λαϊκιστών ηγετών της δεξιάς σε ολόκληρο τον κόσμο, λέγεται πως συμπίπτουν με μια στιγμή «μετα-αλήθειας» (Davies, 2016), η οποία χαρακτηρίζεται από εκτεταμένη απώλεια εμπιστοσύνης απέναντι στους δημόσιους θεσμούς, στην τεχνογνωσία των εμπειρογνωμόνων και στη διαμεσολάβηση του δημόσιου και πολιτικού βίου από διαδικτυακές πλατφόρμες, οι οποίες καθιστούν τους χρήστες τους ευάλωτους στη στοχοποίηση, στη χειραγώγηση και στην παραπληροφόρηση[3].

Είτε η λεγόμενη στιγμή της «μετα-αλήθειας» εκλαμβάνεται ως απόδειξη αποτυχίας είτε ως έκκληση για δράση, ένα πράγμα είναι σαφές: τα δεδομένα δεν μπορούν πλέον να θεωρούνται δεδομένα, το ίδιο και η δημοσιογραφία δεδομένων. Τα δεδομένα δεν παρέχουν απλώς ουδέτερες και άμεσες αναπαραστάσεις του κόσμου, αλλά είναι μάλλον συνυφασμένα με την πολιτική και την κουλτούρα, το χρήμα και την εξουσία. Οι θεσμοί και οι υποδομές που στηρίζουν την παραγωγή δεδομένων –από τις έρευνες έως τις στατιστικές, από την κλιματολογία έως τις πλατφόρμες κοινωνικής δικτύωσης– τελούν υπό αμφισβήτηση. Κατά τη στιγμή της συγγραφής, και ενώ η πανδημία της COVID-19 συνεχίζει να καλπάζει σε ολόκληρο τον κόσμο, οι αριθμοί, τα γραφήματα και οι κατατάξεις έχουν γίνει ευρέως διαδεδομένα, θεματοποιημένα, πολιτικοποιημένα και αποπολιτικοποιημένα –όπως μαρτυρούν τα επιδημιολογικά διαγράμματα που κυκλοφορούν σε καθημερινή βάση, τα οποία αναφέρονται στην πανταχού παρούσα πλέον στρατηγική της δημόσιας υγείας για την «επιπέδωση της καμπύλης». Την ίδια στιγμή, η αστάθεια και η προσωρινότητα αυτών των δεδομένων έχει γίνει ευρέως αντικείμενο αναφοράς, με ανησυχίες σχετικά με την ελλιπή καταγραφή, τη μη καταγραφή και την ταξινόμηση των κρουσμάτων, καθώς και με την αυξανόμενη συνειδητοποίηση των κοινωνικών και πολιτικών επιπτώσεων των διαφορετικών ειδών δεδομένων από πηγές –από νοσοκομειακά στοιχεία έως ερευνητικές εκτιμήσεις, αυτοκαταγραφές και δεδομένα κίνησης από εφαρμογές ιχνηλάτησης επαφών. Η πανδημία έχει διευρύνει την ευαισθητοποίηση όχι μόνο όσον αφορά

2. ΣτΕ (ελληνική έκδοση): Το 2008, ο στατιστικός Νέιτ Σίλβερ, *ιδρυτής και αρχισυντάκτης του FiveThirtyEight,* είχε προβλέψει με επιτυχία το εκλογικό αποτέλεσμα σε 49 από 50 πολιτείες. Έναν χρόνο αργότερα, ο ίδιος ανακηρύχθηκε, από το περιοδικό *Time*, ως ένας από τους 100 ανθρώπους με τη μεγαλύτερη επιρροή στον κόσμο. Το 2012, οι προβλέψεις του για το εκλογικό αποτέλεσμα ήταν σωστές και για τις 50 πολιτείες. Το –έως τότε blog– *FiveThirtyEight* άρχισε να λειτουργεί ως μέσο το 2014, όταν πωλήθηκε στον όμιλο ESPN.

3. Για μια κριτική οπτική του όρου αυτού, βλ. Jasanoff, S. & Simmet, H. R. (2017), «No funeral bells: Public reason in a "post-truth" age», *Social Studies of Science*, 47(5), σσ. 751-770.

στη χρήση αλλά και στην κριτική παρουσίαση αριθμών και δεδομένων. Έτσι, θα μπορούσε κανείς να αναρωτηθεί σχετικά με τη χρήση των δεδομένων στη δημοσιογραφία: ποια δεδομένα, τίνος τα δεδομένα και με ποια μέσα; Δεδομένα για ποια θέματα και για ποιον σκοπό; Ποια είδη θεμάτων είναι πλούσια σε δεδομένα και ποια είναι φτωχά σε δεδομένα και γιατί; Ποιος έχει τη δυνατότητα να επωφεληθεί από τα δεδομένα και ποιος όχι; Τι είδους κοινό συγκεντρώνουν τα δεδομένα, τι είδους ικανότητες υποστηρίζουν, τι είδους πολιτικές θεσπίζουν και τι είδους συμμετοχή γεννούν;

Προς μια κριτική πρακτική δεδομένων

Αντί να περιχαρακώσει αυτά τα ερωτήματα και τους προβληματισμούς, ο παρών τόμος στόχο έχει να «σταθεί στο πρόβλημα», όπως το θέτει η εξέχουσα φεμινίστρια ακαδημαϊκός Ντόνα Χαραγουέι (2016)[4]. Αντί να αντιμετωπίζουμε τη σημασία και τη σπουδαιότητα της δημοσιογραφίας δεδομένων ως απόφανση, μπορούμε να τις αντιμετωπίσουμε ως ένα ερώτημα που δύναται να απαντηθεί με πολλαπλούς τρόπους. Η συλλογή κεφαλαίων που περιλαμβάνονται σε αυτό το βιβλίο πασχίζει να προσφέρει μια πιο πλούσια ιστορία σχετικά με το τι κάνει η δημοσιογραφία δεδομένων, με ποιον το κάνει και για ποιον το κάνει. Μέσα από την επιμέλεια επιδιώξαμε να ενθαρρύνουμε τον προβληματισμό σχετικά με όσα μπορούν να επιτύχουν τα εγχειρήματα δημοσιογραφίας δεδομένων και τις συνθήκες υπό τις οποίες μπορούν να το επιτύχουν. Αυτό συνεπάγεται την καλλιέργεια ενός διαφορετικού είδους ακρίβειας σχετικά με τις πρακτικές της δημοσιογραφίας δεδομένων: τον καθορισμό των καταστάσεων στις οποίες αυτή αναπτύσσεται και λειτουργεί. Η ακρίβεια απαιτεί τη διεύρυνση του πεδίου αναφοράς του βιβλίου, ώστε να συμπεριλάβει όχι μόνο τους τρόπους με τους οποίους τα δεδομένα αναλύονται, δημιουργούνται και χρησιμοποιούνται στο πλαίσιο της δημοσιογραφίας αλλά και τις κοινωνικές, πολιτιστικές, πολιτικές και οικονομικές συνθήκες στις οποίες εντάσσονται οι πρακτικές αυτές.

Ο υπότιτλος αυτού του νέου τόμου είναι «Προς μια Κριτική Πρακτική Δεδομένων» και αντανακλά τόσο τη φιλοδοξία μας ως επιμελητών να ασκήσουμε κριτικό προβληματισμό στις πρακτικές της δημοσιογραφίας δεδομένων, όσο και την ολοένα και πιο κριτική στάση των επαγγελματιών της δημοσιογραφίας δεδομένων. Η έννοια της «κριτικής πρακτικής δεδομένων» κλείνει το μάτι στην έννοια της «κριτικής τεχνικής πρακτικής» του Φίλιπ Έιγκερ, την οποία ο ίδιος περιγράφει σαν να πατάει κανείς «με το ένα πόδι στο χειρωνακτικό έργο του σχεδιασμού και με το άλλο στο αναστοχαστικό έργο της κριτικής» (Agre, 1997, σ. 155). Όπως έχουμε γράψει και αλλού, αυτό που μας ενδιαφέρει σε αυτόν τον τόμο είναι να κατανοήσουμε πώς η κριτική ενασχόληση με τα δεδομένα μπορεί να τροποποιήσει τις πρακτικές δεδομένων, δίνοντας χώρο στη δημόσια φαντασία και στις δημόσιες παρεμβάσεις σχετικά με τις πολιτικές δεδομένων (Gray, 2018· Gray κ.α., 2018).

Παράλληλα με τη συνεισφορά δημοσιογράφων δεδομένων και επαγγελματιών που γράφουν για αυτό που κάνουν, το βιβλίο περιλαμβάνει, επίσης, κεφάλαια από ερευνητές των οποίων το έργο μπορεί να προωθήσει τον κριτικό προβληματισμό σχετικά με τις πρακτικές δημοσιογραφίας δεδομένων, από πεδία όπως η ανθρωπολογία, οι επιστημονικές και τεχνολογικές σπουδές, οι σπουδές (νέων) μέσων, οι σπουδές Διαδικτύου, η μελέτη των πλατφορμών, η κοινωνιολογία της ποσοτικοποίησης, οι σπουδές δημοσιογραφίας, οι μελέτες αυτόχθονων πληθυσμών, οι φεμινιστικές σπουδές, οι ψηφιακές μέθοδοι και η ψηφιακή κοινωνιολογία. Αντί για έναν πιο παραδοσιακό καταμερισμό εργασίας, όπου οι ερευνητές θα προέβαιναν στον κριτικό προβληματισμό και οι επαγγελματίες θα πρότειναν περισσότερο εργαλειακές πληροφορίες και συμβουλές, επιδιώξαμε να ενθαρρύνουμε τους ερευνητές να εξετάσουν την πρακτική σημασία

4. Αναφερόμενο σε αυτό το έργο, το κεφάλαιο της Βέραν, σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 3.2, «Η Ιστορία ενός Αριθμού και πώς να Σταθεί Κανείς στο Πρόβλημα της Τιμής»), διερευνά τον τρόπο με τον οποίο οι δημοσιογράφοι δεδομένων μπορούν να σταθούν στο πρόβλημα της τιμής και των αριθμών.

της δουλειάς τους και να προσφέρουμε χώρο στους επαγγελματίες, για να προβληματιστούν πάνω σε αυτό που κάνουν, αφήνοντας στην άκρη τις καθημερινές προθεσμίες τους.

Καμία από αυτές τις διαφορετικές οπτικές δεν εξαντλεί το πεδίο και στόχος μας είναι να ενθαρρύνουμε τους αναγνώστες να παρακολουθήσουν διαφορετικές πτυχές του τρόπου με τον οποίο γίνεται η δημοσιογραφία δεδομένων. Με άλλα λόγια, το βιβλίο αυτό προορίζεται να λειτουργήσει ως έναυσμα για μια διεπιστημονική συζήτηση και –ελπίζουμε– ως καταλύτης για συνεργασίες.

Δεν υποθέτουμε ότι ο όρος «δημοσιογραφία δεδομένων» αναφέρεται σε ένα μεμονωμένο, ενιαίο σύνολο πρακτικών. Αντίθετα, πρόκειται για έναν ξεχωριστό, περιγραφικό όρο ο οποίος αναφέρεται σε πλειάδα διαφορετικών πρακτικών που μπορούν να μελετηθούν εμπειρικά, να προσδιοριστούν και να γίνουν αντικείμενο πειράματος. Όπως σημειώνεται σε μια πρόσφατη κριτική, πρέπει να διερευνήσουμε το «πώς της ποσοτικοποίησης όσο και το απλό γεγονός της [ύπαρξής της]», τα αποτελέσματα της οποίας «εξαρτώνται από τις προθέσεις και την υλοποίηση» (Berman & Hirschman, 2018). Σκοπός μας δεν είναι να κανονικοποιήσουμε τον τρόπο με τον οποίο γίνεται η δημοσιογραφία των δεδομένων, αλλά να επιστήσουμε την προσοχή στις ποικίλες πτυχές της και να ανοίξουμε χώρο, ώστε να μπορούμε να την κάνουμε διαφορετικά.

Ένα συλλογικό πείραμα

Αξίζει να αναφέρουμε εν συντομία τι δεν είναι αυτός ο τόμος. Δεν είναι απλώς ένα εγχειρίδιο ή διδακτικό βιβλίο με τη συμβατική έννοια: τα κεφάλαια δεν καταλήγουν σε ένα εμπεριστατωμένο σώμα γνώσης, αλλά έχουν μάλλον ως στόχο να υποδείξουν ενδιαφέρουσες κατευθύνσεις για περαιτέρω έρευνα και πειραματισμό. Το βιβλίο δεν είναι απλώς ένας πρακτικός οδηγός με εκπαιδευτικά βοηθήματα ή συμβουλευτικά άρθρα τύπου «πώς να κάνετε το τάδε»: υπάρχει ήδη πληθώρα άμεσα διαθέσιμου υλικού και μαθημάτων σχετικά με διάφορες πτυχές της πρακτικής των δεδομένων (π.χ. ανάλυση δεδομένων και οπτικοποίηση δεδομένων). Δεν είναι απλώς ένα βιβλίο με «παρασκηνιακές» περιπτωσιολογικές μελέτες: υπάρχει πλήθος άρθρων και αναρτήσεων σε ιστολόγια που δείχνουν πώς έγινε το κάθε έργο, συμπεριλαμβανομένων συνεντεύξεων με τους δημιουργούς του. Δεν πρόκειται απλώς για ένα βιβλίο με πρόσφατες ακαδημαϊκές οπτικές: υπάρχει ένα αναδυόμενο σώμα βιβλιογραφίας σχετικά με τη δημοσιογραφία δεδομένων, διάσπαρτο σε πολυάριθμα βιβλία και περιοδικά[5].

Αντ' αυτού, το βιβλίο έχει σχεδιαστεί ως ένα συλλογικό πείραμα σχετικά με τις πρακτικές της δημοσιογραφίας δεδομένων των τελευταίων ετών και ως μια συλλογική πρόσκληση για τη διερεύνηση του τρόπου με τον οποίο αυτές οι πρακτικές μπορούν να τροποποιηθούν. Είναι συλλογικό υπό την έννοια ότι, όπως και στην πρώτη έκδοση, καταφέραμε να συγκεντρώσουμε έναν συγκριτικά μεγάλο αριθμό συνεργατών (περισσότεροι από 70) για έναν μικρό τόμο. Για τη διαδικασία της επιμέλειας, επωφεληθήκαμε από τις προτάσεις των συνεργατών μας μέσω ανταλλαγής ηλεκτρονικών μηνυμάτων. Ένα εργαστήριο με ορισμένους από τους συντελεστές στο Διεθνές Φεστιβάλ Δημοσιογραφίας στην Περούτζια το 2018 έδωσε τη δυνατότητα για ανταλλαγές απόψεων και προβληματισμό. Μια δοκιμαστική έκδοση του βιβλίου κυκλοφόρησε διαδικτυακά, προκειμένου να δώσει την ευκαιρία για τη δημόσια προεπισκόπηση επιλεγμένων κεφαλαίων πριν από τη δημοσίευση της έντυπης έκδοσης και να εκμαιεύσει σχόλια και ζυμώσεις πριν να πάρει το βιβλίο την τελική μορφή του. Μέσα από αυτό που θα μπορούσε να θεωρηθεί ένα είδος επιμελημένης «συντακτικής χιονοστιβάδας», προσπαθήσαμε να παρακολουθήσουμε πώς ασκείται η δημοσιογραφία δεδομένων από διαφορετικούς φορείς, σε διαφορετικά μέρη, γύρω από διαφορετικά θέματα, με διαφορετικά μέσα. Κατά τη διάρκεια αυτής της διαδικασίας συλλέξαμε και ερευνήσαμε πολλές βραχείες λίστες, μακρές λίστες, μέσα

5. https://www.zotero.org/groups/data_journalism_research

ενημέρωσης και σύνολα δεδομένων, προκειμένου να συγκεντρώσουμε διαφορετικές οπτικές σχετικά με τις πρακτικές δημοσιογραφίας δεδομένων. Μολονότι υπήρχαν πολλοί, πολλοί περισσότεροι συντελεστές που θα θέλαμε να συμπεριλάβουμε, έπρεπε να λειτουργήσουμε εντός των περιορισμών ενός βιβλίου που επρόκειτο να τυπωθεί, καθώς και να δώσουμε φωνή σε διαφορετικές γεωγραφικές περιοχές, θέματα, ανησυχίες και φύλα.

Το βιβλίο είναι πειραματικό υπό την έννοια ότι τα κεφάλαια προσφέρουν ποικίλες οπτικές και προκλήσεις για τη δημοσιογραφία δεδομένων, τις οποίες καλούμε τους αναγνώστες να εξερευνήσουν περαιτέρω μέσω της ενεργής διαμόρφωσης των δικών τους συνδυασμών εργαλείων, συνόλων δεδομένων, μεθόδων, κειμένων, κοινού και θεμάτων. Αντί να αποδεχθούμε άκριτα τους τρόπους με τους οποίους αντιλαμβανόμαστε και μαθαίνουμε τα πράγματα, που έχουν «ενσωματωθεί» σε στοιχεία όπως τα επίσημα σύνολα δεδομένων ή τα δεδομένα από τα μέσα κοινωνικής δικτύωσης, ενθαρρύνουμε την αντιμετώπισή τους ως αντικείμενα αναζήτησης από τους ίδιους τους αναγνώστες. Έτσι, ακολουθούμε το πνεύμα της «κριτικής ανάλυσης» και των «επινοητικών μεθόδων», οι οποίες αποσκοπούν στην τροποποίηση των ερωτήσεων που τίθενται και του τρόπου με τον οποίο διατυπώνονται οι προβληματισμοί (Lury & Wakeford, 2012· Rogers, 2018). Η δημοσιογραφία δεδομένων μπορεί να εξεταστεί όχι μόνο ως προς τον τρόπο με τον οποίο αναπαρίστανται τα πράγματα, αλλά και ως προς τον τρόπο με τον οποίο οργανώνει τις σχέσεις –έτσι ώστε να μην καταλήγει απλώς να παράγει θέματα δεδομένων (μέσω της συλλογής, της ανάλυσης, της οπτικοποίησης και της αφήγησης δεδομένων), αλλά να εστιάζει, επίσης, σε όσους και σε όσα έρχονται πιο κοντά μέσα από τα θέματα αυτά (συμπεριλαμβανομένων των ακροατηρίων, των πηγών, των μεθόδων, των θεσμών και των μέσων κοινωνικής δικτύωσης). Έτσι, μπορούμε να θέσουμε το ερώτημα, όπως το διατύπωσε πρόσφατα η Νόρτζε Μάρες: «Ποιες είναι οι μέθοδοι, τα υλικά, οι τεχνικές και οι διαδικασίες που επιμελούμαστε, προκειμένου να δημιουργήσουμε χώρους όπου τα προβλήματα μπορούν να αντιμετωπιστούν με διαφορετικό τρόπο;»[6]. Τα κεφάλαια σε αυτό το βιβλίο δείχνουν πώς η δημοσιογραφία δεδομένων μπορεί να είναι μια επινοητική, ευφάνταστη, συνεργατική τέχνη, αναδεικνύοντας τον τρόπο με τον οποίο οι δημοσιογράφοι δεδομένων εξετάζουν τις επίσημες πηγές δεδομένων, φτιάχνουν και συνθέτουν τα δικά τους δεδομένα, δοκιμάζουν νέες οπτικές και διαδραστικά φορμάτ, προβληματίζονται σχετικά με τα αποτελέσματα της εργασίας τους, και καθιστούν τις μεθόδους τους ανοιχτές στη λογοδοσία και τον κώδικά τους επαναχρησιμοποιήσιμο. Εάν το μέλλον της δημοσιογραφίας δεδομένων είναι αβέβαιο, τότε ελπίζουμε ότι οι αναγνώστες αυτού του βιβλίου θα συμμετάσχουν μαζί μας στην κριτική αποτίμηση του παρόντος και του παρελθόντος της δημοσιογραφίας, καθώς και στην παρέμβαση για τη διαμόρφωση του μέλλοντός της. Όπως συμβαίνει με όλα τα έργα, η επιτυχία, η αποτυχία και η τελική μοίρα αυτού του βιβλίου-πειράματος, τελικά, εξαρτάται από εσάς, τους αναγνώστες του, από τους τρόπους που θα το χρησιμοποιήσετε, τις αντιδράσεις και την ανταπόκριση που θα προκαλέσει.

Η εικόνα του εξωφύλλου του βιβλίου[7] είναι μια φωτογραφία από το έργο *Fixed Points Finding a Home* («Σταθερά Σημεία Βρίσκουν Σπίτι») της Σάρα Ζι, στο Μουσείο Σύγχρονης Τέχνης του Λουξεμβούργου (Mudam), για την οποία ευχαριστούμε την καλλιτέχνιδα, την γκαλερί της και το

6. Ένα ερώτημα που έθεσε η Νόρτζε Μάρες στην εισήγησή της στην ολομέλεια της EASST 2018 στο Λάνκαστερ: https://twitter.com/jwyg/status/1023200997668204544

7. ΣτΕ (ελληνική έκδοση): Οι επιμελητές του πρωτότυπου έργου αναφέρονται στο εξώφυλλο του πρωτότυπου έντυπου τόμου *The Data Journalism Handbook: Towards a Critical Data Practice*, που κυκλοφορεί από τις εκδόσεις του Πανεπιστημίου του Άμστερνταμ. Μια εικόνα του έργου της Ζι, *Fixed Points Finding a Home*, είναι διαθέσιμη εδώ: https://artsandculture.google.com/asset/sarah-sze-fixed-points-finding-a-home-2012-sarah-sze/dQF6IsV03pH1yA

μουσείο, που μας παραχώρησαν την άδεια αναπαραγωγής[8]. Αν και μπορεί η τοποθέτηση ενός έργου γλυπτικής στο εξώφυλλο ενός βιβλίου για τη δημοσιογραφία να μη μοιάζει προφανής επιλογή, σκεφτήκαμε ότι αυτή η εικόνα θα μπορούσε να ενθαρρύνει μια σχεσιακή προοπτική πάνω στη δημοσιογραφία δεδομένων ως ένα είδος επιμελητικής τέχνης, η οποία συλλέγει και εργάζεται με ποικίλα υλικά, κοινότητες και υποδομές, προκειμένου να δημιουργήσει διαφορετικούς τρόπους γνώσης, αφήγησης και θέασης του κόσμου σε διαφορετικές κλίμακες και χρονικές στιγμές. Αντί να επικεντρωθούμε στην παραγωγή της δημοσιογραφίας δεδομένων (π.χ. με στιγμιότυπα οθόνης οπτικοποιήσεων ή διαδραστικών μέσων), θελήσαμε να αντικατοπτρίσουμε τα διαφορετικά είδη διαδικασιών και συλλογικοτήτων που εμπλέκονται, όταν κάνουμε δημοσιογραφία με δεδομένα. Συναντήσαμε τυχαία τις εκθέσεις της Ζι στο Mudam, στην Μπιενάλε της Βενετίας, στο ZKM[9], στην Tate και όχι μόνο, και το έργο της μας απορρόφησε τόσο βαθιά που σκεφτήκαμε πως θα μπορούσε να προσφέρει μια διαφορετική (και ελπίζουμε λιγότερο γνώριμη) οπτική γωνία για την πρακτική της δημοσιογραφίας δεδομένων, η οποία θα είχε απήχηση στις σχεσιακές προοπτικές για τις υποδομές πληροφοριών και τις «συναρμολογήσεις δεδομένων»[10]. Οι εγκαταστάσεις της Ζι ενσαρκώνουν μια ακριβή και παιχνιδιάρικη ευαισθησία απέναντι στην επαναχρησιμοποίηση ευρεθέντων υλικών, η οποία παρουσιάζει έναν οπτικό παραλληλισμό με αυτό που ελπίζαμε να τονίσουμε, όταν επιμελούμασταν τις διαφορετικές πτυχές της δημοσιογραφίας δεδομένων για το βιβλίο. Ο Μπρούνο Λατούρ πρόσφατα έγραψε ότι η προσέγγιση της Ζι στη συναρμολόγηση υλικών μοιάζει να επικυρώνει τις «συνθετικές ασυνέχειες» (Latour, 2020) –κάτι που ταιριάζει πολύ με την ελπίδα μας να ενθαρρύνουμε την «κριτική πρακτική δεδομένων» και να αφηγηθούμε ιστορίες τόσο *με* όσο και *για* τα ποικίλα υλικά και τους φορείς που εμπλέκονται στη δημοσιογραφία δεδομένων, όπως συζητάμε παρακάτω, καθώς και με την προσέγγιση της επιμέλειας για την υποστήριξη των διαφορετικών στυλ, φωνών, διαλέκτων και ενδιαφερόντων των κεφαλαίων αυτού του βιβλίου.

Μια συνοπτική παρουσίαση του βιβλίου

Για να παραμείνουμε πιστοί στην έμφαση που θέλει να δώσει η επιμέλεια στον προσδιορισμό του πλαισίου, σημειώνουμε ότι ο προσανατολισμός του βιβλίου και η επιλογή των κεφαλαίων του είναι επηρεασμένος από τα ενδιαφέροντα, τα δικά μας και εκείνα των φίλων, των συναδέλφων και των δικτύων μας, αυτή τη συγκεκριμένη χρονική στιγμή –συμπεριλαμβανομένων των αυξανόμενων ανησυχιών για την κλιματική αλλαγή, την περιβαλλοντική καταστροφή, την ατμοσφαιρική ρύπανση, τη φοροδιαφυγή, τη (νεο)αποικιοκρατία, τον ρατσισμό, τον σεξισμό, την ανισότητα, τον εξτρακτιβισμό[11], τον αυταρχισμό, την αλγοριθμική αδικία και την εργασία σε ψηφιακές πλατφόρμες. Τα κεφάλαια διερευνούν τον τρόπο με τον οποίο η δημοσιογραφία δεδομένων καθιστά τέτοια ζητήματα κατανοητά και ικανά να βιωθούν, καθώς και τα είδη των

8. Sarah S. (2012) *Fixed Points Finding a Home* (λεπτομέρεια) [Μεικτή τεχνική. Μεταβλητών διαστάσεων]. Παραγγελία και Συλλογή Mudam, Λουξεμβούργο. Δωρεά 2012—Les Amis des Musées d'Art et d'Histoire Luxembourg. © Έργο τέχνης: Σάρα Ζι. Ευγενική παραχώρηση της καλλιτέχνιδας και της Βικτόρια Μίρο. © Φωτογραφία: Αντρέ Λεχόνα/Mudam, Λουξεμβούργο.

9. ΣτΕ (ελληνική έκδοση): Κέντρο για την Τέχνη και την Τεχνολογία των Μέσων στην πόλη Καρλρούη στη Γερμανία

10. Για σχεσιακές προοπτικές ως προς τις υποδομές δεδομένων, για παράδειγμα, βλ. το σημαίνον έργο της Σούζαν Λι Σταρ: Star, S. L., & Ruhleder, K. (1996) «Steps toward an ecology of infrastructure: Design and access for large information spaces», *Information Systems Research*, 7, σσ. 111-134· Star, S. L. (1999) «The ethnography of infrastructure», *American Behavioral Scientist*, 43, σσ. 377-391. Για πιο πρόσφατες εργασίες σχετικά με τις «συναρμολογήσεις δεδομένων», βλ. για παράδειγμα: Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Λονδίνο: Sage· Kitchin, R., & Lauriault, T. (2018) «Towards critical data studies: Charting and unpacking data assemblages and their work», στο Thatcher, J., Shears, A. & Eckert, J. (επιμ.), *Thinking big data in geography*: New regimes, new research, σσ. 3-20. Λίνκολν, Νεμπράσκα: University of Nebraska Press.

11. ΣτΜ: Η διαδικασία εξαγωγής φυσικών πόρων από τη γη προς πώληση στη διεθνή αγορά.

αντιδράσεων που μπορεί να κινητοποιήσει. Η επιλογή των κεφαλαίων, επίσης, αντανακλά τις δικές μας ταλαντεύσεις μεταξύ ακαδημαϊκής έρευνας, δημοσιογραφίας και δράσης, καθώς και τα διαφορετικά στυλ γραφής και πρακτικής δεδομένων που συνδέονται με καθένα από αυτά.

Συνεχίζουμε να πιστεύουμε στη γενεσιουργό δυναμική των συναντήσεων μεταξύ συναδέλφων σε αυτά τα διαφορετικά πεδία, και πολλά από τα κεφάλαια μαρτυρούν επιτυχημένες διατομεακές συνεργασίες. Εκτός από τη διερεύνηση των συνεργειών και των κοινών σημείων, αξίζει επίσης να σημειωθεί εξαρχής (όπως θα παρατηρήσουν οι οξυδερκείς αναγνώστες) ότι υπάρχουν διαφορές, εντάσεις και τριβές μεταξύ των οπτικών που παρουσιάζονται στα διάφορα κεφάλαια, συμπεριλαμβανομένων διαφορετικών ιστοριών και ιστοριών προέλευσης· διαφορετικών απόψεων σχετικά με τις μεθόδους, τα δεδομένα και τις αναδυόμενες τεχνολογίες· διαφορετικών απόψεων για το πόσο επιθυμητή είναι η τυποποίηση και ο πειραματισμός με διαφορετικές προσεγγίσεις· και διαφορετικών απόψεων για το τι είναι η δημοσιογραφία δεδομένων, τι εξυπηρετεί, τις προϋποθέσεις και τους περιορισμούς της, τον τρόπο οργάνωσής της και τις δυνατότητες που παρουσιάζει.

Μετά από αυτήν την εισαγωγή, το βιβλίο ξεκινά με μια «πρόγευση» σχετικά με τη διαχείριση θεμάτων με δεδομένα. Αυτή περιλαμβάνει ποικιλία διαφορετικών φορμάτ για την κατανόηση διαφορετικών θεμάτων σε διαφορετικούς τόπους –μεταξύ άλλων, αναζητούμε συνδέσεις μεταξύ των γεωργικών προϊόντων, του εγκλήματος, της διαφθοράς και της αποικιοκρατίας σε διάφορες χώρες (Σάντσεζ και Βιγιαγκράν), χαρτογραφούμε τους φυλετικούς διαχωρισμούς στις Ηνωμένες Πολιτείες (Γουίλιαμς), πολλαπλασιάζουμε ανθρώπινες αναμνήσεις από δέντρα στην Μπογκοτά (Μαγκάνια), ρίχνουμε μια ματιά στους ανθρώπους και στις σκηνές πίσω από τους αριθμούς των κατεδαφίσεων σπιτιών στην κατεχόμενη Ανατολική Ιερουσαλήμ (Χαντάντ), κινητοποιούμαστε για την οδική ασφάλεια στις Φιλιππίνες (Ρέι) και ιχνηλατούμε τους θανάτους εργαζομένων στην Τουρκία (Νταγκ). Τα κεφάλαια αυτής της ενότητας απεικονίζουν ένα ευρύ φάσμα πρακτικών, από τεχνικές οπτικοποίησης έως τη δημιουργία εκστρατειών και την επαναχρησιμοποίηση επίσημων δεδομένων με διαφορετικές προτεραιότητες ανάλυσης.

Η δεύτερη ενότητα επικεντρώνεται στον τρόπο με τον οποίο οι δημοσιογράφοι συγκεντρώνουν δεδομένα –ένας σημαντικός αναδυόμενος τομέας τον οποίο επιδιώξαμε να αναδείξουμε στο βιβλίο και στη σχετική έρευνα (Gray κ.α., 2018· Gray & Bounegru, 2019). Εξερευνάται η δημιουργία έργων για θέματα όπως τα εγκλήματα με μαχαίρια (Μπαρ) και οι συγκρούσεις για τη γη (Σριβαστάβα και Παλιγουάλ). Επίσης, περιγράφεται πώς μπορεί κανείς να αποκτήσει δεδομένα και να δουλέψει μαζί τους σε χώρες όπου ενδέχεται να είναι πιο δύσκολο να βρεθούν, όπως στην Κούβα (Ρέγιες, Αλμέιντα και Γκέρα) και στην Κίνα (Μα). Η συγκέντρωση δεδομένων, επίσης, μπορεί να είναι ένας τρόπος για αλληλεπίδραση με τους αναγνώστες (Κοέλο) και τη συγκέντρωση φορέων που ενδιαφέρονται για ένα θέμα, το οποίο μπορεί από μόνο του να αποτελεί σημαντικό αποτέλεσμα ενός έργου. Η συλλογή δεδομένων μπορεί να περιλαμβάνει την προσαρμογή άλλων μορφών παραγωγής γνώσης, όπως οι δημοσκοπήσεις και οι έρευνες, στο πλαίσιο της δημοσιογραφίας (Μπόρος). Ένα κεφάλαιο σχετικά με την κυριαρχία των αυτοχθόνων επί των δεδομένων τους (Κουκουτάι και Γουόλτερ) διερευνά κοινωνικά, πολιτιστικά και πολιτικά ζητήματα γύρω από τα επίσημα δεδομένα και τον τρόπο με τον οποίο μπορούν να αξιοποιηθούν άλλες περιθωριοποιημένες οπτικές στην οργάνωση του συλλογικού βίου με δεδομένα. Εκτός από τη χρήση των αριθμών ως υλικού για την αφήγηση ιστοριών, οι δημοσιογράφοι δεδομένων μπορούν, επίσης, να αφηγηθούν ιστορίες για το πώς δημιουργούνται οι αριθμοί (Βέραν).

Η τρίτη ενότητα ασχολείται με τους διαφορετικούς τρόπους εργασίας με δεδομένα. Αυτό περιλαμβάνει αλγορίθμους (Στρέι), κώδικα (Σιμόν) και μηχανές (Μπόρχες-Ρέι). Οι συντελεστές εξετάζουν τα αναδυόμενα ζητήματα και τις ευκαιρίες που προκύπτουν, όταν κανείς δουλεύει με πηγές όπως τα κειμενικά δεδομένα (Μασέντα). Άλλοι εξετάζουν πρακτικές για τη διαφάνεια, τη λογοδοσία και την αναπαραγωγιμότητα των έργων της δημοσιογραφίας δεδομένων (Λίον,

Μαζότε). Οι βάσεις δεδομένων μπορούν, επίσης, να προσφέρουν ευκαιρίες για συνεργατική εργασία σε μεγάλα ερευνητικά έργα (Ντιάζ-Στρακ, Γκαλιέγο και Ρομέρα). Η φεμινιστική σκέψη και πρακτική, επίσης, μπορεί να εμπνεύσει διαφορετικούς τρόπους εργασίας με δεδομένα (Ντ' Ιγκνάζιο).

Η τέταρτη ενότητα είναι αφιερωμένη στην εξέταση των διαφορετικών τρόπων με τους οποίους μπορούν να βιώνονται τα δεδομένα, ξεκινώντας με μια ματιά στα διαφορετικά φορμάτ που μπορεί να λάβει η δημοσιογραφία δεδομένων (Κοέν). Κάποια κεφάλαια προβληματίζονται για τις σύγχρονες πρακτικές οπτικοποίησης (Άις και Ροστ), καθώς και για τον τρόπο με τον οποίο οι αναγνώστες ανταποκρίνονται και συμμετέχουν στην παραγωγή νοήματος με οπτικοποιήσεις (Κένεντι κ.α.). Άλλα κεφάλαια εξετάζουν τον τρόπο με τον οποίο τα δεδομένα παρουσιάζονται στους αναγνώστες διαμεσολαβημένα, μέσω βάσεων δεδομένων (Ραχμάν και Βιαμάγιερ), διαδικτυακών διαδραστικών μέσων (Μπέντλεϊ), τηλεόρασης και ραδιοφώνου (ντε Γιόνγκ), κόμικς (Λούνα) και σκίτσων (Τσαλαμπί και Γκρέι).

Η πέμπτη ενότητα είναι αφιερωμένη στις αναδυόμενες προσεγγίσεις για τη διερεύνηση δεδομένων, πλατφορμών και αλγορίθμων. Πρόσφατα δημοσιογραφικά έργα θεωρούν ότι η ψηφιακή τεχνολογία δεν προσφέρει μόνο νέες τεχνικές και ευκαιρίες για τους δημοσιογράφους αλλά και νέα αντικείμενα για έρευνα. Παραδείγματα αυτού είναι το ευρέως διαδεδομένο έργο της [ερευνητικής ομάδας] Bellingcat και του *BuzzFeed News* σχετικά με το περιεχόμενο που γίνεται viral, την παραπληροφόρηση και την ψηφιακή κουλτούρα[12]. Τα κεφάλαια αυτής της ενότητας εξετάζουν διαφορετικούς τρόπους εκπόνησης ρεπορτάζ σχετικά με τους αλγορίθμους (Ντιακόπουλος), καθώς και τον τρόπο διεξαγωγής μακροχρόνιων συνεργασιών στον τομέα αυτό (Ελμερ). Άλλα κεφάλαια εξετάζουν πώς κανείς μπορεί να δουλέψει με δεδομένα των μέσων κοινωνικής δικτύωσης, για να διερευνήσει τον τρόπο με τον οποίο οι πλατφόρμες συμμετέχουν στη διαμόρφωση της συζήτησης –από αφηγηματικές προσεγγίσεις (Βο) έως τη σχέση μεταξύ της έρευνας με ψηφιακές μεθόδους και της δημοσιογραφίας δεδομένων, και από τον τρόπο με τον οποίο τα «εξαρχής ψηφιακά» («born digital») δεδομένα μπορούν να χρησιμοποιηθούν για έρευνες σχετικά με τις υποδομές διαδικτυακής παρακολούθησης (Ρότζερς) έως ζητήματα αναφορικά με εφαρμογές και σχετιζόμενες πλατφόρμες (Βέλτεβρεντ).

Η έκτη ενότητα αφορά στην οργάνωση της δημοσιογραφίας δεδομένων και ασχολείται με διάφορους τύπους εργασίας στον τομέα, οι οποίοι θεωρούνται απαραίτητοι, αλλά δεν αναγνωρίζονται πάντα ως σημαντικοί. Αυτό συμπεριλαμβάνει την αλλαγή που έχει καταγραφεί στη δημοσιογραφία δεδομένων κατά την τελευταία δεκαετία (Ρότζερς)· τον τρόπο με τον οποίο οι πλατφόρμες και η οικονομία της περιστασιακής απασχόλησης («gig economy») διαμορφώνουν διασυνοριακά ερευνητικά δίκτυα (Κάντεα)· την εμπλοκή μεταξύ της δημοσιογραφίας δεδομένων και των κινημάτων για τα ανοιχτά δεδομένα και την τεχνολογία των πολιτών (Μπαάκ)· τις πρακτικές προγραμματισμού με ανοιχτό κώδικα (Πιτς και Μουσκάτο)· τις πρακτικές μέτρησης του κοινού (Πέτρε)· την αρχειοθέτηση της δημοσιογραφίας δεδομένων (Μπρουσάρντ)· και τον ρόλο του hashtag #ddj στη σύνδεση των κοινοτήτων δημοσιογραφίας δεδομένων στο Twitter (Ω και Σμιθ).

Η έβδομη ενότητα δίνει πληροφορίες για τη δημοσιογραφία δεδομένων ως συνεργατική διαδικασία, περιλαμβανομένων των εκπαιδευτικών προγραμμάτων δημοσιογραφίας δεδομένων και της ανάπτυξής της σε ολόκληρο τον κόσμο. Αυτό συμπεριλαμβάνει κεφάλαια σχετικά με τη διδασκαλία της δημοσιογραφίας δεδομένων στα πανεπιστήμια των Ηνωμένων Πολιτειών (Φίλιπς)· την ενδυνάμωση περιθωριοποιημένων κοινοτήτων, προκειμένου να αφηγηθούν τις ιστορίες τους (Κονσταντάρας, Βάκα)· μια προειδοποίηση ενάντια στην «ψηφιακή καθολικότητα» και την υποτίμηση της καινοτομίας στην «περιφέρεια» (Τσαν)· και διαφορετικές προσεγγίσεις για συνεργασίες μεταξύ δημοσιογράφων και ερευνητών (Ράντκλιφ και Λούις).

12. https://www.bellingcat.com/, https://www.buzzfeednews.com/topic/fake-news

Η δημοσιογραφία δεδομένων δεν λαμβάνει χώρα μέσα σε κενό. Η όγδοη και τελευταία ενότητα επικεντρώνεται στην τοποθέτηση αυτής της πρακτικής σε σχέση με τα διάφορα κοινωνικά, πολιτικά, πολιτιστικά και οικονομικά της πλαίσια. Ένα κεφάλαιο σχετικά με τη γενεαλογία της δημοσιογραφίας δεδομένων στις Ηνωμένες Πολιτείες ενθαρρύνει τον προβληματισμό σχετικά με τις διάφορες ιστορικές πρακτικές και ιδέες που τη διαμορφώνουν (Άντερσον). Άλλα κεφάλαια εξετάζουν τον τρόπο με τον οποίο αξιολογούνται τα έργα δημοσιογραφίας δεδομένων μέσω βραβείων (Λούζεν)· διαφορετικές προσεγγίσεις για τη μέτρηση του αντικτύπου των έργων δημοσιογραφίας δεδομένων (Μπράντσο, Γκριν-Μπάρμπερ)· ζητήματα γύρω από τη δημοσιογραφία δεδομένων και την αποικιοκρατία (Γιανγκ και Κάλισον)· κατά πόσο η δημοσιογραφία δεδομένων μπορεί να ανταποκριθεί στις προγενέστερες φιλοδοξίες της να μετατραπεί σε πεδίο εμπνευσμένου πειραματισμού, διαδραστικότητας και παιχνιδιού (Άσερ)· και τη σχέση της δημοσιογραφίας δεδομένων με τον ψηφιακό φιλελευθερισμό (Μπόγιερ).

Δώδεκα προκλήσεις για μια κριτική πρακτική δεδομένων

Προκειμένου να αξιοποιήσουμε τον χρόνο που περάσαμε εξερευνώντας τις πρακτικές δημοσιογραφίας δεδομένων κατά τη διάρκεια της ανάπτυξης αυτού του τόμου, θα θέλαμε να ολοκληρώσουμε αυτήν την εισαγωγή στο βιβλίο με δώδεκα προκλήσεις για την «κριτική πρακτική δεδομένων». Οι προκλήσεις αυτές εξετάζουν τη δημοσιογραφία δεδομένων σε σχέση με την ικανότητά της να διαμορφώνει σχέσεις ανάμεσα σε διαφορετικούς φορείς, καθώς και να δημιουργεί απεικονίσεις για τον κόσμο. Έχοντας δοκιμαστεί στο πλαίσιο των συνεργασιών μας για «διδασκαλία με γνώμονα την έρευνα» στο King's College του Λονδίνου και στο Public Data Lab[13], αυτές οι δώδεκα προκλήσεις προορίζονται να λειτουργήσουν ως προτροπή για επίδοξους δημοσιογράφους δεδομένων, για ομαδικές εργασίες και έρευνες φοιτητών, για συνεργασίες ερευνητών-δημοσιογράφων και για άλλες δραστηριότητες που φιλοδοξούν να οργανώσουν τη συλλογική έρευνα με δεδομένα, χωρίς να θεωρούν δεδομένες τις υποδομές, τα περιβάλλοντα και τις πρακτικές μέσω των οποίων παράγονται.

- Πώς μπορούν τα έργα δημοσιογραφίας δεδομένων να αφηγηθούν ιστορίες *τόσο με όσο και για* τα δεδομένα, συμπεριλαμβανομένων των διάφορων φορέων, διαδικασιών, θεσμών, υποδομών και μορφών γνώσης μέσω των οποίων δημιουργούνται τα δεδομένα;
- Πώς μπορούν τα έργα δημοσιογραφίας δεδομένων να αφηγηθούν ιστορίες για ζητήματα μεγάλης κλίμακας (π.χ. κλιματική αλλαγή, ανισότητα, πολυεθνική φορολογία, μετανάστευση), επιβεβαιώνοντας ταυτόχρονα την προσωρινότητα και αναγνωρίζοντας τα μοντέλα, τις υποθέσεις και την αβεβαιότητα που εμπλέκονται στην παραγωγή των αριθμών;
- Πώς μπορούν τα έργα δημοσιογραφίας δεδομένων να λάβουν υπόψη τους τον συλλογικό χαρακτήρα των ψηφιακών δεδομένων, των πλατφορμών, των αλγορίθμων και των διαδικτυακών συσκευών, συμπεριλαμβανομένης της αλληλεπίδρασης μεταξύ ψηφιακών τεχνολογιών και ψηφιακού πολιτισμού;
- Πώς μπορούν τα έργα δημοσιογραφίας δεδομένων να καλλιεργήσουν τους δικούς τους τρόπους, ώστε να είναι κατανοητά, ουσιαστικά και συναφή μέσω των δεδομένων, χωρίς απλώς να προωθούν άκριτα τους τρόπους γνώσης που «διαποτίζουν» τα δεδομένα, τα προερχόμενα από τους κυρίαρχους θεσμούς, υποδομές και πρακτικές;
- Πώς μπορούν τα έργα δημοσιογραφίας δεδομένων να αναγνωρίσουν και να πειραματιστούν με τον οπτικό πολιτισμό και τις αισθητικές από τις οποίες τροφοδοτούνται, μεταξύ άλλων μέσω συνδυασμών οπτικοποίησης δεδομένων και άλλων οπτικών υλικών;

13. https://www.kcl.ac.uk/research/engaged-research-led-teaching και http://publicdatalab.org/

- Πώς μπορούν τα έργα δημοσιογραφίας δεδομένων να δημιουργήσουν χώρο για τη συμμετοχή και την παρέμβαση του κοινού στην εξέταση των καθιερωμένων πηγών δεδομένων και στην επανεκτίμηση των ζητημάτων που ελέγχονται μέσω των δεδομένων –και του τρόπου με τον οποίο διεξάγεται αυτός ο έλεγχος;
- Πώς μπορούν οι δημοσιογράφοι δεδομένων να καλλιεργήσουν και να εδραιώσουν συνειδητά το δικό τους στυλ εργασίας με δεδομένα, το οποίο μπορεί να χρησιμοποιεί ως πηγές τομείς όπως η στατιστική, η επιστήμη των δεδομένων και η ανάλυση μέσων κοινωνικής δικτύωσης;
- Πώς μπορεί ο τομέας της δημοσιογραφίας δεδομένων να αναπτύξει πρακτικές μνήμης για την αρχειοθέτηση και τη διατήρηση του έργου του, καθώς και για την τοποθέτησή του σε σχέση με τις πρακτικές και τις κουλτούρες που χρησιμοποιεί ως πηγές;
- Πώς μπορούν τα έργα δημοσιογραφίας δεδομένων να συνεργαστούν γύρω από διακρατικά ζητήματα με τρόπους που να αποφεύγουν τη λογική της πλατφόρμας και της αποικίας και να στηρίζουν τις καινοτομίες στην περιφέρεια;
- Πώς μπορεί η δημοσιογραφία δεδομένων να υποστηρίξει περιθωριοποιημένες κοινότητες, προκειμένου αυτές να χρησιμοποιήσουν δεδομένα, ώστε να αφηγηθούν τις δικές τους ιστορίες με τους δικούς τους όρους, αντί να αφηγείται εκείνη τις ιστορίες τους για λογαριασμό τους;
- Πώς μπορούν τα έργα δημοσιογραφίας δεδομένων να αναπτύξουν τους δικούς τους εναλλακτικούς και εφευρετικούς τρόπους υπολογισμού της αξίας και του αντικτύπου τους στον κόσμο, πέρα από τις μετρήσεις των μέσων κοινωνικής δικτύωσης και τις μεθοδολογίες αντικτύπου που έχουν καθιερωθεί σε άλλους τομείς;
- Πώς θα μπορούσε η δημοσιογραφία δεδομένων να αναπτύξει ένα στυλ αντικειμενικότητας που να επιβεβαιώνει, αντί να ελαχιστοποιεί, τον δικό της ρόλο σε μια παγκόσμια παρέμβαση και στη διαμόρφωση των σχέσεων μεταξύ των διαφόρων παραγόντων της συλλογικής ζωής;

Ευχαριστίες

Ευχαριστούμε ιδιαιτέρως τις Εκδόσεις του Πανεπιστημίου του Άμστερνταμ (Amsterdam University Press, AUP) και ειδικότερα τη Μαρίς Έλιοτ, για την υποστήριξη που παρείχαν σε αυτό το πειραματικό έργο, συμπεριλαμβανομένης της δημοσίευσης μιας διαδικτυακής δοκιμαστικής έκδοσης, καθώς και της υποστήριξής τους για μια ψηφιακή έκδοση του βιβλίου με ανοιχτή πρόσβαση. Οι εκδόσεις του AUP είναι μια εύστοχη επιλογή για τον επιπλέον λόγο ότι αρκετοί από τους συντελεστές του βιβλίου συγκεντρώθηκαν σε ένα διεθνές συνέδριο για τη δημοσιογραφία δεδομένων που πραγματοποιήθηκε στο Άμστερνταμ πίσω το 2010. Η χρηματοδότηση της ανοιχτής πρόσβασης υποστηρίζεται από επιχορήγηση του Ολλανδικού Οργανισμού Επιστημονικής Έρευνας (NWO, 324-98-014), χάρη στον Ρίτσαρντ Ρότζερς του Πανεπιστημίου του Άμστερνταμ. Το όραμα για το βιβλίο γεννήθηκε μέσα από συζητήσεις με φίλους και συναδέλφους που σχετίζονται με το Public Data Lab. Ιδιαίτερα ωφεληθήκαμε από τις συζητήσεις για διάφορες πτυχές αυτού του βιβλίου με τους Αντρέας Μπέρκμπακ, Έρικ Μπόρα, Νόρτζε Μάρες, Ρίτσαρντ Ρότζερς, Τομάζο Βεντουρίνι, Έσθερ Βέλτεβρεντ, Μικέλε Μάουρι, Γκαμπριέλε Κολόμπο και Άνχελες Μπριόνες. Μας δόθηκε, επίσης, χώρος για να αναπτύξουμε τον προσανατολισμό αυτού του βιβλίου μέσω εκδηλώσεων και επισκέψεων στο Πανεπιστήμιο Columbia (σε συνεργασία με τον Μπρούνο Λατούρ), στο Πανεπιστήμιο της Ουτρέχτης, στο Πανεπιστήμιο της Καλιφόρνιας στο Μπέρκλεϊ, στο Πανεπιστήμιο του Στάνφορντ, στο Πανεπιστήμιο του Άμστερνταμ, στο Πανεπιστήμιο του Μαϊάμι, στο Πανεπιστήμιο του Άαλμποργκ στην Κοπεγχάγη, στο Sciences Po στο Παρίσι, στο Πανεπιστήμιο του Κέιμπριτζ, στο London School of Economics, στο Πανεπιστήμιο του Κάρντιφ, στο Πανεπιστήμιο του Λάνκαστερ και στο Διεθνές Φεστιβάλ Δημοσιογραφίας στην Περούτζια. Οι μεταπτυχιακοί φοιτητές που παρακολουθούσαν το μάθημα της δημοσιογραφίας δεδομένων στο King's College του Λονδίνου

μας βοήθησαν να δοκιμάσουμε την έννοια της «κριτικής πρακτικής δεδομένων», η οποία βρίσκεται στο επίκεντρο αυτού του βιβλίου.

Η μακροχρόνια ελπίδα μας να υλοποιήσουμε άλλη μία έκδοση καλλιεργήθηκε και υλοποιήθηκε χάρη στη Ρίνα Τσουμπάκι, η οποία βοήθησε να μας δοθεί υποστήριξη από το Ευρωπαϊκό Κέντρο Δημοσιογραφίας (European Journalism Centre) και το Google News Lab. Είμαστε ευγνώμονες στους Άνταμ Τόμας, Μπιάνκα Λέμενς, Μπίμπα Κλομπ, Λετίσια Γκαμπίνι, Άρνε Γκράουλς και Σάιμον Ρότζερς που μας παρείχαν, αφενός, εκδοτική ανεξαρτησία και, αφετέρου, διαρκή υποστήριξη για την κλιμάκωση των προσπαθειών μας. Η εκδοτική βοήθεια της Ντανιέλα Ντεμάρκι και της Ουάνα Μπουνεγκρού υπήρξε εξαιρετικά πολύτιμη, καθώς μας βοήθησε να χαράξουμε μια σαφή πορεία μέσα από την αχανή ροή κειμένων, υποσημειώσεων, αναφορών, ηλεκτρονικών μηνυμάτων, κοινοποιημένων εγγράφων, ιστορικού εκδόσεων, υπολογιστικών φύλλων και άλλου υλικού.

Πάνω απ' όλα, θα θέλαμε να ευχαριστήσουμε όλους τους επαγγελματίες της δημοσιογραφίας δεδομένων και τους ερευνητές οι οποίοι συμμετείχαν στο έργο (είτε γραπτώς, είτε μέσω αλληλογραφίας, είτε μέσω συζητήσεων), που μας συνόδευσαν και που υποστήριξαν αυτό το πείραμα με τη συνεισφορά τους σε χρόνο, ενέργεια, υλικό και ιδέες. Χωρίς όλα αυτά η πραγματοποίηση του έργου δεν θα ήταν δυνατή. Το βιβλίο αυτό ήταν, και συνεχίζει να είναι, ένα συλλογικό εγχείρημα.

Παραπομπές

Agre, P. E. (1997) «Toward a critical technical practice: Lessons learned in trying to reform AI», στο G. Bowker, S. L. Star, B. Turner & L. Gasser (επιμ.), *Social science, technical systems, and cooperative work: Beyond the great divide,* Νιου Τζέρσει: Erlbaum, σσ. 130-157.

Berman, E. P. & Hirschman, D. (2018) «The sociology of quantification: Where are we now?», *Contemporary Sociology*, 47(3), σσ. 257-266. DOI: https://doi.org/10.1177/0094306118767649

Byers, D. (2014) «Knives out for Nate Silver», *Politico*, 19 Μαρτίου. Διαθέσιμο σε: https://www.politico.com/blogs/media/2014/03/knives-out-for-nate-silver-185394.html (Προσπελάστηκε: 13 Νοεμβρίου 2021).

Davies, W. (2016) «The age of post-truth politics», *The New York Times*, 24 Αυγούστου. Διαθέσιμο σε: https://www.nytimes.com/2016/08/24/opinion/campaign-stops/the-age-of-post-truth-politics.html (Προσπελάστηκε: 13 Νοεμβρίου 2021).

Gray, J. (2018) «Three aspects of data worlds», *Krisis: Journal for Contemporary Philosophy*, 1. Διαθέσιμο σε: https://archive.krisis.eu/three-aspects-of-data-worlds/ (Προσπελάστηκε: 13 Νοεμβρίου 2021).

Gray, J. & Bounegru, L. (2019) «What a difference a dataset makes? *Data journalism and/as data activism*», στο Evans, J., Ruane, S. & Southall, H. (επιμ.), *Data in society: Challenging statistics in an age of globalization,* Μπρίστολ: Policy Press, σσ. 365-374. DOI: https://doi.org/10.5281/zenodo.1415450

Gray, J., Chambers, L., & Bounegru, L. (επιμ.) (2012) *The data journalism handbook: How journalists can use data to improve the news*. Σεβαστούπολη, Καλιφόρνια: O'Reilly Media.

Gray, J., Gerlitz, C., & Bounegru, L. (2018), «Data infrastructure literacy», *Big Data & Society*, 5(2), σσ. 1-13. DOI: https://doi.org/10.1177/2053951718786316

Haraway, D. J. (2016) *Staying with the trouble: Making kin in the Chthulucene.* Ντάραμ, Βόρεια Καρολίνα: Duke University Press. DOI: https://doi.org/10.1215/9780822373780

Latour, B. (2020) «Sarah Sze as the sculptor of Critical Zones», στο Latour, B. & Weibel, P. (επιμ.), Critical zones: *The science and politics of landing on earth*, Κέιμπριτζ, Μασαχουσέτη: MIT Press, σσ. 158-159.

Lury, C. & Wakeford, N. (επιμ.) (2012) *Inventive methods: The happening of the social.* Λονδίνο: Routledge.

Rogers, R. (2018) «Otherwise engaged: Social media from vanity metrics to critical analytics», *International Journal of Communication,* 12, σσ. 450-472. Διαθέσιμο σε: https://dare.uva.nl/search?identifier=e7a7c11b-b199-4d7c-a9cb-fdf1dd74d493 (Προσπελάστηκε: 13 Νοεμβρίου 2021).

Snowden, E. (2016) *Biggest leak in the history of data journalism just went live, and it's about corruption* [Twitter] 3 Απριλίου. Διαθέσιμο σε: https://twitter.com/Snowden/status/716683740903247873 (Προσπελάστηκε: 13 Νοεμβρίου 2021).

Οι συγγραφείς

Η Λιλιάνα Μπούνεγκρου (@bb_liliana) και ο Τζόναθαν Γκρέι (@jwyg) είναι Λέκτορες στο Τμήμα Ψηφιακών Ανθρωπιστικών Επιστημών στο King's College του Λονδίνου, συνιδρυτές του Public Data Lab και ερευνητικοί συνεργάτες στο Digital Methods Initiative στο Πανεπιστήμιο του Άμστερνταμ. Συνέβαλαν εξίσου ως επιμελητές αυτού του τόμου.

2.

Θέματα με Δεδομένα

2.1
Από τον Καφέ στην Αποικιοκρατία: Πώς οι Φτωχοί Ταΐζουν τους Πλούσιους

Raúl Sánchez, Ximena Villagrán

Περίληψη: Με ποιον τρόπο χρησιμοποιήσαμε δεδομένα, για να αποκαλύψουμε παράνομες επιχειρηματικές πρακτικές, παρατεταμένη περιβαλλοντική καταστροφή και συνθήκες δουλείας για τους εργαζομένους στις αγροτικές βιομηχανίες αναπτυσσόμενων χωρών.

Λέξεις-κλειδιά: διασυνοριακές έρευνες, γεωργία, αποικιοκρατία, δημοσιογραφία δεδομένων, περιβαλλοντική καταστροφή

Στις αρχές του 2016, μια μικρή ομάδα δημοσιογράφων αποφασίσαμε να ερευνήσουμε το ταξίδι μιας σοκολάτας, μιας μπανάνας και μιας κούπας καφέ από τις αρχικές τους φυτείες μέχρι τα γραφεία μας. Αφορμή για την έρευνά μας στάθηκαν ρεπορτάζ σύμφωνα με τα οποία όλα αυτά τα προϊόντα παράγονται σε φτωχές χώρες και καταναλώνονται κυρίως σε πλούσιες.

Ξεκινώντας από τα εν λόγω ρεπορτάζ, αποφασίσαμε να θέσουμε μερικά ερωτήματα: ποιες είναι οι συνθήκες εργασίας σε αυτές τις φυτείες; Παρατηρείται συγκέντρωση της ιδιοκτησίας γης σε μια μικρή πληθυσμιακή ομάδα; Τι είδους επιβάρυνση προκαλούν στο περιβάλλον αυτά τα προϊόντα στις συγκεκριμένες χώρες; Έτσι, το *El Diario* και το *El Faro* (δύο ψηφιακά και ανεξάρτητα μέσα ενημέρωσης στην Ισπανία και στο Ελ Σαλβαδόρ) ένωσαν τις δυνάμεις τους, για να διερευνήσουν τη σκοτεινή πλευρά του επιχειρηματικού μοντέλου της αγροτικής βιομηχανίας στις αναπτυσσόμενες χώρες[14].

Το έργο *Enslaved Land* προέκυψε ως αποτέλεσμα μιας διασυνοριακής έρευνας βασισμένης σε δεδομένα, που διήρκεσε έναν χρόνο. Συνοδεύεται από έναν υπότιτλο που μπαίνει απευθείας στο θέμα: «Έτσι οι φτωχές χώρες χρησιμοποιούνται, για να ταΐζουν τις πλούσιες»[15]. Στην πραγματικότητα, η αποικιοκρατία είναι το κυρίως θέμα αυτού του έργου. Ως δημοσιογράφοι, δεν θέλαμε να αφηγηθούμε την ιστορία των φτωχών ιθαγενών χωρίς να εξετάσουμε την πιο συστημική εικόνα. Θέλαμε να εξηγήσουμε πώς η ιδιοκτησία γης, η διαφθορά, το οργανωμένο έγκλημα, οι τοπικές συγκρούσεις και οι αλυσίδες εφοδιασμού ορισμένων προϊόντων εξακολουθούν να αποτελούν μέρος ενός αποικιοκρατικού συστήματος.

Σε αυτό το έργο, ερευνήσαμε πέντε καλλιέργειες οι οποίες καταναλώνονται ευρέως στην Ευρώπη και στις Ηνωμένες Πολιτείες: ζάχαρη, καφές, κακάο, μπανάνα και φοινικέλαιο που παράγονται στη Γουατεμάλα, στην Κολομβία, στην Ακτή του Ελεφαντοστού και στην Ονδούρα. Αφού η έρευνα είχε ως γνώμονα τα δεδομένα, χρησιμοποιήσαμε δεδομένα, για να φτάσουμε από το

14. https://www.eldiario.es/ (στα ισπανικά), https://elfaro.net/ (στα ισπανικά)
15. https://latierraesclava.eldiario.es/ (στα ισπανικά)

σχηματικό μοτίβο στην αληθινή ιστορία. Η επιλογή των καλλιεργειών και των χωρών έγινε με βάση προηγούμενη ανάλυση δεδομένων από 68 εκατομμύρια εγγραφές στη Βάση Δεδομένων Παγκόσμιου Εμπορίου του ΟΗΕ[16] (Εικόνα 2.1.1).

Εικόνα 2.1.1: Διάγραμμα δικτύου που αναπαριστά τις διεθνείς εισαγωγές και εξαγωγές καφέ το 2014. Πηγή: *eldiario.es*.

Αυτή η έρευνα δείχνει πώς έχει αλλάξει η ισορροπία εξουσίας ανάμεσα στις πλούσιες και στις φτωχές χώρες από τον 15ο αιώνα μέχρι σήμερα και αποδεικνύει ότι αυτές οι καλλιέργειες παράγονται χάρη σε συνθήκες εκμετάλλευσης που μοιάζουν με δουλεία για τους εργαζομένους, χάρη σε παράνομες επιχειρηματικές πρακτικές και επιφέροντας παρατεταμένη περιβαλλοντική καταστροφή.

Το επίκεντρο των ιστοριών μας διαμορφώθηκε από τα δεδομένα που χρησιμοποιήσαμε. Στην Ονδούρα, το κλειδί ήταν η χρήση γεωγραφικών πληροφοριών για την αφήγηση της ιστορίας. Συνθέσαμε τον χάρτη χρήσεων γης της χώρας και αντιπαραβάλαμε την επιφάνεια που καλύπτουν οι φυτείες φοινίκων με τις προστατευόμενες περιοχές. Διαπιστώσαμε ότι 70.000 στρέμματα φυτείας για φοινικέλαιο φυτεύτηκαν παράνομα σε προστατευόμενες περιοχές της χώρας. Με αυτόν τον τρόπο, ο δημοσιογράφος μας μπόρεσε να ερευνήσει τις συγκεκριμένες ζώνες με φυτείες φοινίκων σε προστατευόμενες περιοχές. Μέσα στο θέμα χρησιμοποιούνται μεμονωμένες περιπτώσεις για την επισήμανση και την αφήγηση της συστημικής κακοποίησης, όπως η περίπτωση του Μοντσίτο, ενός Ονδουριανού αγρότη που καλλιεργεί αφρικανικά φοινικόδεντρα στο Εθνικό Πάρκο Jeannette Kawas[17].

Αυτό το έργο δεν αφορά μόνο τη χρήση γης. Στη Γουατεμάλα, δημιουργήσαμε μια βάση δεδομένων με όλα τα εργοστάσια ζάχαρης της χώρας. Μπήκαμε στο τοπικό εμπορικό μητρώο, για να μάθουμε τα ονόματα των ιδιοκτητών και των διευθυντών των εργοστασίων. Στη συνέχεια, χρησιμοποιήσαμε τα δημόσια αρχεία επιχειρήσεων, για να συνδέσουμε αυτά τα φυσικά

16. ΣτΕ: UN Comtrade Database είναι η εναλλακτική ονομασία για την ανοιχτή και δωρεάν προσβάσιμη στατιστική βάση δεδομένων εμπορίου αγαθών του ΟΗΕ.
17. ΣτΕ: Έχει πάρει το όνομά του από τη Ζανέτ Κάουας, περιβαλλοντική ακτιβίστρια από την Ονδούρα που δολοφονήθηκε τον Φεβρουάριο 1995.

και νομικά πρόσωπα με υπεράκτιες εταιρείες στον Παναμά, στις Παρθένες Νήσους και στις Μπαχάμες. Προκειμένου να εξετάσει τον τρόπο δημιουργίας και τους όρους διαχείρισης των υπεράκτιων εταιρειών, το *El Faro* αναζήτησε στοιχεία που υπήρχαν στη βάση δεδομένων των *Panama Papers*. Οπότε, χρησιμοποιήσαμε αυτές τις πληροφορίες για να ανασυνθέσουμε τον τρόπο με τον οποίο ένα από τα μεγαλύτερα εργοστάσια της χώρας συνεργάστηκε με τη δικηγορική εταιρεία Mossack Fonseca, για να αποφύγει τους φόρους.

Μια διεθνής έρευνα, με στόχο την αποκάλυψη της διαφθοράς και των επιχειρηματικών ατασθαλιών σε φτωχές χώρες, αποτελεί πρόκληση από πολλές απόψεις. Αναγκαστήκαμε να εργαστούμε σε αγροτικές περιοχές όπου δεν υπάρχει κυβερνητική παρουσία και, στις περισσότερες περιπτώσεις, η διαδικασία του ρεπορτάζ ενείχε κάποιους κινδύνους. Ασχοληθήκαμε με χώρες όπου υπάρχει σημαντική έλλειψη διαφάνειας, απουσιάζουν τα ανοιχτά δεδομένα και, σε ορισμένες περιπτώσεις, η δημόσια διοίκηση δεν γνωρίζει ποιες πληροφορίες έχει στη διάθεσή της.

Η Ονδούρα και η Γουατεμάλα αποτέλεσαν μόνο μία πτυχή της έρευνάς μας. Περισσότερα από δέκα άτομα συνεργάστηκαν για την παραγωγή αυτού του υλικού. Όλες αυτές οι εργασίες συντονίστηκαν από τα γραφεία του *El Diario* στην Ισπανία και του *El Faro* στο Ελ Σαλβαδόρ, σε συνεργασία με δημοσιογράφους στην Κολομβία, στη Γουατεμάλα, στην Ονδούρα και στην Ακτή του Ελεφαντοστού.

Τη δουλειά αυτή ανέλαβαν όχι μόνο δημοσιογράφοι, αλλά και επιμελητές ύλης, φωτογράφοι, ντιζάινερ και προγραμματιστές, οι οποίοι συμμετείχαν στη διαδικασία ανάπτυξης και παραγωγής ενός ολοκληρωμένου διαδικτυακού προϊόντος. Το έργο αυτό δεν θα ήταν δυνατό να περατωθεί χωρίς όλους αυτούς.

Χρησιμοποιήσαμε μια ολοκληρωμένη αφήγηση για καθεμία από τις έρευνες. Για εμάς, ο τρόπος με τον οποίο οι χρήστες διαβάζουν και αλληλεπιδρούν με τα θέματα είναι εξίσου σημαντικός με την ίδια την έρευνα. Επιλέξαμε να συνδυάσουμε δορυφορικές εικόνες, φωτογραφίες, οπτικοποιήσεις δεδομένων και αφηγήσεις, επειδή θέλαμε ο αναγνώστης να κατανοήσει τη σχέση μεταξύ των προϊόντων που καταναλώνει και των αγροτών, των εταιρειών και άλλων φορέων που εμπλέκονται στην παραγωγή τους.

Αυτή η δομή μάς επέτρεψε να συνδυάσουμε τις προσωπικές ιστορίες με την ανάλυση δεδομένων σε μια συναρπαστική αφήγηση. Ένα παράδειγμα είναι η ιστορία του Τζον Πέρες, ενός Κολομβιανού αγρότη ο οποίος έχασε τη γη του, όταν αυτή καταπατήθηκε από παραστρατιωτικές ομάδες και εταιρείες μπανάνας κατά τη διάρκεια μιας ένοπλης σύγκρουσης. Για να διηγηθούμε αυτήν την ιστορία, χρησιμοποιήσαμε έναν χάρτη με δυνατότητα μεγέθυνσης που μεταφέρει τον αναγνώστη από τη φυτεία του αγρότη στον τελικό προορισμό της κολομβιανής παραγωγής μπανάνας.

Το έργο αυτό έδειξε ότι η δημοσιογραφία δεδομένων μπορεί να εμπλουτίσει τις παραδοσιακές τεχνικές ρεπορτάζ και να συνδέσει ατομικές ιστορίες με το ευρύτερο κοινωνικό, οικονομικό και πολιτικό περικείμενο.

Η έρευνά μας δημοσιεύθηκε, επίσης, από την *Plaza Pública* στη Γουατεμάλα και το *Ciper* στη Χιλή, και συμπεριλήφθηκε στη ραδιοφωνική εκπομπή της Γουατεμάλας *ConCriterio*. Αυτή η τελευταία οδήγησε σε μια δημόσια δήλωση της Φορολογικής Υπηρεσίας της Γουατεμάλας που έκανε έκκληση για πόρους, προκειμένου να καταπολεμήσει τις φορολογικές απάτες στη βιομηχανία της ζάχαρης.

Οι συγγραφείς

Ο Ραούλ Σάντσεζ είναι Ισπανός δημοσιογράφος δεδομένων που καλύπτει ιστορίες ανισότητας, φύλου και διαφθοράς με κάθε κόστος, στο elDiario.es.
Η Σιμένα Βιγιαγκράν είναι δημοσιογράφος δεδομένων από τη Γουατεμάλα, συνιδρύτρια του *El Intercambio*.

2.2

Επαναχρησιμοποίηση Απογραφικών Δεδομένων για τον Υπολογισμό του Φυλετικού Διαχωρισμού στις ΗΠΑ

Aaron Williams

Περίληψη: Οπτικοποίηση του φυλετικού διαχωρισμού στις ΗΠΑ με τη χρήση των απογραφικών δεδομένων.

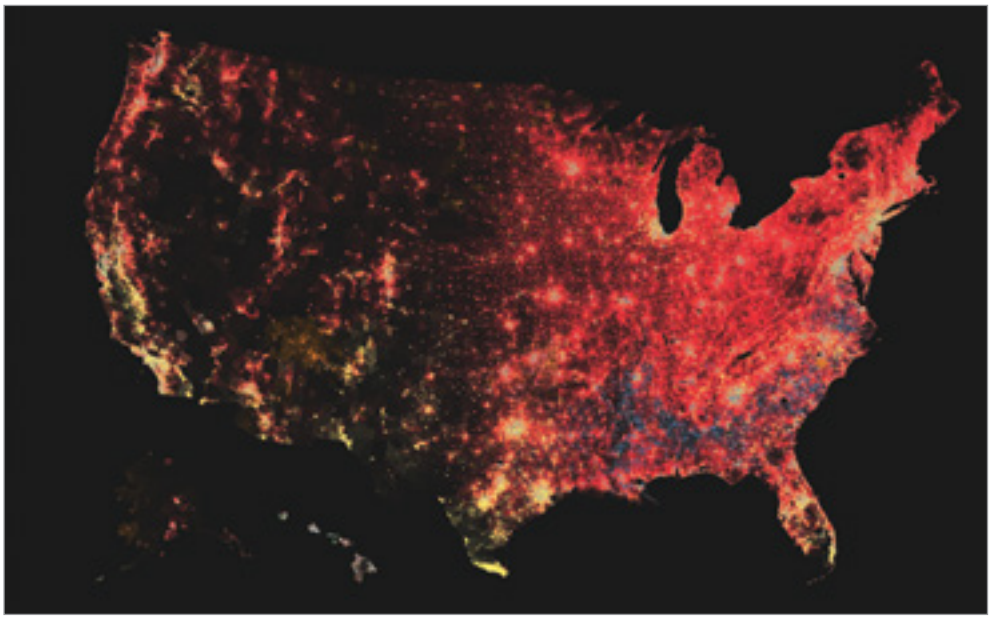
Λέξεις-κλειδιά: προγραμματισμός, χαρτογράφηση, φυλετικός διαχωρισμός, απογραφή, οπτικοποίηση δεδομένων, δημοσιογραφία δεδομένων

Πώς μπορούμε να υπολογίσουμε τον διαχωρισμό ανά φυλή; Ιδιαίτερα στις Ηνωμένες Πολιτείες, καταγράφεται ιστορικά από την ίδρυσή τους μια προσπάθεια διαχωρισμού των ανθρώπων. Καθώς η χώρα άλλαζε και ρατσιστικές πολιτικές, όπως ο διαχωρισμός, τέθηκαν εκτός νόμου, εμφανίστηκαν νέες νομοθεσίες που σκοπό είχαν να κρατήσουν τους Αφροαμερικανούς, καθώς και άλλες ομάδες, χωριστά από τους λευκούς Αμερικανούς. Πολλοί Αμερικανοί έχουν βιώσει τις παρατεταμένες επιπτώσεις αυτών των νόμων, αλλά εγώ επιθυμούσα να μάθω αν υπήρχε τρόπος να υπολογιστεί ο αντίκτυπος με βάση τον τόπο διαμονής των ανθρώπων.

Η αρχική έμπνευση μου γεννήθηκε μετά την ανάγνωση του *We Gon' Be Alright: Notes on Race and Resegregation* του Τζεφ Τσανγκ, ένα βιβλίο με δοκίμια όπου ο συγγραφέας εξερευνά θεματικές που συνδέουν τη φυλή με τον τόπο. Μου έκαναν εντύπωση τα κεφάλαια που μιλούσαν για δημογραφικές αλλαγές οι οποίες έλαβαν χώρα σε μέρη όπως το Σαν Φρανσίσκο, το Λος Άντζελες και η Νέα Υόρκη και θέλησα να δουλέψω επάνω σε ένα έργο που θα ποσοτικοποιούσε τις ιδέες για τις οποίες έγραφε ο Τσανγκ.

Πολλοί χάρτες που δείχνουν διαχωρισμό, στην πραγματικότητα, δεν τον δείχνουν. Αυτοί οι χάρτες συχνά αποτυπώνουν κάθε μέλος μιας συγκεκριμένης φυλής ή εθνότητας σε μια γεωγραφική περιοχή ως μια κουκκίδα η οποία χρωματίζεται ανάλογα με τη φυλή του ατόμου. Καταλήγουν να παρουσιάζουν συναρπαστικούς πληθυσμιακούς χάρτες για τα μέρη όπου ζουν οι άνθρωποι, αλλά δεν μετρούν πόσο ετερόκλητες ή διαχωρισμένες είναι αυτές οι περιοχές.

Πώς το γνωρίζουμε αυτό; Λοιπόν, ο διαχωρισμός και η ετερότητα είναι δύο όροι που έχουν εντελώς διαφορετικούς ορισμούς ανάλογα με τον συνομιλητή σας. Αν και πολλοί άνθρωποι μπορεί να θεωρούν ότι ζουν σε μια διαχωρισμένη περιοχή, η άποψη αυτή ενδέχεται να αλλάζει ανάλογα με τον τρόπο με τον οποίο υπολογίζει κανείς τον διαχωρισμό. Δεν ήθελα να ενεργήσω μονάχα με βάση ανεκδοτολογικά γεγονότα. Έτσι, αναζήτησα τρόπους για να υπολογίσω τον διαχωρισμό με την ακαδημαϊκή έννοια και να βασίσω το ρεπορτάζ μου εκεί.



Εικόνα 2.2.1: Χάρτης πληθυσμιακής πυκνότητας ανά φυλή στις Ηνωμένες Πολιτείες από τα στοιχεία της απογραφής το 2018. Πηγή: *The Washington Post*

Πήρα συνέντευξη από τον Μάικλ Μπέιντερ, αναπληρωτή καθηγητή κοινωνιολογίας στο American University της Ουάσιγκτον, ο οποίος μου έδειξε τον Δείκτη Εντροπίας Πολλαπλών Ομάδων («Multigroup Entropy Index» ή Δείκτης Theil), ένα στατιστικό μέτρο που προσδιορίζει τη χωρική κατανομή πολλαπλών φυλετικών ομάδων ταυτόχρονα. Το χρησιμοποιήσαμε, για να βαθμολογήσουμε κάθε ομάδα τετραγώνων απογραφής[18] στις Ηνωμένες Πολιτείες, σε σύγκριση με τον φυλετικό πληθυσμό της κομητείας στην οποία βρίσκονται.

Το έργο χρειάστηκε περίπου έναν χρόνο για να ολοκληρωθεί. Το μεγαλύτερο διάστημα που προηγήθηκε αφιερώθηκε στη διερεύνηση των δεδομένων και των ποικίλων μέτρων διαχωρισμού. Κατά τη διάρκεια της έρευνάς μου, έμαθα ότι υπάρχουν διάφοροι τρόποι μέτρησης του διαχωρισμού. Για παράδειγμα, ο Δείκτης Εντροπίας Πολλαπλών Ομάδων μετρά την ομαλότητα, συγκρίνοντας τη χωρική κατανομή των πληθυσμών εντός μιας δεδομένης γεωγραφικής περιοχής. Υπάρχουν και άλλες μετρήσεις, όπως αυτές που προκύπτουν από τον Δείκτη Έκθεσης («Exposure Index»), ο οποίος υπολογίζει πόσο πιθανό είναι να έρθουν σε επαφή δύο ομάδες μεταξύ τους στην ίδια γεωγραφική περιοχή. Δεν υπάρχει μόνο ένα μέτρο που να αποδεικνύει ή να μην αποδεικνύει τον διαχωρισμό, αλλά τα μέτρα μπορούν να συνδυαστούν, για να εξηγήσουν τον τρόπο με τον οποίο συγκροτείται μια κοινότητα. Διάβασα πολλές έρευνες σχετικά με τα δημογραφικά στοιχεία των απογραφών και προσπάθησα να αντικατοπτρίσω τις κατηγοριοποιήσεις μου στην υπάρχουσα βιβλιογραφία για το θέμα. Έτσι, επέλεξα τις έξι φυλετικές κατηγορίες που περιλαμβάνονται σε αυτό το έργο με βάση την υπάρχουσα έρευνα σχετικά με τις φυλές και τον διαχωρισμό, η οποία διενεργήθηκε από την Υπηρεσία Απογραφής των ΗΠΑ, και επέλεξα τον Δείκτη Εντροπίας Πολλαπλών Ομάδων, επειδή μου επιτρέπει να συγκρίνω πολλαπλές φυλετικές ομάδες σε μία ενιαία ανάλυση.

18. ΣτΜ: Γεωγραφική μονάδα μέτρησης που χρησιμοποιείται από τις απογραφικές υπηρεσίες των ΗΠΑ και συνήθως περιλαμβάνει πληθυσμό από 600 έως 3.000 άτομα. Είναι η μικρότερη γεωγραφική μονάδα για την οποία δημοσιεύονται δεδομένα.

Αποφάσισα να συγκρίνω τη σύνθεση κάθε ομάδας τετραγώνων απογραφής με τη φυλετική σύνθεση της αντίστοιχης κομητείας. Στη συνέχεια, ο συνάδελφός μου, Αρμάντ Ιμαμντζόμε, και εγώ περάσαμε μήνες δουλεύοντας επάνω στον μηχανισμό που τροφοδοτούσε την ανάλυση δεδομένων. Στο παρελθόν, έχω δει πολλές έρευνες σχετικά με τη δημογραφική απογραφή να διεξάγονται με εργαλεία όπως η Python, η R ή το SPSS, αλλά ήμουν περίεργος αν θα μπορούσα να κάνω αυτή τη δουλειά χρησιμοποιώντας JavaScript. Διαπίστωσα ότι η JavaScript και το περιβάλλον του node.js προσφέρουν ένα πλούσιο σύνολο εργαλείων για την επεξεργασία δεδομένων και τη μετέπειτα εμφάνισή τους στον Ιστό. Μία πρόκληση ήταν το γεγονός ότι αρκετές από τις συναρτήσεις της ανάλυσής μου έπρεπε να είναι χειρόγραφες, αλλά σε αντιστάθμισμα αυτού μπόρεσα να κατανοήσω κάθε βήμα αυτής της ανάλυσης και να χρησιμοποιήσω τις ίδιες συναρτήσεις στον Ιστό. Τόσο το Mapbox όσο και η d3.js διαθέτουν πολύ ισχυρά και ώριμα εργαλεία, για να δουλέψει κανείς με γεωχωρικά δεδομένα, τα οποία αξιοποίησα σε κάθε στάδιο της ανάλυσής μου.

Περίπου δύο μήνες πριν από τη δημοσίευση του θέματος, ταλαντευτήκαμε όσον αφορά στον σχεδιασμό και στη διάταξή του. Σε μια πρώιμη έκδοση αυτού του έργου εφαρμόσαμε την τεχνική του scrollytelling, όπου ο χάρτης καταλαμβάνει ολόκληρη την οθόνη και το κείμενο κυλά επάνω στον χάρτη, όσο οι χρήστες μετακινούνται προς τα κάτω στην οθόνη. Ενώ αυτή η τεχνική είναι καθιερωμένη και χρησιμοποιείται σε μεγάλο βαθμό από την ομάδα μου στην *Post*, εντούτοις μας εμπόδιζε να συμπεριλάβουμε, με έναν ολιστικό τρόπο, τα όμορφα στατικά γραφικά που είχαμε δημιουργήσει. Στο τέλος, επιλέξαμε μια παραδοσιακή διάταξη θέματος, που εξερευνούσε την ιστορία του διαχωρισμού και των διακρίσεων στη στέγαση στις Ηνωμένες Πολιτείες, με περιπτωσιολογικές μελέτες για τρεις πόλεις, και στη συνέχεια συμπεριλάβαμε τον πλήρη διαδραστικό χάρτη με τα ιστορικά δεδομένα στο «κλείσιμο» του θέματος[19].

Αυτό είναι το πιο πολυδιαβασμένο θέμα που έχω δημοσιεύσει ποτέ ως δημοσιογράφος. Νομίζω ότι, αφήνοντας τους αναγνώστες να εξερευνήσουν τα δεδομένα μετά το τέλος του κειμένου, δημιουργήθηκε ένα πρόσθετο επίπεδο εξατομίκευσης το οποίο επέτρεψε στους αναγνώστες να τοποθετηθούν μέσα στην αφήγηση. Η δημοσιογραφία δεδομένων μας επιτρέπει να αφηγούμαστε ιστορίες που ξεπερνούν τις λέξεις, ξεπερνούν τις ιδέες. Μπορούμε να βάλουμε τον αναγνώστη απευθείας μέσα στην ιστορία και να τον αφήσουμε να πει τη δική του.

Ο συγγραφέας

Ο Έρον Γουίλιαμς είναι ερευνητικός δημοσιογράφος δεδομένων που ειδικεύεται στην ανάλυση και οπτικοποίηση δεδομένων για την *Washington Post*.

19. https://www.washingtonpost.com/graphics/2018/national/segregation-us-cities/

2.3
Πολλαπλασιάζοντας τις Αναμνήσεις με την Ανακάλυψη Δέντρων στην Μπογκοτά

María Isabel Magaña

Περίληψη: Πώς χρησιμοποιήσαμε δεδομένα σχετικά με τα δέντρα για να δημιουργήσουμε αναμνήσεις, να προωθήσουμε τη διαφάνεια και να συμπεριλάβουμε τους πολίτες στην αφήγηση στην Μπογκοτά της Κολομβίας.

Λέξεις-κλειδιά: δημοσιογραφία δεδομένων, ιθαγένεια, διαφάνεια, ανοιχτή διακυβέρνηση, πολυτροπική αφήγηση, δέντρα

Η Μπογκοτά φιλοξενεί σχεδόν το 16% του πληθυσμού της Κολομβίας σε μόλις 1.775 τετραγωνικά χιλιόμετρα. Αντιλαμβάνεστε την κατάσταση, η πόλη είναι πολυπληθής και μαινόμενη. Όμως, είναι, επίσης, μια καταπράσινη πόλη, που περιβάλλεται από βουνά και διαθέτει πολλά είδη δέντρων. Τις περισσότερες φορές, τα δέντρα περνούν απαρατήρητα από τους κατοίκους μέσα στην καθημερινότητά τους. Ή τουλάχιστον αυτό συνέβαινε με τα μέλη της ομάδας δεδομένων μας, με εξαίρεση μία προγραμματίστριά μας που λατρεύει τα δέντρα και αποκλείεται να διασχίσει τον δρόμο χωρίς να τα παρατηρήσει. Γνωρίζει όλα τα είδη και όλα τα στοιχεία που τα αφορούν. Η αγάπη της για τη φύση μέσα στο χάος της πόλης ήταν αυτό που μας έκανε να αναλογιστούμε: έχει μιλήσει κανείς, ποτέ, για τα δέντρα που φυτεύονται σε ολόκληρη την πόλη;

Και αυτή η απλή ερώτηση ήταν ο καταλύτης για τόσες άλλες: τι γνωρίζουμε για αυτά; Ποιος είναι υπεύθυνος για τη φροντίδα τους; Βοηθούν στην εξάλειψη της ρύπανσης της πόλης; Χρειαζόμαστε περισσότερα δέντρα στην πόλη; Είναι αλήθεια ότι μόνο οι πλούσιες γειτονιές έχουν ψηλά δέντρα; Υπάρχουν ιστορικά δέντρα στην πόλη;

Ξεκινήσαμε την έρευνά μας με δυο διαφορετικούς στόχους: πρώτον, να δημιουργήσουμε μια σύνδεση ανάμεσα στους πολίτες και στους πράσινους γίγαντες που βλέπουν καθημερινά· και, δεύτερον, να κατανοήσουμε την πραγματικότητα των σχεδίων της πόλης για δεντροφύτευση και συντήρηση[20].

Για τον σκοπό αυτό, αναλύσαμε την απογραφή των αστικών δεντροφυτεύσεων στην Μπογκοτά, που διεξήγαγε ο Βοτανικός Κήπος το 2007. Αποτελεί το μοναδικό διαθέσιμο σύνολο πληροφοριών, το οποίο και ενημερώνεται κάθε μήνα. Ο Βοτανικός Κήπος αρνήθηκε να μας παραχωρήσει τα πλήρη στοιχεία, ακόμη και μετά την υποβολή πολλαπλών αιτημάτων στο πλαίσιο της ελευθερίας της πληροφόρησης, γεμάτων με νομικά επιχειρήματα. Η θέση τους ήταν ξεκάθαρη: τα δεδομένα ήδη βρίσκονταν διαθέσιμα στη δική τους πύλη οπτικοποίησης δεδομένων. Το δικό μας επιχείρημα: από εκεί μπορεί κανείς να κατεβάσει μόνο 10.000 εγγραφές και η βάση δεδομένων αποτελείται από 1,2 εκατομμύρια εγγραφές. Είναι δημόσια δεδομένα, μπορείτε απλά να μας τα δώσετε! Η απάντησή τους: δεν θα σας τα δώσουμε, αλλά θα βελτιώσουμε την εφαρμογή μας, ώστε να μπορείτε να κατεβάσετε 50.000 εγγραφές.

20. http://especiales.datasketch.co/arboles-bogota/

Η λύση μας; Επικοινωνήσαμε με άλλους οργανισμούς που είχαν βοηθήσει τον Βοτανικό Κήπο να συλλέξει τα δεδομένα. Ένας από αυτούς τους φορείς ήταν η πλατφόρμα Ideca, που συλλέγει όλες τις πληροφορίες οι οποίες σχετίζονται με το κτηματολόγιο της πόλης. Μας έδωσαν ολόκληρο το σύνολο των δεδομένων σε χρόνο μηδέν. Εμείς, προφανώς, αποφασίσαμε να το δημοσιεύσουμε, ώστε να έχουν όλοι πρόσβαση σε αυτό (πείτε πως ήταν η μικρή μας εκδίκηση απέναντι στην αδιαφάνεια). Ο Βοτανικός Κήπος το έμαθε και διέκοψε κάθε περαιτέρω συζήτηση μαζί μας, ενώ εμείς αποφασίσαμε να μην κυνηγήσουμε το θέμα νομικά.

Επιπλέον, συμπεριλάβαμε δημόσια δεδομένα από το Γραφείο Δημάρχου της Μπογκοτά και την Εθνική Απογραφή, για να διασταυρώσουμε πληροφορίες που θα μπορούσαμε να αναλύσουμε σε σχέση με τα δέντρα. Τέλος, πραγματοποιήσαμε συνεντεύξεις με ειδικούς περιβαλλοντολόγους και δασολόγους οι οποίοι μας επέτρεψαν να κατανοήσουμε τις προκλήσεις που αντιμετωπίζει η πόλη. Είχαν κάνει πολλή δουλειά και πολλές έρευνες, αναλύοντας όχι μόνο την πράξη των προγραμμάτων δεντροφύτευσης, αλλά και την ιστορία πίσω από τα δέντρα στην πόλη. Και το μεγαλύτερο τμήμα αυτού του έργου περνούσε, σε μεγάλο βαθμό, απαρατήρητο από τις Αρχές, τους δημοσιογράφους και πολλούς άλλους.

Το τελικό αποτέλεσμα ήταν ένα έργο δεδομένων οκτώ ενοτήτων που έδειχνε την πραγματικότητα των σχεδίων δεντροφύτευσης της πόλης. Χαρτογραφούσε κάθε δέντρο –με πληροφορίες για το ύψος, το είδος και τα οφέλη του για την πόλη–, κατέρριπτε πολλούς μύθους γύρω από τη φύτευση δέντρων και έλεγε τις ιστορίες ορισμένων ιστορικών δέντρων της πόλης. Χρησιμοποιήσαμε το Leaflet και το SoundCloud για τα διαδραστικά στοιχεία. Ο σχεδιασμός υλοποιήθηκε από την ταλαντούχα ομάδα προγραμματιστών μας. Χρησιμοποιήσαμε, επίσης, το StoryMapJS, για να επιτρέψουμε στους χρήστες να εξερευνήσουν τα ιστορικά δέντρα της πόλης.

Επιλέξαμε πώς και ποια κομμάτια ήταν σημαντικά για το θέμα, αφού ερευνήσαμε πολλά άλλα παρόμοια έργα και, στη συνέχεια, συνεργαστήκαμε με ντιζάινερ, για να δημιουργήσουμε μια καλή εμπειρία χρήστη. Ήταν το πρώτο μας μεγάλο έργο δεδομένων και ένα μεγάλο τμήμα του περιλάμβανε δοκιμές και λάθη, καθώς και διερευνητικές εργασίες.

Το πιο σημαντικό είναι ότι συμπεριλάβαμε στο έργο μας τους πολίτες, καλώντας τους να μας βοηθήσουν να δημιουργήσουμε έναν συλλογικό κατάλογο δέντρων και να μοιραστούν τις δικές τους ιστορίες για τα δέντρα που είχαμε χαρτογραφήσει. Το κάναμε χρησιμοποιώντας τα μέσα κοινωνικής δικτύωσης, καλώντας τους να προσθέσουν πληροφορίες σχετικά με τα είδη δέντρων σε ένα υπολογιστικό φύλλο. Οι κάτοικοι της Μπογκοτά συνεχίζουν να μας βοηθούν στον εμπλουτισμό του καταλόγου μέχρι σήμερα. Επιπλέον, κοινοποιήσαμε έναν αριθμό WhatsApp όπου οι άνθρωποι μπορούσαν να στέλνουν φωνητικές σημειώσεις με τις ιστορίες τους για τα δέντρα. Λάβαμε σχεδόν εκατό φωνητικά μηνύματα από ανθρώπους που διηγούνταν ιστορίες για τα δέντρα όπου έδωσαν το πρώτο τους φιλί, όπου έμαθαν να σκαρφαλώνουν, τα δέντρα που τους προστάτεψαν από κλέφτες ή που τους έλειψαν επειδή κόπηκαν. Αποφασίσαμε να συμπεριλάβουμε αυτά τα αρχεία ήχου ως ένα επιπλέον επίπεδο στην εφαρμογή οπτικοποίησης, έτσι ώστε οι χρήστες να μπορούν, επίσης, να γνωρίσουν τα δέντρα της πόλης μέσω ανθρώπινων ιστοριών.

Το κύριο άρθρο και το εικαστικό, στη συνέχεια, αναδημοσιεύτηκαν από μια εθνική εφημερίδα (τόσο σε έντυπη όσο και σε ηλεκτρονική μορφή) και κοινοποιήθηκαν από τις τοπικές Αρχές και πολλούς κατοίκους που ήθελαν να αφηγηθούν τις ιστορίες τους και να μετασχηματίσουν τη σχέση που έχουν άλλοι κάτοικοι με το περιβάλλον τους. Μέχρι στιγμής, οι πολίτες έχουν χρησιμοποιήσει τον χάρτη για να διερευνήσουν τη φύση της πόλης και να τεκμηριώσουν τις δικές τους έρευνες σχετικά με τα δέντρα.

Για τον οργανισμό μας, αυτό ήταν ένα από τα πιο απαιτητικά έργα που έχουμε αναπτύξει ποτέ. Όμως, ήταν και ένα από τα πιο πολύτιμα, επειδή δείχνει πώς η δημοσιογραφία δεδομένων μπορεί να είναι κάτι περισσότερο από απλοί αριθμοί: μπορεί, επίσης, να παίξει ρόλο στη δημιουργία, στη συλλογή και στην ανταλλαγή πολιτιστικών στοιχείων και αναμνήσεων, να βοηθήσει

τους ανθρώπους να παρατηρήσουν πράγματα για τους τόπους όπου ζουν (πέρα από τα γραφήματα και τα διαγράμματα), καθώς και να πολλαπλασιάσει και να αλλάξει τις σχέσεις μεταξύ ανθρώπων, φυτών και ιστοριών στους αστικούς χώρους.

Η συγγραφέας

Η Μαρία Ιζαμπέλ Μαγκάνια είναι Κολομβιανή δημοσιογράφος δεδομένων και καθηγήτρια στη Σχολή Δημοσιογραφίας του Πανεπιστημίου της Σαβάνα της Μπογκοτά. Δημιούργησε την πρώτη πλατφόρμα Ελευθερίας της Πληροφόρησης στην Κολομβία: QueremosDatos.co.

2.4

Πίσω από τους Αριθμούς: Κατεδαφίσεις Σπιτιών στην Κατεχόμενη Ανατολική Ιερουσαλήμ

Mohammed Haddad

Περίληψη: Εάν έχετε υπομονή και μπορείτε να μετρήσετε, τότε μπορείτε να κάνετε δημοσιογραφία δεδομένων.

Λέξεις-κλειδιά: ανθρώπινες ιστορίες, δημοσιογραφία δεδομένων, Μέση Ανατολή, χαρτογράφηση, κατεδαφίσεις σπιτιών

Κοιτώντας το παρακάτω γράφημα (Εικόνα 2.4.1), θα δείτε μια σειρά από σταθερές πορτοκαλί και μαύρες ράβδους που παρουσιάζουν μια μεγάλη αύξηση το 2016. Εάν ρίξετε μια πιο προσεκτική ματιά στη λεζάντα, θα καταλάβετε ότι αυτό το διάγραμμα δείχνει τον αριθμό των κτιρίων που καταστράφηκαν και των ανθρώπων που επλήγησαν από την πολιτική κατεδαφίσεων κατοικιών του Ισραήλ.

Όπως το έθεσε ο Νέιθαν Γιο, συγγραφέας του *Flowing Data*, «τα δεδομένα είναι μια αφαίρεση της πραγματικής ζωής» (2013). Ο κάθε αριθμός αντιπροσωπεύει μια οικογένεια και αφηγείται μια ιστορία.

Το *Broken Homes* είναι, έως σήμερα, το πιο ολοκληρωμένο έργο καταγραφής των κατεδαφίσεων κατοικιών στην Ανατολική Ιερουσαλήμ, μια παλαιστινιακή γειτονιά που βρίσκεται υπό ισραηλινή κατοχή εδώ και 50 χρόνια[21].

Σε στενή συνεργασία με τον OHE, το *Al Jazeera* κατέγραψε κάθε κατεδάφιση σπιτιού στην Ανατολική Ιερουσαλήμ το 2016, που αποδείχθηκε έτος-ρεκόρ, με 190 κατεστραμμένα κτίρια και περισσότερους από 1.200 Παλαιστίνιους να έχουν εκτοπιστεί ή πληγεί.

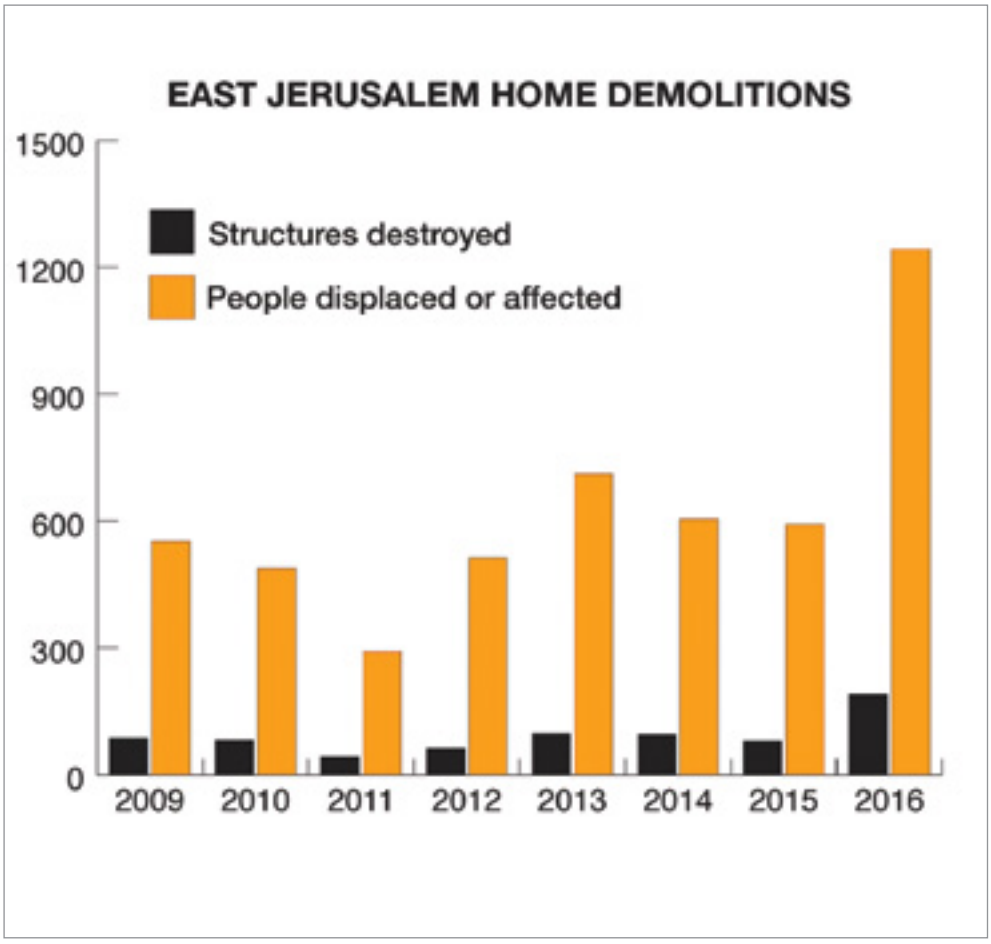
Αποφασίσαμε να καταπιαστούμε με αυτό το έργο, αφότου γίναμε μάρτυρες της κλιμάκωσης της βίας μεταξύ Ισραηλινών και Παλαιστινίων στα τέλη του 2015. Ο στόχος ήταν διττός: να κατανοήσουμε πώς η πολιτική κατεδαφίσεων κατοικιών του Ισραήλ θα επηρεαζόταν από την αύξηση των εντάσεων και να αφηγηθούμε στους αναγνώστες τις ανθρώπινες ιστορίες πίσω από τα δεδομένα.

Το έργο αποκαλύπτει τον αντίκτυπο των κατεδαφίσεων στις παλαιστινιακές οικογένειες μέσω βιντεοσκοπημένων μαρτυριών, φωτογραφιών 360 μοιρών και ενός διαδραστικού χάρτη που υποδεικνύει την τοποθεσία, τη συχνότητα και τον αντίκτυπο κάθε κατεδάφισης.

Ο παραγωγός μας στην Ντόχα ξεκίνησε μια συνεργασία με τον OHE στα τέλη του 2015 για την ανάπτυξη ενός πλαισίου για το έργο. Ο OHE συλλέγει συστηματικά δεδομένα σχετικά με τις

21. https://interactive.aljazeera.com/aje/2017/jerusalem-2016-home-demolitions/index.html

κατεδαφίσεις σπιτιών. Μολονότι κάποια από αυτά είναι διαθέσιμα στο Διαδίκτυο, κάποια άλλα –συμπεριλαμβανομένων των συντεταγμένων GPS– καταχωρούνται μόνο εσωτερικά.



Εικόνα 2.4.1: Διάγραμμα που δείχνει τις κατεδαφίσεις σπιτιών στην Ανατολική Ιερουσαλήμ, από το 2009 έως το 2006. Πηγή: *Al Jazeera*.

Θέλαμε να παρουσιάσουμε κάθε τοποθεσία κατεδάφισης σε έναν χάρτη. Ως εκ τούτου, αρχίσαμε να συγκεντρώνουμε μηνιαία σύνολα δεδομένων από τον ΟΗΕ. Για κάθε περιστατικό, συμπεριλάβαμε την ημερομηνία κατεδάφισης, τον αριθμό των ανθρώπων και των δομών που επλήγησαν, μια σύντομη περιγραφή του τι συνέβη και ένα σημείο στον χάρτη της Ανατολικής Ιερουσαλήμ που δείχνει την τοποθεσία.

Διασταυρώσαμε αυτά τα στοιχεία με τα δελτία ειδήσεων και άλλες τοπικές πληροφορίες σχετικά με κατεδαφίσεις σπιτιών. Στη συνέχεια, επιλέξαμε να προβάλλουμε μια υπόθεση κάθε μήνα, για να δείξουμε τις διαφορετικές πτυχές της ισραηλινής πολιτικής –από τιμωρητικές έως διοικητικές κατεδαφίσεις, που επηρεάζουν τους πάντες, από μικρά παιδιά έως ηλικιωμένους κατοίκους.

Ο ρεπόρτερ μας στο πεδίο ταξίδεψε σε όλη την Ανατολική Ιερουσαλήμ μέσα σε εκείνη τη χρονιά, προκειμένου να έρθει σε επαφή με πολλές από τις οικογένειες που επλήγησαν, για να εξετάσει σε μεγαλύτερο βάθος τις απώλειές τους, να φωτογραφίσει και να καταγράψει τα σημεία κατεδάφισης.

Εικόνα 2.4.2: Πανοραμική φωτογραφία σπιτιού που κατεδαφίστηκε τον Μάιο 2016. Πηγή: *Al Jazeera*.

Υπήρξε ευρύ φάσμα αποκρίσεων από οικογένειες που είχαν πληγεί. Οι συνεντεύξεις έπρεπε να πραγματοποιηθούν στον φυσικό χώρο της κατεδάφισης, μια εμπειρία δυνητικά δύσκολη για τους πληγέντες. Επομένως, απαιτήθηκε ευαισθησία και υπομονή σε όλα τα στάδια, από τη διοργάνωση των συναντήσεων μέχρι την καταγραφή του υλικού.

Συνολικά, οι οικογένειες ανταποκρίθηκαν καλά στο έργο. Ήταν πολύ γενναιόδωρες με τον χρόνο τους και την κοινοποίηση των εμπειριών τους. Σε μία περίπτωση, ένας άντρας είχε γράψει μια λίστα με πράγματα που ήθελε να μας πει. Σε κάποια άλλη περίπτωση, χρειάστηκαν μερικές προσπάθειες, για να πειστεί η οικογένεια να συμμετάσχει. Μία οικογένεια αρνήθηκε να συναντηθεί μαζί μας και, έτσι, αναγκαστήκαμε να έρθουμε σε επαφή με τον ΟΗΕ και να βρούμε κάποια άλλη οικογένεια, πρόθυμη να μιλήσει για την κατεδάφιση του σπιτιού της.

Πολλοί ειδησεογραφικοί οργανισμοί, συμπεριλαμβανομένου του *Al Jazeera*, έχουν καλύψει τις κατεδαφίσεις μεμονωμένων σπιτιών κατά τη διάρκεια των ετών. Ένας από τους κύριους λόγους που αυτή τη φορά ακολουθήσαμε μια προσέγγιση βασισμένη στα δεδομένα ήταν προκειμένου να αποδώσουμε με σαφήνεια το μέγεθος της υπόθεσης, καταμετρώντας κάθε κατεδάφιση, μία προς μία. Αυτό το πλαίσιο και η φρέσκια οπτική είναι ιδιαίτερα σημαντικά κατά την κάλυψη ενός εν εξελίξει θέματος, ώστε να μη χάνουν οι αναγνώστες την επαφή με αυτό.

Μια συμβουλή προς τους επίδοξους δημοσιογράφους δεδομένων: η προσέγγιση ενός θέματος με βάση τα δεδομένα δεν χρειάζεται να είναι ούτε τεχνικής φύσεως ούτε δαπανηρή. Μερικές φορές η απλή παρακολούθηση και η καταμέτρηση των επαναλήψεων ενός περιστατικού με την πάροδο του χρόνου λέει πολλά για την κλίμακα ενός προβλήματος. Εφόσον η μεθοδολογία συλλογής δεδομένων παραμένει συνεπής, υπάρχουν πολλές ιστορίες που μπορείτε να διηγηθείτε, χρησιμοποιώντας δεδομένα τα οποία σε άλλη περίπτωση ενδεχομένως να μη δημοσιεύατε. Επίσης, να είστε υπομονετικοί. Εμείς συγκεντρώναμε τα δεδομένα επί έναν ολόκληρο χρόνο, προκειμένου να αφηγηθούμε αυτήν την ιστορία. Το πιο σημαντικό είναι να οργανώσετε διεξοδικά τα δεδομένα που χρειάζεστε πριν να στείλετε δημοσιογράφους στο πεδίο. Και τις περισσότερες φορές δεν θα χρειαστείτε ειδικό εξοπλισμό. Εμείς χρησιμοποιήσαμε ένα iPhone για να τραβήξουμε όλες τις φωτογραφίες 360 μοιρών και για να καταγράψουμε τις ακριβείς συντεταγμένες GPS.

Το δημοσίευμα –που κυκλοφόρησε τον Ιανουάριο του 2017 στα αγγλικά, στα αραβικά και στα βοσνιακά– αποτελεί ζοφερή προειδοποίηση για το τι μέλλει γενέσθαι, καθώς το Ισραήλ συνεχίζει να αρνείται την έκδοση οικοδομικών αδειών στο 98% των Παλαιστίνιων αιτούντων, εντείνοντας την πίεση σε έναν μεγάλο και αυξανόμενο πληθυσμό.

Παραπομπές

Yau, N. (2013) «Understanding data—Context», *Big Think*, 28 Ιουνίου. Διαθέσιμο σε: https://bigthink.com/experts-corner/understanding-data-context (Προσπελάστηκε: 13 Νοεμβρίου 2021).

Ο συγγραφέας

Ο Μοχάμεντ Χαντάντ είναι δημοσιογράφος δεδομένων και διαδραστικών μέσων στο Al Jazeera Media Network με έδρα τη Μέση Ανατολή.

2.5

Χαρτογραφώντας Τροχαία Ατυχήματα για την Προώθηση της Οδικής Ασφάλειας στις Φιλιππίνες

Aika Rey

Περίληψη: Πώς μια ιστορία δεδομένων σχετικά με τα περιστατικά οδικών ατυχημάτων στην επαρχία Καγκαγιάν των Φιλιππίνων οδήγησε σε θετικά μέτρα και σε κοινωνικές αλλαγές.

Λέξεις-κλειδιά: δημοσιογραφία δεδομένων, οδική ασφάλεια, συμμετοχή της κοινότητας, χαρτογράφηση, Φιλιππίνες, πολιτική και κοινωνική αλλαγή

Τα στοιχεία δείχνουν ότι οι θάνατοι από τροχαία ατυχήματα στις Φιλιππίνες παρουσιάζουν σταθερή αύξηση με την πάροδο των ετών. Οι τραυματισμοί από τροχαία δυστυχήματα αποτελούν πλέον την κύρια αιτία θανάτου μεταξύ των νεαρών Φιλιππινέζων.

Για τον λόγο αυτό, δημιουργήσαμε ένα microsite όπου συγκεντρώσαμε σχετικές πληροφορίες για την οδική ασφάλεια. Συλλέξαμε και αναλύσαμε δεδομένα, αντλήσαμε ιδέες, δημοσιεύσαμε ιστορίες και σχεδιάσαμε ευκαιρίες συμμετοχής των πολιτών –τόσο στον φυσικό χώρο όσο και ψηφιακά– προκειμένου να ενημερώσουμε το κοινό για την οδική ασφάλεια.

Ξεκινήσαμε, επίσης, να προβάλλουμε μια σειρά βίντεο με τίτλο *Right of Way* («Προτεραιότητα στον Δρόμο»), η οποία καταπιάνεται με θέματα αυτοκινητιστών και επιβατικού κοινού στη μητροπολιτική Μανίλα. Έτσι γεννήθηκε η εκστρατεία #SaferRoadsPH του *Rappler*.

Η συλλογή σχετικών δεδομένων για τους θανάτους και τους τραυματισμούς από τροχαία ατυχήματα ήταν μια πρόκληση. Ελλείψει μιας ολοκληρωμένης εθνικής βάσης δεδομένων για τον αριθμό των τροχαίων ατυχημάτων, χτυπήσαμε πόρτες και συγκεντρώσαμε στοιχεία από περισσότερες από δώδεκα εθνικές και τοπικές κυβερνητικές υπηρεσίες, συμπεριλαμβανομένων αστυνομικών τμημάτων σε διάφορες πόλεις και επαρχίες.

Τα δεδομένα που προήλθαν από αυτά τα αποθετήρια δεν είναι τυποποιημένα. Σημαντικό μέρος της δουλειάς ήταν ο καθαρισμός των δεδομένων προς ανάλυση. Μια μεγάλη πρόκληση ήταν ο τρόπος χαρτογράφησης των δεδομένων, όταν οι πληροφορίες για την τοποθεσία είναι είτε ελλιπείς είτε δεν καταγράφονται με συνέπεια[22].

Χρησιμοποιώντας την εφαρμογή καθαρισμού δεδομένων ανοιχτού κώδικα OpenRefine, καταφέραμε να δημιουργήσουμε μια κανονικοποιημένη βάση δεδομένων με πληροφορίες που αποκτήθηκαν από τις διάφορες κυβερνητικές υπηρεσίες. Αυτό μας επέτρεψε να προσδιορίσουμε τις τοποθεσίες, τις ημερομηνίες και τον αριθμό των ατόμων που επηρεάστηκαν από τα ατυχήματα. Αν και παραμένει ελλιπής, η συλλογή μας είναι ίσως η μεγαλύτερη συλλογή δεδομένων για τα τροχαία ατυχήματα στις Φιλιππίνες αυτή τη στιγμή.

22. Δημοσιεύσαμε μια πλήρη επεξήγηση των πηγών δεδομένων μας εδώ: https://r3.rappler.com/move-ph/issues/road-safety/171778-road-crash-incidents-cagayan-valley

Ωστόσο, αυτό που έκανε την προσέγγισή μας ξεχωριστή είναι το γεγονός ότι, πέρα από τις ιστορίες, τις αναλύσεις και τις οπτικοποιήσεις που βασίστηκαν στη συλλογή των δεδομένων μας, επιδιώξαμε να παρουσιάσουμε αυτά τα δεδομένα απευθείας στις ενδιαφερόμενες κοινότητες, όχι μόνο διαδικτυακά, αλλά και με κατά τόπους δραστηριότητες. Στην πορεία, η ανάλυση δεδομένων οδήγησε σε δραστηριότητες συμμετοχής των πολιτών.

Μια συγκεκριμένη ιστορία που ξεχώρισε κατά την κάλυψη του θέματος ήταν η εις βάθος ανάλυση για την επαρχία Καγκαγιάν, η οποία βρίσκεται περίπου 600 χιλιόμετρα βόρεια της Μανίλα και είναι η περιοχή που πλήττεται περισσότερο από τα θανατηφόρα τροχαία δυστυχήματα. Επισκεφθήκαμε σημαντικούς φορείς στη συγκεκριμένη επαρχία, για να συγκεντρώσουμε δεδομένα σχετικά με τα περιστατικά τροχαίων ατυχημάτων, καθώς και για να πραγματοποιήσουμε συνεντεύξεις με θύματα, με την τοπική αστυνομία και με αξιωματούχους δημόσιων υπηρεσιών.

Μετά από αυτές τις ενέργειες, το Rappler διοργάνωσε τον Ιούνιο του 2017 ένα φόρουμ ευαισθητοποίησης για την οδική ασφάλεια στην πρωτεύουσα της επαρχίας, το Τουγκεγκαράο, για να παρουσιάσουμε τα ευρήματά μας. Το φόρουμ αποσκοπούσε στην ενημέρωση του κοινού σχετικά με θέματα οδικής ασφάλειας στην επαρχία, καθώς και στην άσκηση επιρροής σε σημαντικούς αξιωματούχους, ώστε να αντιμετωπιστεί το κενό της νομοθεσίας[23]. Πέρα από τα γραφήματα που δείχνουν τις ώρες της ημέρας κατά τις οποίες συμβαίνουν τα περισσότερα σοβαρά ατυχήματα, στο φόρουμ παρουσιάσαμε έναν θερμικό χάρτη («heatmap»), που δημιουργήθηκε με τη χρήση του Google Fusion Tables, ο οποίος δείχνει τις τοποθεσίες με τον μεγαλύτερο αριθμό περιστατικών στην επαρχία Καγκαγιάν (Εικόνα 2.5.1).

Αξιωματούχοι που ήταν παρόντες απέδωσαν τους αριθμούς αυτούς, μεταξύ άλλων, στην απουσία λωρίδων για πεζούς. Ένας έλεγχος στα σχολεία της πόλης έδειξε πως δεν υπάρχουν λωρίδες για πεζούς μπροστά από τα σχολεία. Μετά το φόρουμ, πραγματοποιήθηκε ένα κοινωνικό πείραμα όπου οι ντόπιοι σχεδίασαν με κιμωλία λωρίδες για πεζούς μπροστά από ένα σχολείο. Οι αστυνομικοί ήθελαν να διαπιστώσουν εάν οι οδηγοί θα σταματούσαν εκεί καθώς περνούσαν οι μαθητές. Το Rappler δημοσίευσε αργότερα ένα βίντεο στο Facebook σχετικά με αυτό το πείραμα[24].

Εικόνα 2.5.1: Στις περιοχές της επαρχίας Καγκαγιάν με έντονο κόκκινο χρώμα καταγράφονται περισσότερα περιστατικά τροχαίων ατυχημάτων. Πηγή: *Rappler*

23. https://www.rappler.com/move-ph/issues/road-safety/171778-road-crash-incidents-cagayan-valley
24. https://www.rappler.com/move-ph/issues/road-safety/172432-cagayan-police-pedestrian-lane-chalk

Το βίντεο προκάλεσε μεγάλο ενδιαφέρον. Ένας αναγνώστης του Rappler που είδε το βίντεο επικοινώνησε μαζί μας και προθυμοποιήθηκε να προσφέρει χρώματα, για να σχεδιαστούν λωρίδες πεζών μέσα στην πόλη. Μήνες αργότερα, χάρη στις συντονισμένες προσπάθειες της τοπικής αυτοδιοίκησης και των εθελοντών, τα σχολεία της πόλης, τελικά, απέκτησαν λωρίδες πεζών. Το βάψιμο των λωρίδων ολοκληρώθηκε στις 30 Σεπτεμβρίου 2017. Δύο χρόνια αργότερα, οι Αρχές της πόλης ενέκριναν ένα τοπικό διάταγμα για την οδική ασφάλεια.

Το έργο αυτό έδειξε ότι η ενημέρωση με βάση τα δεδομένα δεν τελειώνει απαραίτητα όταν ο συντάκτης πατήσει το κουμπί της δημοσίευσης. Αποτελεί απόδειξη ότι ο συνδυασμός τής δημοσιογραφίας δεδομένων με τη διαδικτυακή και φυσική συμμετοχή της κοινότητας μπορεί να οδηγήσει σε θετική πολιτική και κοινωνική αλλαγή.

Η συγγραφέας

Η Άικα Ρέι είναι Φιλιππινέζα δημοσιογράφος του *Rappler*, όπου έχει ασχοληθεί με έρευνες βασισμένες σε δεδομένα σχετικά με τον εθνικό προϋπολογισμό, το περιβάλλον και τις μεταφορές.

2.6
Παρακολούθηση των Εργατικών Θανάτων στην Τουρκία

Pınar Dağ

Περίληψη: Καταγραφή θανάτων εργαζομένων στην Τουρκία, με σκοπό την προώθηση της βελτίωσης των συνθηκών εργασίας.

Λέξεις-κλειδιά: Τουρκία, Σόμα, συνθήκες εργασίας, δημοσιογραφία δεδομένων, ανοιχτά δεδομένα, ελευθερία πληροφόρησης

Στον απόηχο της καταστροφής του ορυχείου Σόμα στην Τουρκία το 2014, διαπιστώθηκε ότι ήταν εξαιρετικά δύσκολο να τεκμηριωθούν οι συνθήκες εργασίας των εργαζομένων. Παρατηρήθηκαν ασυμφωνίες ως προς τα στοιχεία σχετικά με τη συνδικαλιστική οργάνωση των εργαζομένων και έλλειψη δεδομένων σχετικά με τους θανάτους εργαζομένων κατά τις προηγούμενες δεκαετίες. Το διαθέσιμο υλικό ήταν συχνά άναρχο και ελλιπές σε λεπτομέρειες. Θέλαμε να καταστήσουμε τα δεδομένα αυτά ευρέως προσβάσιμα και να ρίξουμε φως στους θανάτους εργαζομένων σε άλλους κλάδους.

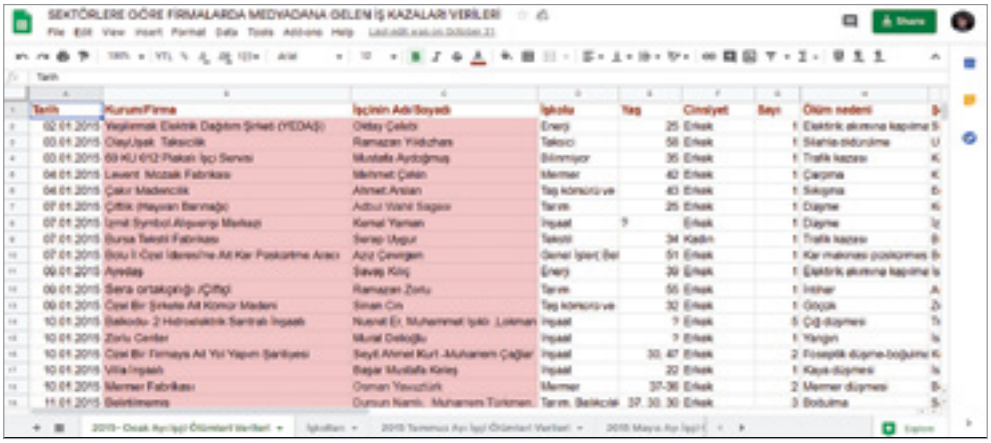
Με αυτό το σκεπτικό, ένας προγραμματιστής, ένας συντάκτης και εγώ αναπτύξαμε το έργο *Open Database of Deceased Workers in Turkey* («Ανοιχτή Βάση Δεδομένων για τους Αποβιώσαντες Εργαζομένους στην Τουρκία»), που φιλοξενείται από την πύλη δημοσιογραφίας δεδομένων *Dağ Medya* και το οποίο συγκέντρωσε δεδομένα από πολλαπλές πηγές, τα επαλήθευσε και παραχώρησε σε όλους ανοιχτή πρόσβαση και χρήση[25].

Στην Τουρκία, τουλάχιστον 130 εργαζόμενοι τον μήνα πεθαίνουν από διάφορες αιτίες. Ο σημαντικότερος στόχος του έργου ήταν η ενημέρωση σχετικά με τους θανάτους αυτούς και τη συχνότητά τους, καθώς και η δημόσια αναγνώριση των θυμάτων και των άσχημων συνθηκών εργασίας που αυτά αντιμετώπισαν.

Το έργο περιλαμβάνει χάρτες με δυνατότητα ενσωμάτωσης, γραφήματα και δεδομένα σε διάφορα φορμάτ[26]. Καλύπτει τους θανάτους εργαζομένων σε περισσότερους από 20 κλάδους από το 2011 έως το 2014. Μετά την ολοκλήρωση του έργου, συνεχίσαμε να κοινοποιούμε τους θανάτους των εργαζομένων μέσω τακτικής παρακολούθησης των μέσων ενημέρωσης κάθε μήνα. Το σημαντικότερο είναι ότι η βάση δεδομένων περιλαμβάνει τα ονόματα των εταιρειών που τους απασχολούσαν (Εικόνα 2.6.1).

Το έργο ξεκίνησε το 2015. Αρχίσαμε υποβάλλοντας αιτήματα στο πλαίσιο της ελευθερίας της πληροφόρησης («freedom of information», FOI) και συλλέγοντας δεδομένα από αξιόπιστες ΜΚΟ που εξήγαγαν δεδομένα από διάφορες πηγές και τα διέθεταν στο κοινό.

25. http://community.globaleditorsnetwork.org/content/open-database-deceased-workers-turkey-0
26. http://platform24.org/veri-gazeteciligi/451/turkiyede-isci-olumleri-veritabani-hazirlandi (στα τουρκικά)



Εικόνα 2.6.1: Συνεργατικό υπολογιστικό φύλλο με ονόματα εταιρειών, με βάση την αποδελτίωση των μέσων από το Google Alerts. Πηγή: Pinar Dağ.

Η πρώτη πρόκληση που αντιμετωπίσαμε ήταν η δυσκολία να λάβουμε ανοιχτά δεδομένα μέσω αιτημάτων FOI. Κάποιες φορές χρειάστηκαν δύο εβδομάδες και άλλες φορές χρειάστηκαν τέσσερις μήνες, για να λάβουμε δεδομένα κατόπιν των αιτημάτων μας. Στη συνέχεια, προέκυψε μια πιο απροσδόκητη πρόκληση, την οποία αναφέρω, επειδή φέρνει στην επιφάνεια αντικρουόμενες οπτικές για αυτό το είδος εργασίας. Όταν ανακοινώσαμε το εγχείρημα, ένας από τους φορείς του οποίου τα δεδομένα χρησιμοποιούσαμε –το İSİG Meclisi (Παρατηρητήριο Εργασίας για την Υγεία και την Ασφάλεια)– δυσαρεστήθηκε με το γεγονός[27]. Ο λόγος που επικαλέστηκαν ήταν ότι το έργο μας απλώς αναδημοσίευσε δεδομένα που είχαν συγκεντρώσει οι ίδιοι. Αντιλαμβάνονταν την κατ' αυτόν τον τρόπο χρήση των δεδομένων τους από εμάς ως εκμετάλλευση της εργασίας τους. Οι αντιρρήσεις για το έργο εξακολούθησαν να υφίστανται, παρά το γεγονός ότι ζητήσαμε άδεια και ότι τα δεδομένα τους είναι δημοσίως διαθέσιμα. Ενώ μας κατηγόρησαν ότι «πορνοποιούσαμε» τους θανάτους των εργαζομένων με τις οπτικοποιήσεις και τους πίνακές μας, εμείς θεωρήσαμε ότι το έργο μας προσέφερε δημόσια προστιθέμενη αξία, αυξάνοντας την εμβέλεια των δεδομένων που εκείνοι είχαν συλλέξει, μέσω οπτικά προσβάσιμων και άμεσα ανακτήσιμων μορφών.

Αν και οι ανθρώπινες ιστορίες είναι ζωτικής σημασίας, πιστεύαμε ότι τα αδόμητα, ανεπεξέργαστα δεδομένα ήταν σημαντικά, προκειμένου να προσφέρουμε μια πιο συστηματική εικόνα αυτών των αδικοπραγιών. Δυσκολευτήκαμε να φτάσουμε σε συναίνεση γύρω από αυτή τη λογική και αγωνιστήκαμε να υποστηρίξουμε την αξία των συνεργατικών πρακτικών ανοιχτού διαμοιρασμού δεδομένων. Καταλήξαμε να δημοσιεύουμε μηνιαία στοιχεία για τους θανάτους των εργαζομένων, συγκρίνοντας τα επίσημα στοιχεία που συγκεντρώσαμε μέσω αιτημάτων FOI με τα στοιχεία που συλλέξαμε από τη δική μας παρακολούθηση αυτών των περιπτώσεων.

Μετά από αυτό το έργο, οι φορείς στους οποίους είχαμε απευθυνθεί με αιτήματα FOI άρχισαν να διαθέτουν τα δεδομένα τους με πιο δομημένο τρόπο και να τα οπτικοποιούν. Θεωρήσαμε πως αυτό αποτελεί απόδειξη ότι πετύχαμε έναν από τους στόχους μας: να καταστήσουμε τις πληροφορίες αυτές ευρύτερα προσβάσιμες. Τελικά, το έργο έφτασε στους φιναλίστ των Βραβείων Δημοσιογραφίας Δεδομένων το 2015.

Η συγγραφέας

Η Πινάρ Νταγκ είναι Λέκτορας στο Τμήμα Νέων Μέσων του Πανεπιστημίου Kadir Has και συνιδρύτρια του Data Literacy Association (DLA).

27. http://www.isigmeclisi.org/ (στα τουρκικά)

3.

Συγκεντρώνοντας τα Δεδομένα

3.1

Δημιουργώντας το Δικό σας Σύνολο Δεδομένων: Καταγραφή των Εγκλημάτων με Μαχαίρια στο Ηνωμένο Βασίλειο

Caelainn Barr

Περίληψη: Δημιουργία συνόλων δεδομένων για έρευνες και μια δυνατή αφήγηση.

Λέξεις-κλειδιά: δημοσιογραφία δεδομένων, έγκλημα, λογοδοσία, φυλή, Ηνωμένο Βασίλειο, βάσεις δεδομένων

Στις αρχές του 2017, δύο συνάδελφοι, οι Γκάρι Γιάνγκ και Ντέμιεν Γκέιλ, με πλησίασαν στην αίθουσα σύνταξης του *Guardian*. Ήθελαν να ερευνήσουν τα εγκλήματα με μαχαίρια στο Ηνωμένο Βασίλειο. Μολονότι δεν έλειπαν τα δημοσιεύματα που περιέγραφαν λεπτομερώς τους θανάτους των θυμάτων από εγκλήματα με μαχαίρια, η παρακολούθηση της καταδίωξης των υπόπτων και οι εκθέσεις για τις δίκες και τις καταδίκες των δραστών, κανένας δεν είχε εξετάσει όλες τις ανθρωποκτονίες στο σύνολό τους.

Η πρώτη μου ερώτηση ήταν πόσα παιδιά και έφηβοι είχαν σκοτωθεί από μαχαίρια τα τελευταία χρόνια. Το ερώτημα έμοιαζε απλό, αλλά όταν βάλθηκα να βρω τα δεδομένα, σύντομα έκανα μια διαπίστωση: κανένας δεν ήξερε να μου πει. Τα δεδομένα υπήρχαν κάπου αλλά όχι στον δημόσιο τομέα. Σε αυτό το στάδιο είχα δύο επιλογές: είτε να τα παρατήσω είτε να φτιάξω ένα σύνολο δεδομένων από το μηδέν, χρησιμοποιώντας στοιχεία στα οποία είχα πρόσβαση, δεδομένα που μπορούσα να δημιουργήσω και να επαληθεύσω η ίδια. Αποφάσισα να φτιάξω το δικό μου σύνολο δεδομένων.

Γιατί να δημιουργήσετε ένα δικό σας σύνολο δεδομένων;

Η δημοσιογραφία δεδομένων δεν χρειάζεται να βασίζεται αποκλειστικά σε προϋπάρχοντα σύνολα δεδομένων. Στην πραγματικότητα, υπάρχουν πολλοί πολύ καλοί λόγοι για να δημιουργήσετε τα δικά σας δεδομένα. Υπάρχει μεγάλος πλούτος πληροφοριών σε δεδομένα που δεν δημοσιεύονται τακτικά ή, σε ορισμένες περιπτώσεις, δεν συλλέγονται καν.

Κατά τη δημιουργία του δικού σας συνόλου δεδομένων, δημιουργείτε ένα μοναδικό σύνολο πληροφοριών, μια μοναδική πηγή, με την οποία μπορείτε να διερευνήσετε την ιστορία σας. Τα δεδομένα και οι σχετικές ιστορίες πιθανότατα θα είναι αποκλειστικές και μπορεί να σας δώσουν ένα ανταγωνιστικό πλεονέκτημα, για να βρείτε ιστορίες στις οποίες άλλοι δημοσιογράφοι απλά δεν έχουν πρόσβαση. Το να έχετε στην κατοχή σας ένα μοναδικό σύνολο δεδομένων, επίσης, μπορεί να σας βοηθήσει να εντοπίσετε τις τάσεις που οι ειδικοί και οι υπεύθυνοι χάραξης πολιτικής δεν έχουν καταφέρει να διακρίνουν.

Τα δεδομένα είναι πηγή πληροφοριών στη δημοσιογραφία. Η βασική αρχή για τη χρήση δεδομένων στη δημοσιογραφία είναι η δομημένη σκέψη. Προκειμένου να αξιοποιηθούν τα δεδο-

μένα στο έπακρο, ο δημοσιογράφος πρέπει να σκεφτεί δομικά από την αρχή ενός έργου: ποια είναι η ιστορία που θέλω να καταφέρω να πω και τι χρειάζομαι για να γίνει αυτό;

Το κλειδί για την επιτυχή δημιουργία ενός συνόλου δεδομένων για την ιστορία σας είναι να έχετε μια δομημένη προσέγγιση του θέματος και να εξετάζετε κάθε πηγή δεδομένων με μια αίσθηση δημοσιογραφικής περιέργειας.

Η δημιουργία του δικού σας συνόλου δεδομένων περιλαμβάνει πολλές από τις απαιτούμενες δεξιότητες της δημοσιογραφίας δεδομένων –δομημένη σκέψη, σχεδιασμένη αφήγηση και εύρεση δεδομένων με δημιουργικούς τρόπους. Επίσης, δεν είναι ιδιαίτερα δύσκολο να ξεκινήσει κανείς να ασχολείται με τη δημοσιογραφία δεδομένων, καθώς αυτό μπορεί να γίνει με ή χωρίς δεξιότητες προγραμματισμού. Εάν μπορείτε να πληκτρολογήσετε σε ένα υπολογιστικό φύλλο και να ταξινομήσετε έναν πίνακα, είστε σε καλό δρόμο για να καλλιεργήσετε τις βασικές δεξιότητες που απαιτούνται για τη δημοσιογραφία δεδομένων.

Αυτό δεν σημαίνει ότι η δημοσιογραφία δεδομένων είναι απλή διαδικασία. Τα αξιόπιστα και λεπτομερή έργα δεδομένων μπορεί να συνιστούν μια πολύ σύνθετη και χρονοβόρα εργασία, αλλά με μερικές βασικές δεξιότητες μπορείτε να αποκτήσετε ισχυρά θεμέλια στη χρήση δεδομένων για την αφήγηση ιστοριών.

Δημιουργώντας το δικό σας σύνολο δεδομένων βήμα προς βήμα

Οργανώστε ό,τι απαιτείται. Το πρώτο βήμα για τη δημιουργία ή τη συλλογή δεδομένων για την ανάλυσή σας είναι να εκτιμήσετε τι απαιτείται και κατά πόσο μπορεί να αποκτηθεί. Στην αρχή κάθε έργου αξίζει να κρατήσετε ένα σημείωμα, το οποίο να περιγράφει τι περιμένετε να πει η ιστορία, πού πιστεύετε ότι βρίσκονται τα δεδομένα, πόσος χρόνος θα χρειαστεί για να τα εντοπίσετε και ποιες είναι οι πιθανές παγίδες. Το σημείωμα θα σας βοηθήσει να εκτιμήσετε πόσο χρόνο θα διαρκέσει η εργασία και αν το αποτέλεσμα αξίζει τον κόπο. Μπορεί, επίσης, να χρησιμεύσει ως κάτι στο οποίο μπορείτε να επιστρέψετε σε δεύτερο χρόνο, όταν θα βρίσκεστε στα μέσα του έργου.

Σκεφτείτε τη βασική γραμμή. Στην αρχή μιας ιστορίας δεδομένων, όταν τα δεδομένα δεν υπάρχουν, θα πρέπει να αναρωτηθείτε ποια είναι η βασική σας γραμμή. Είναι σημαντικό να γνωρίζετε τι πρέπει να περιλαμβάνουν τα δεδομένα, καθώς αυτό θέτει τις παραμέτρους για τα ερωτήματα που μπορείτε να θέσετε στα δεδομένα. Αυτό είναι ουσιώδες, καθώς τα δεδομένα απαντούν μόνο σε ερωτήματα σχετικά με τα στοιχεία που περιλαμβάνουν. Επομένως, για να δημιουργήσετε ένα σύνολο δεδομένων που θα ικανοποιεί τις ανάγκες σας, πρέπει να είστε πολύ σαφείς σχετικά με το τι θέλετε να μπορέσετε να διερευνήσετε και ποιες πληροφορίες χρειάζεστε για να το κάνετε.

Πού βρίσκονται τα δεδομένα; Το επόμενο βήμα είναι να σκεφτείτε πού μπορεί να βρείτε τα δεδομένα σε οποιαδήποτε μορφή ή σύνθεση. Ένας τρόπος για να το κάνετε αυτό είναι να αναλογιστείτε τη μέχρι τώρα πορεία σας. Πώς καταλαβαίνετε ότι μπορεί να υπάρχει κάποια ενδιαφέρουσα ιστορία εδώ; Από πού έρχεται η ιδέα και υπάρχει, άραγε, μια πιθανή πηγή δεδομένων πίσω από αυτήν;

Η έρευνα θα σας βοηθήσει να αποσαφηνίσετε τι είναι διαθέσιμο. Για τον λόγο αυτό, χτενίστε τις πηγές πληροφοριών σχετικά με το θέμα που σας ενδιαφέρει, και μιλήστε με ακαδημαϊκούς, ερευνητές και στατιστικολόγους οι οποίοι συλλέγουν τα δεδομένα ή δουλεύουν με αυτά. Αυτό θα σας βοηθήσει να εντοπίσετε ελλείψεις και πιθανές κακοτοπιές στη χρήση των δεδομένων. Θα πρέπει να σας δώσει ιδέες για τις πηγές και τρόπους απόκτησης των δεδομένων. Όλη αυτή η προετοιμασία, πριν να αρχίσετε να δημιουργείτε το σύνολο των δεδομένων σας θα είναι πολύτιμη, αν χρειαστεί να συνεργαστείτε με δύσκολες κυβερνητικές υπηρεσίες ή αν αποφασίσετε να ακολουθήσετε μια άλλη προσέγγιση για τη συλλογή των δεδομένων.

Ηθικοί προβληματισμοί. Κατά τον σχεδιασμό και την εύρεση πηγών για οποιαδήποτε ιστορία, οφείλουμε να σταθμίσουμε τους ηθικούς προβληματισμούς, και η εργασία με δεδομένα δεν

διαφέρει σε αυτό. Όταν δημιουργούμε ένα σύνολο δεδομένων, πρέπει να αναλογιζόμαστε αν η πηγή και η μέθοδος που χρησιμοποιούμε για τη συλλογή των πληροφοριών είναι η ακριβέστερη και η πληρέστερη δυνατή.

Αυτό ισχύει και για την ανάλυση –εξετάστε τις πληροφορίες από πολλαπλές οπτικές γωνίες και μη «βασανίζετε» τα δεδομένα, για να τα εξαναγκάσετε να πουν κάτι που δεν αντικατοπτρίζει επακριβώς την πραγματικότητα. Κατά την παρουσίαση του θέματος, φροντίστε να είστε διαφανείς σχετικά με την προέλευση, την ανάλυση και τους περιορισμούς των δεδομένων. Όλοι αυτοί οι προβληματισμοί θα σας βοηθήσουν να χτίσετε μια πιο δυνατή ιστορία και να καλλιεργήσετε την εμπιστοσύνη του αναγνώστη.

Συλλέξτε τα δεδομένα. Μόλις εντοπιστεί μια πιθανή πηγή, το επόμενο βήμα είναι η συλλογή των δεδομένων. Αυτό μπορεί να γίνει χειροκίνητα μέσω της εισαγωγής δεδομένων σε ένα υπολογιστικό φύλλο· με τη μετατροπή πληροφοριών που εσωκλείονται σε αρχεία .pdf σε δομημένα δεδομένα τα οποία μπορείτε να αναλύσετε· με την προμήθεια εγγράφων μέσω μιας ανθρώπινης πηγής ή του νόμου περί ελευθερίας της πληροφόρησης («Freedom of Information Act», FOIA)· με τη δημιουργία προγραμμάτων για τη συλλογή δεδομένων από έγγραφα ή από ιστοσελίδες· ή με αυτοματοποιημένη συλλογή δεδομένων μέσω διεπαφής προγραμματισμού εφαρμογών («Application Programming Interface», API).

Να είστε ευγενικοί με τον εαυτό σας! Μη θυσιάζετε την απλότητα χωρίς να υπάρχει λόγος. Προσπαθήστε να βρείτε τον πιο εύκολο τρόπο, για να μεταφέρετε τις πληροφορίες σε ένα σύνολο δεδομένων που θα μπορέσετε να αναλύσετε. Εάν είναι δυνατόν, φροντίστε η διαδικασία να μπορεί να επαναληφθεί εύκολα, καθώς αυτό θα σας βοηθήσει να ελέγξετε τη δουλειά σας και να προσθέσετε στοιχεία στο σύνολο δεδομένων σε μεταγενέστερο στάδιο, εάν χρειαστεί.

Κατά την απόκτηση των δεδομένων, ανατρέξτε στο προσχέδιο της ιστορίας σας και αναρωτηθείτε αν τα δεδομένα αυτά θα σας επιτρέψουν να διερευνήσετε πλήρως το θέμα. Περιέχουν τις πληροφορίες που μπορεί να οδηγήσουν στη βασική γραμμή που σας ενδιαφέρει;

Δομή. Η βασική διαφορά ανάμεσα στις πληροφορίες που περιέχονται στα κείμενα μιας στοίβας χάρτινων εγγράφων και σε ένα σύνολο δεδομένων είναι η δομή. Η δομή και η επανάληψη είναι απαραίτητα στοιχεία για τη δημιουργία ενός καθαρού συνόλου δεδομένων, έτοιμου για ανάλυση.

Το πρώτο βήμα είναι να εξοικειωθείτε με τις πληροφορίες. Αναρωτηθείτε τι περιέχει το υλικό –τι θα σας επιτρέψει να πείτε; Τι δεν θα μπορείτε να πείτε με αυτά τα δεδομένα; Υπάρχει ένα άλλο σύνολο δεδομένων με το οποίο ίσως θέλετε να συνδυάσετε τις πληροφορίες; Μπορείτε να οργανώσετε τη δημιουργία αυτού του συνόλου δεδομένων με τέτοιο τρόπο, ώστε να μπορέσετε να το συνδυάσετε με άλλα σύνολα δεδομένων;

Σκεφτείτε πώς θα πρέπει να μοιάζει το σύνολο δεδομένων στο τέλος της διαδικασίας. Αναλογιστείτε τις στήλες ή τις μεταβλητές που θέλετε να είστε σε θέση να αναλύσετε. Αναζητήστε έμπνευση στη μεθοδολογία και στη δομή που διέπει άλλα παρόμοια σύνολα δεδομένων.

Καταρχάς φροντίστε να διευρύνετε τις επιλογές σας, λαμβάνοντας υπόψη όλα τα δεδομένα που θα μπορούσατε να συγκεντρώσετε. Στη συνέχεια, περιορίστε τις επιλογές σας, υπολογίζοντας τι είναι απαραίτητο για την ιστορία και πόσος χρόνος θα χρειαστεί για να το βρείτε. Βεβαιωθείτε ότι τα δεδομένα που συλλέγετε είναι συγκρίσιμα. Επιλέξτε ένα φορμάτ και μείνετε σταθεροί σε αυτό –είναι κάτι που θα σας εξοικονομήσει χρόνο στο τέλος! Επίσης, σκεφτείτε τις διαστάσεις του συνόλου δεδομένων που δημιουργείτε. Οι ώρες και οι ημερομηνίες θα σας επιτρέψουν να αναλύσετε τις πληροφορίες σε βάθος χρόνου· οι γεωγραφικές πληροφορίες θα σας επιτρέψουν να χαρτογραφήσετε τα δεδομένα, για να αναζητήσετε χωροταξικές τάσεις.

Να παρακολουθείτε τη δουλειά σας και να την ελέγχετε καθώς προχωράτε. Κρατήστε σημειώσεις σχετικά με τις πηγές που χρησιμοποιήσατε για τη δημιουργία του συνόλου δεδομένων

σας και να κρατάτε πάντα αντίγραφο των πρωτότυπων εγγράφων και συνόλων δεδομένων. Συντάξτε μια μεθοδολογία και ένα «λεξικό» δεδομένων, για να παρακολουθείτε τις πηγές σας, τον τρόπο επεξεργασίας των δεδομένων και το περιεχόμενο κάθε στήλης. Αυτό θα σας βοηθήσει να επισημάνετε ερωτήματα και να ξεδιαλύνετε πιθανά σφάλματα καθώς συγκεντρώνετε και αρχίζετε να αναλύετε τα δεδομένα.

Να μη θεωρείτε τίποτα αυτονόητο και να επιβεβαιώνετε όλα τα ευρήματά σας με περαιτέρω έρευνα. Μη διστάσετε να μιλήσετε με ειδικούς και στατιστικολόγους, για να ελέγξετε την προσέγγιση και τα ευρήματά σας. Το βάρος της αξιοπιστίας της εργασίας σας είναι ακόμη μεγαλύτερο, όταν έχετε συγκεντρώσει οι ίδιοι τα δεδομένα. Οπότε, λάβετε όλα τα μέτρα για να διασφαλίσετε ότι τα δεδομένα, η ανάλυση και η συγγραφή του κειμένου είναι σωστά.

Περιπτωσιολογική μελέτη: Πέρα από τη λεπίδα

Στις αρχές του 2017 η ομάδα έργων δεδομένων μαζί με τον Γκάρι Γιάνγκ, τον Ντέμιεν Γκέιλ και την ομάδα κοινοτικής δημοσιογραφίας («community journalism»)[28] του *Guardian* ξεκίνησαν να καταγράφουν τους θανάτους κάθε παιδιού και εφήβου που έχασε τη ζωή του από μαχαίρι στο Ηνωμένο Βασίλειο.

Προκειμένου να κατανοήσει πραγματικά το ζήτημα και να διερευνήσει τα βασικά θέματα γύρω από την εγκληματικότητα με μαχαίρια, η ομάδα χρειαζόταν δεδομένα. Θέλαμε να μάθουμε ποιοι είναι οι νέοι που πεθαίνουν στο Ηνωμένο Βασίλειο ως αποτέλεσμα μαχαιρωμάτων. Είναι μικρά παιδιά ή έφηβοι; Τι γίνεται με το φύλο και την εθνικότητα; Πού και πότε σκοτώνονται αυτοί οι νέοι άνθρωποι;

Μετά από συνομιλίες με στατιστικολόγους, αστυνομικούς και εγκληματολόγους κατέστη σαφές ότι τα δεδομένα υπήρχαν, αλλά δεν ήταν δημόσια. Μεγάλο μέρος του εργασιακού μου χρόνου κατά τη διάρκεια του επόμενου έτους θα αναλωνόταν στην προσπάθειά μου να συνθέσω μια απάντηση στο ερώτημα αυτό.

Τα δεδομένα που χρειαζόμουν βρίσκονταν στην κατοχή του Υπουργείου Εσωτερικών, σε ένα σύνολο δεδομένων που ονομάζεται Δείκτης Ανθρωποκτονιών. Τα στοιχεία κοινοποιούνταν στο Υπουργείο Εσωτερικών από τις αστυνομικές δυνάμεις της Αγγλίας και της Ουαλίας. Υπήρχαν δύο πιθανές οδοί για να πάρω τις πληροφορίες –να στείλω αίτημα στο πλαίσιο της ελευθερίας της πληροφόρησης («freedom of information», FOI) στο Υπουργείο Εσωτερικών ή να στείλω αιτήματα σε κάθε αστυνομική αρχή. Για να καλύψω όλα τα ενδεχόμενα, τα έκανα και τα δύο. Αυτό θα μας έδινε τα ιστορικά στοιχεία από το 1977.

Προκειμένου να παρακολουθήσουμε τους θανάτους κατά το τρέχον έτος, έπρεπε να αρχίσουμε να τους καταμετρούμε καθώς συνέβαιναν. Δεδομένου ότι δεν υπήρχαν δημόσια ή κεντρικά συγκεντρωμένα δεδομένα, αποφασίσαμε να παρακολουθούμε τα στοιχεία οι ίδιοι, μέσω αστυνομικών αναφορών, αποδελτίωσης ειδήσεων, Google Alerts, Facebook και Twitter.

Συζητήσαμε μεταξύ μας ποια ήταν τα στοιχεία που θέλαμε να μάθουμε –το όνομα, η ηλικία και η ημερομηνία του περιστατικού ήταν πράγματα που έπρεπε οπωσδήποτε να καταγράψουμε. Όμως, άλλες πτυχές των συνθηκών των θανάτων δεν ήταν τόσο προφανείς. Συζητήσαμε τι πιστεύαμε ότι γνωρίζαμε ήδη για τα εγκλήματα με μαχαίρια –αφορούσαν κυρίως άντρες, με δυσανάλογο αριθμό μαύρων αντρών θυμάτων. Για να ελέγξουμε τις υποθέσεις μας, προσθέσαμε στο σύνολο δεδομένων στήλες για το φύλο και την εθνικότητα. Επαληθεύσαμε όλα τα στοιχεία, συγκρίνοντας τις λεπτομέρειες με τις αστυνομικές αρχές σε ολόκληρο το Ηνωμένο Βασίλειο.

28. ΣτΜ: Η κοινοτική δημοσιογραφία προσανατολίζει τις δραστηριότητές της στις τοπικές κοινότητες, γειτονιές, σε μεμονωμένα προάστια ή σε μικρές πόλεις. Εάν καλύπτει ευρύτερα θέματα, επικεντρώνεται στην επίδραση που έχουν αυτά στους αναγνώστες μιας τοπικής κοινότητας.

Με αυτόν τον τρόπο, αποκαλύφθηκαν κάποιες περιπτώσεις που δεν είχαμε εντοπίσει και, έτσι, μπορέσαμε να διασταυρώσουμε τα ευρήματά μας πριν από τη δημοσίευση.

Ύστερα από μια σειρά απορριφθέντων αιτημάτων FOI και παρατεταμένων καθυστερήσεων, τα δεδομένα τελικά δόθηκαν στη δημοσιότητα από το Υπουργείο Εσωτερικών. Ανέφεραν την ηλικία, την εθνικότητα και το φύλο όλων των ανθρώπων που σκοτώθηκαν από μαχαίρια ανά περιοχή αστυνομικής δύναμης επί σχεδόν 40 χρόνια. Αυτό, σε συνδυασμό με το δικό μας τρέχον σύνολο δεδομένων, μας επέτρεψε να εξετάσουμε ποιοι σκοτώνονταν και ποιες ήταν οι τάσεις με την πάροδο του χρόνου.

Τα δεδομένα αποκάλυψαν ότι 39 παιδιά και έφηβοι έχασαν τη ζωή τους από εγκλήματα με μαχαίρια στην Αγγλία και στην Ουαλία το 2017, καθιστώντας το έτος αυτό μία από τις χειρότερες χρονιές όσον αφορά στους θανάτους νέων ανθρώπων εδώ και σχεδόν μία δεκαετία. Τα στοιχεία προκάλεσαν ανησυχία για μια κρυφή κρίση δημόσιας υγείας εν μέσω πολυετών περικοπών στην αστυνομία.

Τα στοιχεία, επίσης, έθεσαν υπό αμφισβήτηση τις ευρέως διαδεδομένες παραδοχές σχετικά με το ποιους επηρεάζουν τα εγκλήματα με μαχαίρια. Τα δεδομένα έδειξαν ότι στην Αγγλία και στην Ουαλία, στα δέκα χρόνια που προηγήθηκαν του 2015, το ένα τρίτο των θυμάτων ήταν μαύροι. Ωστόσο, εκτός της πρωτεύουσας, οι θάνατοι από μαχαιρώματα ανάμεσα στους νέους δεν αφορούσαν κυρίως μαύρα αγόρια, καθώς την ίδια περίοδο λιγότερα από ένα στα πέντε θύματα εκτός Λονδίνου ήταν μαύροι.

Παρόλο που τα εγκλήματα με μαχαίρια ήταν ένα πολυσυζητημένο θέμα, τα στοιχεία δεν ήταν άμεσα διαθέσιμα στους πολιτικούς και τους αρμόδιους για τη χάραξη πολιτικής, γεγονός που προκάλεσε ερωτήματα σχετικά με τον τρόπο που θα μπορούσε να διαμορφωθεί μια αποτελεσματική στρατηγική, όταν οι βασικές λεπτομέρειες σχετικά με τους ανθρώπους που επηρεάζουν τα εγκλήματα με μαχαίρια δεν ήταν προσβάσιμες.

Τα δεδομένα αποτέλεσαν τη βάση του βραβευμένου έργου μας, το οποίο επαναπροσδιόρισε τη δημόσια συζήτηση για τα εγκλήματα με μαχαίρια. Το έργο δεν θα ήταν δυνατό να υλοποιηθεί χωρίς τη δημιουργία του δικού μας συνόλου δεδομένων.

Η συγγραφέας

Η Καέλαν Μπαρ είναι αρχισυντάκτρια έργων δεδομένων στην Guardian News & Media.

3.2

Η Ιστορία ενός Αριθμού και πώς να Σταθεί Κανείς στο Πρόβλημα της Τιμής

Helen Verran

Περίληψη: Οι αριθμοί είναι φαινομενικά απλές και ευθείες μετρήσεις της αξίας, αλλά προσέξτε: οι αριθμοί υποκρύπτουν ηθικά και πολιτικά προβλήματα.

Λέξεις-κλειδιά: αξία οικοσυστημάτων, περιβαλλοντική διακυβέρνηση της Αυστραλίας, κοινωνιολογία της ποσοτικοποίησης, δημοσιογραφία δεδομένων, σπουδές επιστήμης και τεχνολογίας

Στις αρχές του αιώνα, το αυστραλιανό κράτος ανέπτυξε μια περιβαλλοντική πολιτική που προέβλεπε την πλήρη επιδότηση του εργατικού κόστους για τους ιδιοκτήτες γης, εφόσον αυτοί αναλάμβαναν ειδικά συμφωνημένες εργασίες αρχιτεκτονικής τοπίου που αποσκοπούσαν στην ανάσχεση της περιβαλλοντικής υποβάθμισης. Ωστόσο, δεδομένης της σχεδόν απόλυτης κυριαρχίας των οπαδών του νεοφιλελευθερισμού σε αυτόν τον τομέα της πολιτικής εκείνη την εποχή, οι πολιτικές περιγράφηκαν με την ιλιγγιώδη διγλωσσία της αξίας του φυσικού κεφαλαίου και των υπηρεσιών των οικοσυστημάτων. Στα έγγραφα για την ακολουθητέα πολιτική, αυτή περιγράφεται ως «αγορά περιβαλλοντικών παρεμβάσεων για την ενίσχυση της κρατικής περιβαλλοντικής αξίας». Έτσι, το 2009, μια κρατική κυβερνητική υπηρεσία προέβη στον εξής σχεδόν ακατανόητο ισχυρισμό σχετικά με την επιτυχία της συγκεκριμένης πολιτικής: το 2009 η συνεισφορά στο ακαθάριστο εγχώριο προϊόν (ΑΕΠ) της Αυστραλίας, από συναλλαγές στις οποίες η πολιτεία επένδυσε σε περιβαλλοντικές παρεμβάσεις για την αύξηση της αξίας των οικοσυστημάτων από τους κατόχους αγροτικής γης στην Περιοχή Διαχείρισης Φυσικών Πόρων (Natural Resource Management Region, NRMR) του Κορανγκαμάιτ, υπολογίστηκε σε 4,94 εκατομμύρια αυστραλιανά δολάρια.

Το νούμερο, για το οποίο μιλάω εδώ, προέκυψε από μια ανακοίνωση στον Τύπο που εξέδωσε η κυβέρνηση της αυστραλιανής πολιτείας της Βικτόρια το 2009. Το δελτίο Τύπου ανακοίνωνε την επιτυχία της επένδυσης της πολιτειακής κυβέρνησης για τη διατήρηση του περιβάλλοντος σε μία από τις 57 Περιοχές Διαχείρισης Φυσικών Πόρων (ΠΔΦΠ) στην Αυστραλία. Η περιβαλλοντική διοικητική περιοχή με τις χλοώδεις πεδιάδες από βασάλτη, που εκτείνεται από την Ανατολή έως τη Δύση στη νοτιοκεντρική Βικτόρια, ονομάζεται Κορανγκαμάιτ –αυτός ο όρος της γλώσσας των Αβορίγινων αντικατέστησε την ονομασία που είχαν δώσει στην περιοχή οι πρώτοι Βρετανοί κτηνοτρόφοι, όταν εισέβαλαν στη χώρα από την Τασμανία στα μέσα του 19ου αιώνα. Τότε ήταν που ονόμασαν την περιοχή «Australia Felix» (Τυχερή Αυστραλία) και βάλθηκαν να κόψουν όλα τα δέντρα. Οι καταπατητές, οι οποίοι στη συνέχεια έγιναν ιδιοκτήτες γης στην περιοχή, θα μετατρέπονταν σε λιγότερο από έναν αιώνα σε ένα είδος αποικιακής κτηματικής αριστοκρατίας. Το 2008, στο πλαίσιο του προγράμματος EcoTender στην ΠΔΦΠ του Κορανγκαμάιτ, η κυβέρνηση της Βικτόρια αγόρασε την αξία χρήσης των υπηρεσιών οικοσυστημάτων από τους απογόνους αυτών των καταπατητών σε δημοπρασίες pay-as-bid (πληρωμή του ποσού της προσφοράς). Το 2009 η συμβολή των συναλλαγών αυτών στο ΑΕΠ της Αυστραλίας υπολογίστηκε στα 4,94 εκατομμύρια δολάρια Αυστραλίας. Η ανακοίνωση αυτής της τιμής ήταν η αφορμή για το δελτίο Τύπου, όπου συνάντησα για πρώτη φορά τον συγκεκριμένο αριθμό.

Αμφιβάλλω εάν κάποιος δημοσιογράφος παρατήρησε την είδηση που δημοσιεύτηκε σε εκείνη την ενημέρωση, συμπεριλαμβανομένης της αριθμητικής της τιμής –ο αριθμός αυτός δεν αποτελεί ακριβώς καυτή είδηση. Στο πλαίσιο ενός δελτίου Τύπου, η αναφορά μιας συγκεκριμένης αριθμητικής τιμής καθησυχάζει. Οι εθνικοί λογαριασμοί είναι σημαντικοί και πραγματικοί –εάν αυτή η περιφερειακή κυβερνητική παρέμβαση αποδίδεται ως μια συγκεκριμένη τιμή που συμβάλλει στην εθνική οικονομία, τότε είναι σαφές ότι η κυβερνητική παρέμβαση είναι κάτι καλό. Ο προσδιορισμός της τιμής εδώ κάνει τις βελτιώσεις που επιφέρουν οι κυβερνητικές παρεμβάσεις να φαντάζουν ως κάτι πραγματικό. Το συμπέρασμα είναι ότι η συγκεκριμένη πολιτική οδηγεί σε χρηστή περιβαλλοντική διαχείριση. Φυσικά, η πραγματική αξία στην οποία παραπέμπει ο αριθμός (4,94 εκατομμύρια δολάρια Αυστραλίας), δηλαδή αυτό που εμμέσως ισχυρίζεται ότι καταδεικνύει, δεν ενδιαφέρει κανέναν ιδιαίτερα. Το γεγονός ότι ένας αριθμός φαίνεται να αντιστοιχεί σε κάτι «εκεί έξω» που μπορεί να αποτιμηθεί είναι αρκετό για να επιτύχει εφησυχασμό.

Η δική μου εξιστόρηση για τον εν λόγω αριθμό προσφέρει μια λεπτομερή περιγραφή των κοινωνικοτεχνικών μέσων με τα οποία αυτός δημιουργήθηκε. Το ενοχλητικό αποτέλεσμα του θέματος είναι ότι αποκάλυψε πως αυτό το κοινότοπο νούμερο (που κυκλοφόρησε ασχολίαστο από τα μέσα ενημέρωσης) δεν είναι παρά μια ισχνή συγκάλυψη. Τα σοβαρά προβλήματα καραδοκούν. Πριν να αρχίσω να αφηγούμαι την ιστορία μου και να εξηγήσω τη φύση αυτών των σοβαρών προβλημάτων που φαίνεται να σκιάζουν κάθε αποτίμηση, ακόμη και μια τόσο μπανάλ αποτίμηση, επιτρέψτε μου να απαντήσω προκαταβολικά σε ορισμένα ερωτήματα που φαντάζομαι ότι μπορεί να έχουν αρχίσει να αναδύονται για τους αναγνώστες του Εγχειριδίου Δημοσιογραφίας Δεδομένων.

Καταρχάς, αναγνωρίζω ότι η αφήγηση μιας ιστορίας για το πώς ένας αριθμός έχει ζωντανέψει, αντί της εξεύρεσης κάποιων μέσων για την προώθηση της οπτικοποίησης της σημασίας του αριθμού αυτού σε συγκεκριμένο πλαίσιο, είναι μάλλον μια ασυνήθιστη προσέγγιση στη σύγχρονη δημοσιογραφία δεδομένων. Μπορώ να φανταστώ έναν δημοσιογράφο δεδομένων να αμφιβάλλει ότι μια τέτοια αφήγηση μπορεί να λειτουργήσει. Ίσως μια πρώτη απάντηση είναι να σας υπενθυμίσω ότι δεν πρόκειται για μια επιλογή του τύπου «ή το ένα ή το άλλο» και ότι το να εργάζεται κανείς συνδυάζοντας τις πηγές διήγησης και οπτικοποίησης κατά την αποκωδικοποίηση και την ερμηνεία είναι ένας αποτελεσματικός τρόπος για να δώσετε στους αναγνώστες να καταλάβουν την ιδέα σας. Κατά την παρουσίαση μιας τέτοιας διασύνδεσης, οι δημοσιογράφοι θα πρέπει πάντα να θυμούνται ότι υπάρχουν δύο βασικές θέσεις από τις οποίες μπορεί να μιλήσει κανείς κατά την ανάμειξη διηγήσεων και οπτικού υλικού. Κάποιος θα μπορούσε να προχωρήσει σαν το οπτικό υλικό να είναι ενσωματωμένο στην εξιστόρηση. Επομένως, στην περίπτωση αυτή, απευθυνόμαστε *στο* οπτικό υλικό, το οποίο φαίνεται να αναπαριστά ή να εικονογραφεί κάτι στην ιστορία. Αλλιώς, μπορείτε να προχωρήσετε σαν η εξιστόρηση να είναι ενσωματωμένη στο οπτικό υλικό. Τότε, μιλάτε *από μέσα*, διαγραμματικά. Αυτή είναι μια λιγότερο συνηθισμένη στρατηγική στη δημοσιογραφία δεδομένων, αλλά μπορώ να φανταστώ ότι η ιστορία που αφηγούμαι εδώ θα μπορούσε κάλλιστα να χρησιμοποιηθεί με αυτόν τον τρόπο. Φυσικά, η εναλλαγή μεταξύ αυτών των θέσεων μέσα σε ένα μεμονωμένο έργο ίσως είναι η πιο αποτελεσματική στρατηγική (για μια περιγραφή αυτής της εναλλαγής, βλ. Verran & Winthereik, 2016).

Δεύτερον, μπορεί να σας φανεί περίεργο να λέτε την ιστορία ενός πολύ συγκεκριμένου αριθμού, όταν αυτό που σαφώς έχει σημασία όσον αφορά στη λήψη αποφάσεων και στον σχεδιασμό πολιτικών (και αυτό που ενδιαφέρει τους δημοσιογράφους δεδομένων) είναι το τι μπορεί να βγει από σύνολα δεδομένων, με την επιστράτευση του ενός ή του άλλου αλγορίθμου. Αυτή η ανησυχία μπορεί να σας ωθήσει να αναρωτηθείτε για τις σχέσεις μεταξύ αριθμών και συνόλων δεδομένων. Η απάντηση σε ένα τέτοιο ερώτημα είναι αρκετά απλή και όχι πολύ ενδιαφέρουσα. Υπάρχουν πολλοί αριθμοί σε ένα σύνολο δεδομένων. Η σχέση είναι μια σχέση ένα-προς-πολλά, μολονότι οι αριθμοί είναι συγκεντρωμένοι σε πολύ ακριβείς πίνακες. Το πιο ενδιαφέρον ερώτημα αφορά στη σχέση μεταξύ αριθμών και αλγορίθμων. Η απάντησή μου θα ήταν ότι, ενώ οι αλγόριθ-

μοι επιστρατεύουν ένα πρωτόκολλο που επεξεργάζεται τον τρόπο λειτουργίας των σχέσεων οι οποίες είναι ενσωματωμένες σε μια βάση δεδομένων, οι αριθμοί εκφράζουν ένα πρωτόκολλο που καθορίζει τον τρόπο λειτουργίας των σχέσεων της συλλογικής ύπαρξης. Η αρίθμηση είναι μια μορφή δημιουργίας αλγορίθμου και αντίστροφο[29]. Θα μπορούσαμε να πούμε ότι οι αριθμοί είναι για τους αλγορίθμους ό,τι είναι ένας σπόρος για το φυτό που μπορεί να βλαστήσει από αυτόν –εναλλακτικά, μπορούμε να πούμε ότι έχουν μια σχέση κότας και αυγού. Αν και υπάρχουν ορισμένα διαφορετικά κοινωνικοτεχνικά χαρακτηριστικά για τη δημιουργία τιμής απαρίθμησης με αναλογικά μέσα (αναμειγνύοντας γνωστικούς, γλωσσικούς και γραφικούς πόρους), για τη συμβατική απαρίθμηση, όπως αυτή διδάσκεται στα παιδιά του δημοτικού σχολείου, και για τη δημιουργία τιμής απαρίθμησης με ψηφιακούς υπολογισμούς, αυτό που έχει σημασία εδώ είναι η ομοιότητα: τα 4,94 εκατομμύρια δολάρια Αυστραλίας έχουν παραχθεί αλγοριθμικά και εκφράζουν ένα συγκεκριμένο σύνολο σχέσεων ενσωματωμένο σε ένα συγκεκριμένο σύνολο δεδομένων, αλλά εξακολουθούν να παρουσιάζονται ως ένας απλός αριθμός[30].

Τώρα, λοιπόν, θα επιστρέψω στην ιστορία μου. Η οικεία περιγραφή της κατασκευής αριθμών που αφηγούμαι εδώ σαν ιστορία θα επέτρεπε σε έναν δημοσιογράφο να αναγνωρίσει ότι η καλή είδηση που η κυβέρνηση εκβιάζει πονηρά, με τη δημοσίευση δελτίου Τύπου, δεν είναι μια ξεκάθαρη υπόθεση. Βλέπουμε ότι πιθανόν ένα αποκαλυπτικό πολιτικό δημοσίευμα θα ήταν περισσότερο κατάλληλο. Οι λεπτομέρειες για τον τρόπο με τον οποίο κατασκευάζεται ο αριθμός αποκαλύπτουν ότι αυτό το πρόγραμμα περιβαλλοντικής παρέμβασης με σύμπραξη δημόσιου και ιδιωτικού τομέα συνεπάγεται ότι το κράτος πληρώνει πολύ πλούσιους ιδιοκτήτες γης, για να πραγματοποιήσουν εργασίες που θα αυξήσουν την αξία της ιδιοκτησίας τους.

Το ερώτημα που μπορεί να εγείρει η ιστορία μου είναι πώς θα μπορούσε ένας δημοσιογράφος είτε να επευφημήσει είτε να ξεγυμνώσει αυτόν τον αριθμό με καλή πίστη. Όταν τελειώσω με την ιστορία, θα υποστηρίξω ότι αυτό δεν είναι το σωστό ερώτημα.

Η διήγηση της ιστορίας ενός αριθμού

Ποια είναι η σειρά κοινωνικοτεχνικών διαδικασιών, μέσω των οποίων δημιουργείται η αξία του φυσικού κεφαλαίου και των υπηρεσιών των οικοσυστημάτων σε αυτό το πρόγραμμα σύμπραξης δημόσιου και ιδιωτικού τομέα, ώστε η αξία αυτή να μπορεί να αποτελέσει αντικείμενο συναλλαγής μεταξύ της κυβέρνησης ως αγοραστή και του ιδιοκτήτη γης ως πωλητή; Και πώς ακριβώς το αντίτιμο των συναλλαγών συμβάλλει στο συνολικό οριακό κέρδος που επιτυγχάνεται στο σύνολο της αυστραλιανής οικονομικής δραστηριότητας, δηλαδή στο ΑΕΠ της Αυστραλίας; Απαντώ σε αυτό το διττό ερώτημα με μια βήμα προς βήμα παρουσίαση των όσων απαιτούνται, για να δημιουργήσει ένας κάτοχος γης ένα προϊόν («αξία του φυσικού κεφαλαίου και των υπηρεσιών των οικοσυστημάτων») το οποίο θα είναι ανταγωνιστικό σε μια δημοπρασία που διοργανώνει η κυβέρνηση για την ανάθεση σύμβασης. Η βρόμικη δουλειά που απαιτείται για να παραχθεί

29. Η ιδέα ότι οι αριθμοί και οι αλγόριθμοι έχουν μια ομοιότητα είναι ενδεχομένως καινούρια για πολλούς αναγνώστες, καθώς έχουμε συνηθίσει να σκεφτόμαστε τους αριθμούς ως «αφαιρέσεις». Η δική μου (ασυνήθιστη) περιγραφή των αριθμών τους παρουσιάζει ως πολύ συνηθισμένες υλικές σημειωτικές οντότητες που κατοικούν στο εδώ και στο τώρα. Για μια περιγραφή των διαφορετικών πρωτοκόλλων που κινητοποιούν τις σχέσεις μέσα σε μια ενιαία στιγμή συλλογικής ύπαρξης, βλ. Watson, H. (1990) «Investigating the social foundations of mathematics: Natural number in culturally diverse forms of life», *Social Studies of Science*, 20(2), σσ. 283-312. DOI: https://doi.org/10.1177/030631290020002004 ή Verran, H. (2001) «Two consistent logics of numbering», στο *Science and an African logic*, σσ. 177-205, Σικάγο: University of Chicago Press.

30. Για έναν απολογισμό των διαφορετικών κοινωνικοτεχνικών χαρακτηριστικών τριών αριθμών που εμφανίζονται ποικιλοτρόπως σε αναλογικά ή ψηφιακά περιβάλλοντα, βλ. Verran, H. (2015) «Enumerated entities in public policy and governance», στο E. Davis & P. J. Davis (επιμ.), Mathematics, substance and surmise: Views on the meaning and ontology of mathematics, σσ. 365-379, Νέα Υόρκη: Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-21473-3_18.

αυτό το προϊόν, περιλαμβάνει το να λασπωθείς από το χώμα, να φυτέψεις δενδρύλλια, να επιδιορθώσεις φράχτες και γενικά να καταβάλλεις κάθε δυνατή προσπάθεια, προκειμένου να αποκατασταθεί η ζημιά που προκλήθηκε στη γη –ίσως από τους παππούδες και τις γιαγιάδες του ιδιοκτήτη, οι οποίοι απερίσκεπτα και άπληστα αποψίλωσαν τη γη από τα δέντρα και την έσπειραν με υδροβόρα φυτά, με την ελπίδα για περισσότερα σιτηρά ή περισσότερο μαλλί και μεγαλύτερη οικογενειακή περιουσία. Η αξία του φυσικού κεφαλαίου και των υπηρεσιών των οικοσυστημάτων γεννάται παρεμβαίνοντας στις περιβαλλοντικές διαδικασίες.

Η τιμή, η οποία είναι το προϊόν προς διακίνηση, ξεκινάει από την εργασία των δημοσίων υπαλλήλων που απασχολούνται σε ένα τμήμα της κυβέρνησης της πολιτείας της Βικτόρια (τότε επρόκειτο για το Τμήμα Αειφορίας και Περιβάλλοντος). Συλλογικά οι εν λόγω υπάλληλοι αποφασίζουν ποιες θα είναι οι κρατικές περιοχές εντός των οποίων η διοίκηση θα «τρέξει» τις δημοπρασίες. Για τον σκοπό αυτό, το EnSym, μια πλατφόρμα μοντελοποίησης περιβαλλοντικών συστημάτων, είναι ένα σημαντικό εργαλείο. Η υπολογιστική του ικανότητα είναι ένα θαύμα· γνωρίζει τη «φύση εκεί έξω» όπως κανένας επιστήμονας δεν τη γνώρισε ποτέ. Μπορούν να παραχθούν ακριβείς και εστιασμένες αναπαραστάσεις –πιθανότατα εν μία νυκτί.

> Το λογισμικό αυτό έχει δημιουργηθεί από την ομάδα ecoMarkets και ενσωματώνει την επιστήμη, τα πρότυπα, τις μετρήσεις και τις πληροφορίες που έχουν αναπτυχθεί στο πλαίσιο του Τμήματος Αειφορίας Περιβάλλοντος, καθώς και πολλά κορυφαία διεθνή και εθνικά επιστημονικά μοντέλα. Το EnSym περιλαμβάνει τρία βασικά εργαλεία: το «Site Assessment Tool» για εργασίες πεδίου, το «Landscape Preference Tool» για την ιεράρχηση προτεραιοτήτων και τη δημιουργία μετρήσεων και το «BioSim» για τον σχεδιασμό λεκανών απορροής (DSE, 2018).

Η ιεράρχηση και η χαρτογράφηση των κρατικών περιοχών στις οποίες θα διεξαχθούν οι δημοπρασίες, ο προσδιορισμός και η ποσοτικοποίηση των περιβαλλοντικών ωφελειών, οι οικολογικές τιμές που θα μπορούσαν να αυξηθούν μέσω των επίγειων εργασιών διατήρησης και αναβλάστησης, καταγράφονται σε αριθμητική μορφή. Αντιπροσωπεύουν ιδιότητες του οικοσυστήματος στα εδάφη «εκεί έξω». Μάλιστα, το υπολογιστικό πρόγραμμα μπορεί να κάνει κάτι περισσότερο από αυτό: μπορεί, επίσης, να δημιουργήσει ένα σενάριο για ανθρώπινη παρέμβαση. Όπως το σενάριο ενός θεατρικού έργου απαιτεί παραγωγή, το ίδιο ισχύει και για αυτό το σενάριο. Και, καθώς αυτό το σενάριο ζωντανεύει, η «φύση εκεί έξω» μοιάζει να έρχεται πιο κοντά. Παύει να είναι μια εντελώς απομακρυσμένη «φύση εκεί έξω» και γίνεται φύση ως υποδομή της ανθρώπινης ζωής, μια υποδομή στην οποία μπορούμε να επέμβουμε, για να της φτιάξουμε τα «υδραυλικά».

Όταν το σενάριο για μια χορογραφημένη παραγωγή συλλογικής ανθρώπινης προσπάθειας είναι έτοιμο, η κυβέρνηση καλεί τους ιδιοκτήτες γης στην περιοχή του έργου σε εκδήλωση ενδιαφέροντος ως επόμενο βήμα. Σε απάντηση στις υποβληθείσες εκδηλώσεις ενδιαφέροντος, ένας κυβερνητικός αξιωματούχος επισκέπτεται όλα τα ακίνητα. Μπορούμε να φανταστούμε ότι αυτός ο υπάλληλος μεταφέρει το γενικό σενάριο που παράγεται από το EnSym σε ένα πραγματικό μέρος σε μια δεδομένη χρονική στιγμή. Έχει μπροστά του ένα τρομερό μεταφραστικό έργο.Ο υπεύθυνος πεδίου αξιολογεί πιθανούς χώρους για έργα που θα μπορούσαν να αποτελέσουν σκηνικό για την παραγωγή του σεναρίου. Ο στόχος είναι να ενισχυθεί η παραγωγή των συγκεκριμένων υπηρεσιών οικοσυστημάτων. Επομένως, ο υπεύθυνος πρέπει να εκτιμήσει την πιθανότητα συγκεκριμένες δράσεις σε ένα συγκεκριμένο μέρος να προκαλέσουν αύξηση της παροχής υπηρεσιών από το οικοσύστημα, αυξάνοντας με αυτόν τον τρόπο την αξία της συγκεκριμένης υπηρεσίας οικοσυστημάτων που παράγεται από το συγκεκριμένο κτήμα –και, προσθέτοντας πολλές τέτοιες αυξήσεις στο πλαίσιο του προγράμματος παρέμβασης, από το κράτος στο σύνολό του.

Ο ιδιοκτήτης της γης και ο κυβερνητικός αξιωματούχος καταστρώνουν μαζί ένα σχέδιο. Με συνεχείς διαπραγματεύσεις, καταρτίζεται ένα επίσημο πλάνο διαχείρισης για συγκεκριμένα αγροτεμάχια. Ο υπάλληλος πεδίου συντάσσει αυτό το σχέδιο σε όρους σύμβασης. Οι κάτοχοι γης προσδι-

ορίζουν λεπτομερώς τις πραγματικές εργασίες που θα εκτελέσουν για την υλοποίηση του σχεδίου. Έτσι, ένα προϊόν που λαμβάνει τη μορφή μιας συγκεκριμένης «αξίας του φυσικού κεφαλαίου και των υπηρεσιών των οικοσυστημάτων» σχεδιάζεται και προσδιορίζεται ως μια σειρά συγκεκριμένων εργασιών που πρέπει να ολοκληρωθούν σε μια συγκεκριμένη χρονική περίοδο: να φυτευτούν τόσα φυτά από αυτό το σύνολο ειδών, σε αυτή τη σειρά, σε αυτή τη συγκεκριμένη γωνιά αυτού του συγκεκριμένου περιβόλου, και να περιφραχτούν για να δημιουργηθεί ένα οικόπεδο διατήρησης των τάδε διαστάσεων, χρησιμοποιώντας αυτά τα υλικά.Οι κάτοχοι γης υπολογίζουν το κόστος των έργων που ορίζει το κράτος, το οποίο αναμφίβολα περιλαμβάνει μια γενναιόδωρη πληρωμή εργασίας. Καταλήγουν σε ένα αντίτιμο που πρέπει να πληρώσει η κυβέρνηση για να αγοράσει αυτό το προϊόν, μια συγκεκριμένη «αξία του φυσικού κεφαλαίου και των υπηρεσιών των οικοσυστημάτων». Εδώ, προσδιορίζουν το χρηματικό ποσό που είναι διατεθειμένοι να δεχτούν, για να αναλάβουν τις συγκεκριμένες εργασίες και, ως εκ τούτου, να παραδώσουν την αξία του φυσικού κεφαλαίου και των υπηρεσιών των οικοσυστημάτων μέχρι την καθορισμένη ημερομηνία. Υποβάλλουν σχετικά έγγραφα στην κυβέρνηση σε σφραγισμένο φάκελο.

Πώς, λοιπόν, γίνεται η δημοπρασία που ακολουθεί; Εδώ το EnSym είναι και πάλι σημαντικό για την αξιολόγηση των προσφορών. Το EnSym δεν είναι απλά γνώστης της «φύσης εκεί έξω» και συγγραφέας σεναρίων παρέμβασης σε αυτό το «εκεί έξω» που φαντάζει ως υποδομή, αλλά είναι, επίσης, ένας απομακρυσμένος κριτικός παρατηρητής ο οποίος μπορεί να αξιολογήσει τις προσφορές που έχουν γίνει για την παραγωγή αυτού του σεναρίου, όπως η Warner Bros μπορεί να αξιολογήσει τις ανταγωνιστικές προσφορές για την παραγωγή μιας ταινίας. Οι προσφορές κατατάσσονται σύμφωνα με έναν υπολογισμένο «δείκτη περιβαλλοντικών ωφελειών» και την τιμή που προτείνει ο ιδιοκτήτης της γης. Πρέπει να υποθέσουμε ότι η κυβέρνηση αγοράζει το προϊόν που προσφέρει τον υψηλότερο «δείκτη περιβαλλοντικών ωφελειών» ανά μονάδα κόστους.

> Τώρα, ως προς την αξιολόγηση της προσφοράς: όλες οι προσφορές αξιολογούνται αντικειμενικά, με βάση την εκτιμώμενη αλλαγή στα περιβαλλοντικά αποτελέσματα· την αξία αυτής της αλλαγής· την αξία των περιουσιακών στοιχείων που επηρεάζονται από αυτές τις αλλαγές (σημαντικότητα)· το κόστος σε δολάρια, δηλαδή την τιμή που καθορίζεται από τον κάτοχο της γης (DSE, 2018).

Όταν ανακοινωθούν τα αποτελέσματα της δημοπρασίας, οι επιλεγμένοι πλειοδότες υπογράφουν μια τελική σύμβαση, με βάση το σχέδιο διαχείρισης και το υποβληθέν χρονοδιάγραμμα εργασιών, όπου ορίζεται η χωρική και χρονική οργάνωση. Όταν υπογραφούν όλα τα έγγραφα, εφαρμόζονται οι ρυθμίσεις περί υποβολής εκθέσεων και μπορεί να δρομολογηθεί η πληρωμή: «Το Τμήμα Αειφορίας Περιβάλλοντος διαβιβάζει την πληρωμή στους εγγεγραμμένους κατόχους γης με την παραλαβή του τιμολογίου. Οι πληρωμές εξαρτώνται από την ικανοποιητική πρόοδο των δράσεων που καθορίζονται στη Σύμβαση Διαχείρισης» (DSE, 2018).

Αυτό είναι κάτι καλό, σωστά;

Ό,τι παρουσίασα είναι μια ακριβής περιγραφή του τρόπου αγοραπωλησίας της αξίας του φυσικού κεφαλαίου και των υπηρεσιών των οικοσυστημάτων. Αυτό με φέρνει πίσω στο δελτίο Τύπου. Μια γρήγορη ανάγνωση του δελτίου Τύπου μπορεί να αφήσει στον αναγνώστη την εντύπωση ότι τα 4,94 εκατομμύρια δολάρια Αυστραλίας είναι η τιμή της πρόσθετης φυσικής κεφαλαιακής αξίας που έχει παραχθεί από το συγκεκριμένο κυβερνητικό πρόγραμμα. Με μια πρώτη ματιά, τα 4,94 εκατομμύρια δολάρια Αυστραλίας φαίνεται να είναι το οριακό κέρδος από τη φυσική κεφαλαιακή αξία της Αυστραλίας που επιτεύχθηκε στο πλαίσιο του προγράμματος. Όμως, αυτό είναι λάθος. Τα 4,94 εκατομμύρια δολάρια Αυστραλίας δεν είναι η ονομαστική αξία μιας φυσικής κεφαλαιακής αξίας. Παρακάτω θα εξηγήσω σε τι παραπέμπει αυτός ο αριθμός. Σε αυτό το σημείο, θέλω να σταθώ στο προϊόν που αγοράστηκε και πωλήθηκε σε αυτή τη δημοπρασία. Αυτό το προϊόν είναι το πρόβλημα στο οποίο θέλω να σταθώ.

Θέλω να ρωτήσω για την αξία της αύξησης της «αξίας του φυσικού κεφαλαίου και των υπηρεσιών των οικοσυστημάτων» που πέτυχε αυτό το περίπλοκο και μάλλον δαπανηρό κυβερνητικό πρόγραμμα. Μια προσεκτική ανάγνωση των λεπτομερειών του έργου με το οποίο προκύπτει αυτή η αύξηση της τιμής αποκαλύπτει ότι πουθενά και σε καμία στιγμή της διαδικασίας αυτή η τιμή δεν κατονομάζεται ούτε προσδιορίζεται. Το προϊόν που αγοράζεται και πωλείται τόσο σχολαστικά είναι μια απουσία . Και το χειρότερο είναι ότι δεν υπάρχει κυριολεκτικά κανένας τρόπος ώστε να γίνει κάτι άλλο. Το πρόγραμμα είναι μια πολύ περίτεχνη λογιστική άσκηση για τη διανομή χρημάτων. Όταν αυτό γίνει σαφές σε έναν εξωτερικό παρατηρητή, γίνεται επίσης φανερό ότι ποτέ δεν έμεινε κρυμμένο το τι είναι αυτή η άσκηση στην πραγματικότητα. Κατά βάθος, το πρόγραμμα αυτό είναι ένα νόμιμο μέσο για τη μεταφορά χρημάτων από τα κρατικά ταμεία στα χέρια των ιδιωτών ιδιοκτητών γης.

Αναγνωρίζοντας ότι πρόκειται για ένα πρόγραμμα περιβαλλοντικής διακυβέρνησης σε μια φιλελεύθερη κοινοβουλευτική δημοκρατία, στην οποία η κοινωνική τεχνολογία του πολιτικού κόμματος είναι καθοριστικής σημασίας, επιτρέψτε μου ως αφηγήτριά σας να φορέσω προσωρινά ένα κομματικό-πολιτικό καπέλο. Το Κορανγκαμάιτ είναι μια εκλογική περιφέρεια που ιστορικά ταλαντεύεται ανάμεσα στο αριστερό κεντρώο κόμμα (Εργατικό Κόμμα) και στο δεξιό κεντρώο κόμμα (Φιλελεύθερο Κόμμα) για την εκλογή αντιπροσώπου των κατοίκων της περιοχής στο Κοινοβούλιο της Βικτόρια. Είναι σαφώς προς το συμφέρον οποιασδήποτε κυβέρνησης –αριστερής ή δεξιάς– να προσεγγίσει τους ψηφοφόρους του εκλογικού σώματος. Και δεν υπάρχει καλύτερος τρόπος για να γίνει αυτό από το να βρεθούν τρόποι να μεταφερθούν νόμιμα πόροι από το κράτος στους τραπεζικούς λογαριασμούς των ψηφοφόρων. Το γεγονός ότι δεν υφίσταται δυνατότητα υπολογισμού της τιμής του προϊόντος που αγοράζει το κράτος και πωλούν οι ιδιοκτήτες γης εδώ, δεν έχει καμία σημασία, με βάση αυτήν την ανάγνωση.

Επιτρέψτε μου, λοιπόν, να συνοψίσω. Από οικονομική άποψη το πρόγραμμα αυτό υποτίθεται πως παράγει αξία του φυσικού κεφαλαίου και των υπηρεσιών των οικοσυστημάτων. Εάν το περιγράψει κανείς έτσι, αυτή είναι μια καλή είδηση. Τα χρήματα των φορολογουμένων χρησιμοποιήθηκαν σωστά, για να βελτιωθεί το περιβάλλον και να φυτευτούν δέντρα, ώστε να μειωθεί η υπερβολική παραγωγή διοξειδίου του άνθρακα στη Βικτόρια. Το πρόβλημα είναι ότι η αύξηση της αξίας του φυσικού κεφαλαίου της Βικτόριας δεν μπορεί να προσδιοριστεί, να διατυπωθεί ως αριθμός, παρά το γεγονός ότι πρόκειται για ένα προϊόν που αγοράζεται και πωλείται. Φαίνεται ότι, μολονότι υπάρχουν ακόμη τεχνικές δυσκολίες, προφανώς αυτό είναι κάτι καλό.

Όμως, χρησιμοποιώντας μια διαφορετική οικονομική προσέγγιση, το πρόγραμμα αυτό μπορεί εξίσου νόμιμα να περιγραφεί ως χρηματοδότηση της εργασίας της δεντροφύτευσης για την αύξηση της αξίας των ακινήτων των ιδιωτών ιδιοκτητών γης. Είναι ένα μέσο παρέμβασης για την αποκατάσταση της ζημίας που προκλήθηκε από προηγούμενα κυβερνητικά προγράμματα, τα οποία επιδοτούσαν την κακώς νοούμενη εργασία του καθαρισμού της γης, από την οποία κατά πάσα πιθανότητα επωφελήθηκαν οι παππούδες και οι γιαγιάδες των ιδιοκτητών γης, δημιουργώντας ένα όφελος το οποίο οι ιδιοκτήτες εξακολουθούν να απολαμβάνουν. Σύμφωνα με αυτήν την ανάγνωση, η κυβερνητική πολιτική που εφαρμόζεται στο EcoTender είναι ένα ακριβό πρόγραμμα που σκοπό έχει να μοιράσει με νομιμοφανή τρόπο τα χρήματα των φορολογουμένων. Προφανώς, αυτό είναι *κάτι κακό*.

Για να μην ασεβούμε απέναντι σε αριθμούς και αλγορίθμους: Στεκόμαστε στα προβλήματα της τιμής

Τι πρέπει να κάνει ένας δημοσιογράφος, λοιπόν; Γράφοντας ως ακαδημαϊκός και όχι ως δημοσιογράφος, μπορώ να απαντήσω σε αυτό το προφανές ερώτημα μόνο αόριστα. Καταρχάς, επανέρχομαι στον ισχυρισμό μου ότι η ονομαστική αξία του αριθμού που χρησιμοποιείται στο δελτίο Τύπου είναι μια ισχνή συγκάλυψη, για να αποπροσανατολίσει την κοινή γνώμη από τα προβλήματα που ελλοχεύουν. Κατά τη γνώμη μου, η αποτίμηση φέρνει πάντα ηθικά προ-

βλήματα που δεν μπορούν ποτέ να «κουκουλώνονται» για πολύ καιρό. Το σωστό ερώτημα που πρέπει να θέσουμε νομίζω είναι: «Πώς θα μπορούσε ένας δημοσιογράφος δεδομένων να ανταποκριθεί σε αυτό το ηθικό πρόβλημα;».

Πρώτον, ξεκαθαρίζω το θέμα των 4,94 εκατομμυρίων δολαρίων Αυστραλίας: ποιος είναι αυτός ο αριθμός; Από πού προέρχεται αυτή η εύσχημη νομισματική αξία; Έτσι περιγράφεται σε ένα ακαδημαϊκό άρθρο που σχολιάζει κριτικά το πρόγραμμα EcoTender:

> Σύμφωνα με αυτό το βασισμένο στην αγορά μοντέλο, η οικονομική αξία από τις υπηρεσίες οικοσυστημάτων δημιουργείται, όταν το ανά μονάδα κόστος συμμόρφωσης με τη σύμβαση διατήρησης είναι μικρότερο από την τιμή ανά μονάδα που κατακυρώνεται στους επιτυχόντες συμμετέχοντες στη δημοπρασία. Ενώ [για τους εν λόγω πωλητές] χάνεται κάποια οικονομική αξία μέσω της πιθανότητας μη παραγωγής εμπορεύσιμων αγαθών, ο περιορισμός συμμετοχής των ορθολογικών ιδιοκτητών γης εξασφαλίζει ότι θα υπάρξει καθαρή αύξηση της [οικονομικής] αξίας που δημιουργείται κατά τη διεξαγωγή της δημοπρασίας (Stoneham κ.α., 2012).

Σύμφωνα με την οικονομική μοντελοποίηση αυτής της πολιτικής, γίνεται η παραδοχή ότι οι ιδιοκτήτες γης θα υπολογίσουν αποτελεσματικά το κόστος που θα τους βαρύνει κατά την υλοποίηση του κυβερνητικού σεναρίου για την παρέμβαση στη φύση ως υποδομή –δημιουργώντας μια πιο αποτελεσματική απόδοση της λειτουργίας αυτής της φυσικής υποδομής. Όλοι υποθέτουν ότι ο ιδιοκτήτης γης θα έχει κέρδος, αν και, φυσικά, είναι πάντα πιθανό ο ιδιοκτήτης γης να έχει κάνει λάθος υπολογισμούς και, αντί για κέρδος, να έχει ζημία, αλλά αυτό δεν ενδιαφέρει την κυβέρνηση ως αγοραστή της αξίας που παράγεται από την εργασία των ιδιοκτητών γης.

Αυτό που ενδιαφέρει την κυβέρνηση είναι πώς αυτή η οικονομική συναλλαγή μπορεί να διατυπωθεί με εύσχημο τρόπο. Αυτό είναι ένα αρκετά μεγάλο πρόβλημα, όταν το προϊόν που αγοράζεται και πωλείται υφίσταται μονάχα στο πλαίσιο μιας δημοπρασίας. Η λύση σε αυτήν την προβληματική μορφή ύπαρξης του προϊόντος είναι η περίτεχνη, σύνθετη και πολύπλοκη τεχνολογία του εθνικού λογιστικού συστήματος. Δημιουργώντας μια αγορά για την αξία του φυσικού κεφαλαίου και των υπηρεσιών των οικοσυστημάτων, η κυβέρνηση θέλει να δείξει ότι συμβάλλει στην προστασία της φύσης. Και οι εθνικοί λογαριασμοί είναι το βολικό μέρος όπου αυτό μπορεί να παρουσιαστεί με νομισματικούς όρους. Ο «δείκτης περιβαλλοντικών ωφελειών», η συγκεκριμένη τιμή βάσει της οποίας η κυβέρνηση αγόρασε ένα συγκεκριμένο προϊόν (μια αξία του φυσικού κεφαλαίου και των υπηρεσιών των οικοσυστημάτων), είναι εφήμερη. Υπάρχει μόνο σαν αναλαμπή, μία στιγμή κατά τη διάρκεια της δημοπρασίας (Roffe, 2015). Παρ' όλη αυτή τη δυσκολία στη μορφή της ύπαρξής του, με ευφυή επινοητικότητα, επιτυγχάνεται τόσο ο τρόπος αγοραπωλησίας ενός προϊόντος που έχει μία και μοναδική εφήμερη στιγμή ύπαρξης, όσο και η ενσωμάτωση στοιχείων της συγκεκριμένης περίπτωσης οικονομικής δραστηριότητας στους εθνικούς λογαριασμούς, αν και ορισμένοι οικονομολόγοι έχουν σοβαρές επιφυλάξεις για την ακρίβειά τους (Stoneham κ.α., 2012).

Τα 4,94 εκατομμύρια δολάρια Αυστραλίας αποκλίνουν από τη δραστηριότητα του προγράμματος EcoTender και από το περιβάλλον που αυτό έχει σχεδιαστεί να αποκαταστήσει. Ωστόσο, σαφώς, αν η κυβέρνηση προβεί σε μια δήλωση ότι τα προγράμματά της έχουν αποκαταστήσει με επιτυχία ένα υποβαθμισμένο και κατεστραμμένο περιβάλλον, είναι καλύτερο να βρει έναν τρόπο να δείξει το μέγεθος αυτής της αποκατάστασης. Φαίνεται ότι οποιοσδήποτε αριθμός είναι καλύτερος από το τίποτα, σε αυτήν την περίπτωση. Και, σίγουρα, αυτός είναι ένας ευχάριστος, θετικός αριθμός. Ένας δυσάρεστος, αρνητικός αριθμός που αναμφίβολα είναι διαθέσιμος στους κυβερνητικούς λογιστές (το κόστος εκτέλεσης του κυβερνητικού προγράμματος) δεν θα είχε ποτέ θέση εδώ. Γιατί συνεχίζουμε με αυτόν τον περίεργα παράταιρο αριθμό; Μήπως το παρατραβάμε; Πού είναι το κακό σε ένα μικρό ταχυδακτυλουργικό κόλπο που γίνεται σχετικά εύκολα αντιληπτό; Η ανησυχία μου εδώ είναι ότι πρόκειται για την κατάχρηση ενός αριθμού που

δείχνει να είναι σκόπιμη. Αποτυγχάνει να σεβαστεί τους αριθμούς και αρνείται να αναγνωρίσει το πρόβλημα που η αρίθμηση –ή, εν προκειμένω, ο αλγόριθμος– συνεπάγεται πάντα. Καταστρέφει ένα πρωτόκολλο.

Η εξιστόρησή μου για έναν αριθμό που βρήκα κατά την επίσκεψή μου σε έναν κυβερνητικό ιστότοπο αποκάλυψε απερίφραστα ένα κυβερνητικό πρόγραμμα το οποίο παράγει ταυτόχρονα κοινωνικά πλεονεκτήματα και μειονεκτήματα. Η ταχυδακτυλουργική παράθεση αριθμών (χρησιμοποιώντας την ακριβή τιμή 4,94 εκατομμύρια δολάρια Αυστραλίας στο δελτίο Τύπου) παραπέμπει πλαγίως σε κάτι που πάντα απειλεί να μας κατατροπώσει: στην αποτίμηση ως πεδίο ηθικών εντάσεων και προβλημάτων.

Μήπως ο μεγάλος ισχυρισμός εδώ είναι ότι η αξία αποτελεί ένα ηθικό πρόβλημα που δεν μπορεί ποτέ να «κουκουλωθεί» για πολύ; Η θεωρία της αξίας είναι ένα τεράστιο θέμα, που έχει αρχαίες ρίζες σε όλες τις φιλοσοφικές παραδόσεις, και πρόκειται για μια κουνελότρυπα τεραστίων διαστάσεων, στην οποία αρνούμαι να εισέλθω. Επισημαίνω απλώς πως οι ισχυρισμοί, που ακούγονται συχνά τα τελευταία 30 χρόνια, ότι το αόρατο χέρι της αγοράς τιθασεύει τα ηθικά προβλήματα τα οποία σχετίζονται με την αξία, αποτελούν μια επικίνδυνη υπερβολή. Οι αγορές μπορεί να βρουν τρόπους να δαμάσουν στιγμιαία και εφήμερα την τιμή –όπως αποκαλύπτει η ιστορία μου. Όμως, το πρόβλημα με την αξία πάντα επιστρέφει. Το καθήκον του δημοσιογράφου δεδομένων είναι να φροντίσει για το θέμα αυτό.

Ακολουθούν μερικές προτάσεις για το πώς ένας δημοσιογράφος δεδομένων μπορεί να σέβεται τους αριθμούς και τους αλγορίθμους –ως πρωτόκολλα:

- Όταν έρχεστε αντιμέτωποι με κάτι αδιατάρακτο, όπου δεν υπάρχει καμία ένδειξη ηθικής έντασης, αλλά κάτι παραμονεύει, τότε τεντώστε τα αυτιά και ανοίξτε τα μάτια σας.
- Συντονιστείτε με τους αριθμούς και τους αλγορίθμους επί τόπου· βρείτε έναν τρόπο σκέψης για έναν αριθμό που σας τραβάει την προσοχή.
- Βρείτε τρόπους να κρυφοκοιτάξετε καλύτερα τους αριθμούς μέσα από τις χαραμάδες.
- Βρείτε τρόπο να σέβεστε τους αριθμούς και τους αλγορίθμους, όντας πειθαρχημένα περίεργοι.
- Αναγνωρίστε το γεγονός ότι οι αριθμοί έχουν προκαθορισμένη φύση και ιδιαίτερες ικανότητες που αναδύονται, όταν τους συναντά κανείς· ότι η ακολουθία των πρακτικών με τις οποίες δημιουργούνται, έχει σημασία.

Να είστε βέβαιοι πως, όταν μπορέσετε να τα κάνετε όλα αυτά αρκετά καλά, σας περιμένουν εκπλήξεις. Ενδιαφέροντα πράγματα συμβαίνουν μέσα στους αριθμούς καθώς αυτοί διαμορφώνονται.

Παραπομπές

DSE (2018) [Department of Sustainability and Environment, νυν Department of Environment, Land, Water and Planning (DELWP), Victoria State Government, Australia] I*nnovative market approaches: EcoMarkets. EcoTender and BushTender.* Διαθέσιμο σε: https://www.environment.vic.gov.au/innovative-market-approaches/ (Προσπελάστηκε: 16 Νοεμβρίου 2021).

Roffe, J. (2015). *Abstract market theory.* Λονδίνο: Palgrave Macmillan.

Stoneham, G., O'Keefe, A., Eigenraam, M. & Bain, D. (2012) *«Creating physical environmental asset accounts from markets for ecosystem conservation»*, Ecological Economics, 82, σσ. 114-122. DOI: https://doi.org/10.1016/j.ecolecon.2012.06.017

Verran, H. (2001) «Two consistent logics of numbering», στο *Science and an African logic,* σσ. 177-205), Σικάγο: University of Chicago Press.

Verran, H. (2015) «Enumerated entities in public policy and governance», στο E. Davis & P. J. Davis (επιμ.), *Mathematics, substance and surmise: Views on the meaning and ontology of mathematics,* σσ. 365-379, Νέα Υόρκη: Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-21473-3_18.

Verran, H. & Winthereik, B. R. (2016) «Innovation with words and visuals. A baroque sensibility», στο J. Law & E. Ruppert (επιμ.) *Modes of knowing*, σσ. 197-223. Μάντσεστερ: Mattering Press.

Watson, H. (1990) «Investigating the social foundations of mathematics: Natural number in culturally diverse forms of life», *Social Studies of Science,* 20(2), σσ. 283-312. DOI: https://doi.org/10.1177/030631290020002004

Η συγγραφέας

Η Έλεν Βέραν είναι Καθηγήτρια στο Πανεπιστήμιο Charles Darwin και προβληματίζεται για το πώς δημιουργούνται οι αριθμοί από τότε που συνάντησε ορισμένους εκπληκτικούς αριθμούς στη Γιορούμπαλαντ της Νιγηρίας πριν από πολλά χρόνια.

3.3

Κυριαρχία των Αυτοχθόνων επί των Δεδομένων τους: Επιπτώσεις για τη Δημοσιογραφία Δεδομένων

Tahu Kukutai, Maggie Walter

Περίληψη: Αυτό το κεφάλαιο παρουσιάζει ορισμένα από τα πιθανά δεινά της ψηφιοποίησης και εξετάζει πώς η κυριαρχία των αυτοχθόνων επί των δεδομένων τους, ως ένας αναδυόμενος τομέας της επιστήμης και του ακτιβισμού, μπορεί να μετριάσει τους κινδύνους, εξασφαλίζοντας παράλληλα τρόπους αποκόμισης ωφελειών.

Λέξεις-κλειδιά: κυριαρχία των αυτοχθόνων επί των δεδομένων τους, ακτιβισμός, δημοσιογραφία δεδομένων, στατιστική παρακολούθηση, αυτόχθονες πληθυσμοί

Οι ψηφιακές τεχνολογίες, συμπεριλαμβανομένων των τεχνολογιών παρακολούθησης και πληροφοριών και της τεχνητής νοημοσύνης («artificial intelligence», AI), γίνονται ολοένα και περισσότερο χαρακτηριστικό της ζωής των αυτοχθόνων πληθυσμών, ειδικά για τους λαούς στις ανεπτυγμένες και μεταβατικές οικονομίες. Ωστόσο, ενώ οι τεχνολογίες που βασίζονται στα δεδομένα μπορούν να προωθήσουν την καινοτομία και να βελτιώσουν την ανθρώπινη ευημερία, είναι απίθανο οι αυτόχθονες να μοιραστούν ισότιμα αυτά τα οφέλη, δεδομένου ότι οι πληθυσμοί τους βρίσκονται, σχεδόν καθολικά, τοποθετημένοι στο κοινωνικοοικονομικό, πολιτιστικό και πολιτικό περιθώριο. Η αυξανόμενη χρήση συνδεδεμένων και ολοκληρωμένων μεγάλων δεδομένων («big data») από τις κυβερνήσεις και τις επιχειρήσεις, επίσης, εγκυμονεί σημαντικούς κινδύνους για τους αυτόχθονες πληθυσμούς. Τέτοιοι είναι η οικειοποίηση της πολιτιστικής γνώσης και της πνευματικής ιδιοκτησίας, η εκμετάλλευση της γης και άλλων φυσικών πόρων και η διαιώνιση των διακρίσεων, του στίγματος και της συνεχιζόμενης περιθωριοποίησης. Αυτοί οι κίνδυνοι ενισχύονται από τις δημοσιογραφικές πρακτικές αφήγησης που ανακυκλώνουν καθιερωμένα σχήματα λόγου σχετικά με τη δυσλειτουργικότητα των ιθαγενών. Σε αυτό το κεφάλαιο παρουσιάζονται ορισμένα από τα πιθανά δεινά της ψηφιοποίησης και εξετάζεται πώς η κυριαρχία των αυτοχθόνων επί των δεδομένων τους («indigenous data sovereignty», ID-SOV), ως ένας αναδυόμενος τομέας της επιστήμης και του ακτιβισμού, μπορεί να μετριάσει τους κινδύνους, εξασφαλίζοντας παράλληλα τρόπους αποκόμισης ωφελειών. Καταλήγουμε προτείνοντας ότι η έρευνα και τα δίκτυα ID-SOV αποτελούν, επίσης, πολύτιμες πηγές δεδομένων και τεχνογνωσίας δεδομένων, που μπορούν να τροφοδοτήσουν περισσότερο ισότιμες, κριτικές και δίκαιες προσεγγίσεις στη δημοσιογραφία σχετικά με τους αυτόχθονες πληθυσμούς και τα ζητήματα που τους αφορούν.

Αυτόχθονες πληθυσμοί και δεδομένα

Υπολογίζεται ότι υπάρχουν περίπου 370 εκατομμύρια αυτόχθονες πληθυσμοί παγκοσμίως, οι οποίοι βρίσκονται σε όλες τις ηπείρους και μιλούν χιλιάδες διαφορετικές γλώσσες (United Nations, 2009). Σε παγκόσμιο επίπεδο, ο πραγματικός αριθμός είναι αδύνατον να προσδιορι-

στεί, καθώς η πλειονότητα των χωρών με αυτόχθονες πληθυσμούς δεν τους αναφέρουν ως τέτοιους στις εθνικές συλλογές δεδομένων (Mullane-Ronaki, 2017). Ανεξάρτητα από αυτές τις «ερήμους δεδομένων» των αυτοχθόνων και τη σημαντική παγκόσμια διακύμανση της πολιτικής αυτονομίας και του βιοτικού τους επιπέδου, υπάρχουν άφθονες αποδείξεις ότι οι αυτόχθονες συχνά ανήκουν στις φτωχότερες πληθυσμιακές ομάδες στις πατρίδες τους και κουβαλούν το μεγαλύτερο βάρος των ασθενειών, της υπέρμετρης φυλάκισης και ενός ευρύτατου φάσματος ανισοτήτων (Anderson κ.α., 2016· Stephens κ.α., 2006). Αυτή η κοινή τοποθέτηση της περιθωριοποίησης δεν είναι τυχαία· σχετίζεται άμεσα με την ιστορία τους ως αποικιοκρατούμενων και εκδιωγμένων λαών. Ωστόσο, οι καταστροφικές συνέπειες της αποικιοκρατίας και των συντρόφων της, της λευκής υπεροχής και του ρατσισμού, σπάνια αναγνωρίζονται, πόσο μάλλον επικρίνονται, στις κύριες δημοσιογραφικές απεικονίσεις των αυτοχθόνων λαών και κοινοτήτων.

Οι αυτόχθονες λαοί ανέκαθεν δραστηριοποιούνταν σε αυτό που σήμερα ονομάζουμε δεδομένα, με αρχαίες παραδόσεις καταγραφής και προστασίας των πληροφοριών και των γνώσεων μέσω, για παράδειγμα, της τέχνης, της ξυλογλυπτικής, των τοτέμ, των τραγουδιών, των ψαλμών, του χορού και των προσευχών. Οι σκόπιμες προσπάθειες εξάλειψης αυτών των πρακτικών και των συστημάτων γνώσης ήταν αναπόσπαστο μέρος των αποικιοκρατικών διαδικασιών. Την ίδια στιγμή, οι αυτόχθονες πληθυσμοί έγιναν γνωστοί μέσα από τα γραπτά των Ευρωπαίων περιηγητών, εξερευνητών και επιστημόνων, οι οποίοι παρουσιάστηκαν ως πιο αντικειμενικοί, επιστημονικοί και αξιόπιστοι «γνώστες» των αυτοχθόνων πληθυσμών και των πολιτισμών τους. Με την πάροδο του χρόνου οι φυλετικές ιεραρχίες που δικαιολογούσαν και συντηρούσαν την αποικιοκρατία εδραιώθηκαν και ενσωματώθηκαν μέσω ιδεολογικών δομών, θεσμικών ρυθμίσεων (π.χ. δουλεία, φυλετικός διαχωρισμός) και κρατικών πρακτικών κατηγοριοποίησης. Για παράδειγμα, οι Αβορίγινες και οι κάτοικοι των Νήσων του Στενού Τόρες στην Αυστραλία αποκλείονταν ρητά από την εθνική απογραφή μέχρι το 1971 και ο αποκλεισμός αυτός συνδέθηκε με παρόμοιους αποκλεισμούς από βασικά δικαιώματα του πολίτη, όπως η σύνταξη γήρατος (Chesterman & Galligan, 1997). Στη σύγχρονη εποχή, η αρμοδιότητα της απόφασης περί του εάν και πώς οι αυτόχθονες πληθυσμοί καταμετρώνται, κατηγοριοποιούνται, αναλύονται και αντιμετωπίζονται εξακολουθεί να ανήκει στις κυβερνήσεις –και όχι στους ίδιους τους αυτόχθονες πληθυσμούς. Η μεταφορά της εξουσίας επί των δεδομένων των αυτοχθόνων από το εθνικό κράτος πίσω στους αυτόχθονες λαούς βρίσκεται στο επίκεντρο του ID-SOV.

Ορισμός του ID-SOV

Η ορολογία του ID-SOV είναι σχετικά πρόσφατη, με την πρώτη σημαντική δημοσίευση για το θέμα να εμφανίζεται μόλις το 2015 (Kukutai & Taylor, 2016). Το ID-SOV αφορά στο δικαίωμα των αυτοχθόνων πληθυσμών να κατέχουν, να ελέγχουν, και να διαθέτουν δεδομένα που προέρχονται από αυτούς και τα οποία αναφέρονται στα μέλη, στα συστήματα γνώσεων, στα έθιμα ή στα εδάφη τους, και να έχουν πρόσβαση σε αυτά (First Nations Information Governance Centre, 2016· Snipp, 2016)[31]. Το ID-SOV στηρίζεται στα εγγενή δικαιώματα αυτοδιάθεσης και διακυβέρνησης των αυτοχθόνων πληθυσμών, της χώρας τους (συμπεριλαμβανομένων των εδαφών, των υδάτων και του εναέριου χώρου) και των πόρων τους, όπως περιγράφεται στη Διακήρυξη των Ηνωμένων Εθνών για τα Δικαιώματα των Αυτοχθόνων Λαών («United Nations Declaration on the Rights of Indigenous Peoples», UNDRIP)[32]. Το ID-SOV συνεπάγεται την επιδίωξη να χρησιμοποιούνται τα δεδομένα με τρόπους που υποστηρίζουν και ενισχύουν τη συλλογική ευημερία και την αυτοδιάθεση των αυτοχθόνων πληθυσμών –όπως υπογραμμίζεται από τις ΜΚΟ, τις κοινότητες και τις φυλές των αυτοχθόνων (First Nations Information

31. Στην Αοτεαρόα-Νέα Ζηλανδία, το δίκτυο ID-SOV Te Mana Raraunga ορίζει τα δεδομένα των Μαορί ως «ψηφιακές ή ψηφιοποιήσιμες πληροφορίες γνώσεων που αφορούν τον λαό των Μαορί ή προέρχονται από αυτόν, τη γλώσσα, τα πολιτισμικά μας στοιχεία, τους πόρους ή τα περιβάλλοντά μας» (Te Mana Raraunga, 2018a).
32. https://www.un.org/esa/socdev/unpfii/documents/DRIPS_en.pdf

Governance Centre, 2016· Hudson κ.α., 2016). Στην πράξη, το ID-SOV σημαίνει ότι οι αυτόχθονες πληθυσμοί πρέπει να είναι οι υπεύθυνοι για τη λήψη αποφάσεων σχετικά με τη χρήση και την αξιοποίηση των δεδομένων που τους αφορούν. Ως εκ τούτου, το ID-SOV γεννά ερωτήματα όπως: σε ποιον ανήκουν τα δεδομένα; Ποιος έχει την εξουσία να λαμβάνει αποφάσεις σχετικά με τον τρόπο πρόσβασης στα δεδομένα και υπό ποιες συνθήκες; Ποιοι είναι οι προβλεπόμενοι δικαιούχοι των δεδομένων και της χρήσης τους;

Το ID-SOV εξετάζει, επίσης, ακανθώδη ζητήματα σχετικά με την εξισορρόπηση των ατομικών δικαιωμάτων (συμπεριλαμβανομένων των δικαιωμάτων προστασίας της ιδιωτικής ζωής), των κινδύνων και των ωφελειών, με τα αντίστοιχα δικαιώματα, κινδύνους και ωφέλειες των ομάδων στις οποίες ανήκουν. Η επικέντρωση στα συλλογικά δικαιώματα και συμφέροντα είναι σημαντική, διότι υπερβαίνει τη στενή επικέντρωση στην προστασία και στον έλεγχο των προσωπικών δεδομένων που διαπερνά πολιτικές και κανονιστικές προσεγγίσεις όπως ο Γενικός Κανονισμός για την Προστασία Δεδομένων της Ευρωπαϊκής Ένωσης («General Data Protection Regulation», GDPR). Οι αγγλοευρωπαϊκές νομικές έννοιες της ιδιωτικής ζωής και της ιδιοκτησίας μεταφράζονται ανεπαρκώς στο πλαίσιο των κοινοτήτων των αυτοχθόνων, όπου τα άτομα αποτελούν μέρος μιας ευρύτερης ομάδας η οποία ορίζεται, για παράδειγμα, από κοινές γενεαλογίες ή γονίδια. Σε ένα τέτοιο πλαίσιο, η κοινή χρήση δεδομένων που κωδικοποιούν πληροφορίες σχετικά με άλλα μέλη της ομάδας δεν μπορεί να βασίζεται μόνο στην προσωπική συγκατάθεση, αλλά πρέπει επίσης να λαμβάνει υπόψη τα συλλογικά δικαιώματα και συμφέροντα (Hudson κ.α., 2020). Στενά συνδεδεμένη με το ID-SOV είναι η έννοια της διακυβέρνησης των δεδομένων των αυτοχθόνων, η οποία μπορεί να οριστεί, σε γενικές γραμμές, ως οι αρχές, οι δομές, οι μηχανισμοί λογοδοσίας, τα νομικά όργανα και οι πολιτικές μέσω των οποίων οι αυτόχθονες πληθυσμοί ασκούν έλεγχο στα δεδομένα τους (Te Mana Raraunga, 2018a). Η διακυβέρνηση των δεδομένων των αυτοχθόνων, στην ουσία της, είναι ένας τρόπος λειτουργικοποίησης του ID-SOV (Carroll κ.α., 2017). Μόνο μέσω της διακυβέρνησης των δεδομένων των αυτοχθόνων μπορούν να διεκδικηθούν τα δικαιώματα και τα συμφέροντά τους σε σχέση με τα δεδομένα (Walter, 2018).

Στατιστική παρακολούθηση και αυτόχθονες πληθυσμοί

Η σκιαγράφηση του προφίλ των αυτοχθόνων πληθυσμών και η στόχευση των υπηρεσιών δεν είναι κάτι καινούριο. Η επιτήρηση από το κράτος, τους θεσμούς και τους φορείς του αποτελεί παλαιόθεν μόνιμο χαρακτηριστικό της αποικιοκρατίας (Berda, 2013). Ακόμη και κατά τη διάρκεια του επίσημου αποκλεισμού των Αβορίγινων και των κατοίκων των Νησιών του Στενού Τόρες από την εθνική απογραφή στην Αυστραλία, η επιτήρηση των πληθυσμών των Αβορίγινων δεν έπαψε ποτέ να λαμβάνει χώρα (Briscoe, 2003). Αυτό που είναι νέο στη σκηνή της κοινωνικής πολιτικής είναι οι αδιαφανείς, πολύπλοκες και ολοένα και περισσότερο αυτοματοποιημένες διαδικασίες που διαμορφώνουν τη στόχευση και την κατάρτιση προφίλ (Henman, 2018). Ως «υποκείμενα δεδομένων» (Van Alsenoy κ.α., 2009), οι αυτόχθονες πληθυσμοί περιλαμβάνονται σε ένα ευρύ φάσμα συσσώρευσης δεδομένων, από αυτοπροσδιοριζόμενες πολιτικές και κοινωνικές ομάδες (π.χ. φυλές, εθνοτικές/φυλετικές ομάδες) μέχρι ομάδες συμφερόντων που ορίζονται από αναλυτές δεδομένων βάσει χαρακτηριστικών, συμπεριφοράς ή/και συνθηκών.

Η θέση που κατέχουν οι αυτόχθονες πληθυσμοί εντός αυτών των διαδικασιών δεδομένων δεν είναι μια καλή θέση. Αντιθέτως, ενώ οι πηγές των δεδομένων για τους αυτόχθονες πληθυσμούς εξελίσσονται με ταχείς ρυθμούς, τα χαρακτηριστικά των δεδομένων αυτών, ως αμείλικτη περιγραφική καταμέτρηση των διαφόρων σοβαρών κοινωνικοοικονομικών και υγειονομικών ανισοτήτων που υφίστανται οι αυτόχθονες πληθυσμοί, παραμένουν τα ίδια. Ο Γουόλτερ τα ονόμασε δεδομένα 5D (Walter, 2016), επειδή εστιάζουν στη Διαφορά, στην Ανισότητα, στη Μειονεξία, στη Δυσλειτουργία και στην Εξαθλίωση (Difference, Disparity, Disadvantage, Dysfunction, Deprivation). Τα στοιχεία που υποστηρίζουν αυτόν τον ισχυρισμό μπορούν εύκολα να βρεθούν μέσω μιας αναζήτησης στο Google με τον όρο «στατιστικές των αυτοχθόνων» ή εισάγοντας το

όνομα ενός ιθαγενούς λαού στην αναζήτηση (π.χ. ιθαγενείς Αμερικανοί, Αβορίγινες και κάτοικοι των νήσων του Στενού Τόρες, Μαορί, ιθαγενείς της Χαβάης, Πρώτα Έθνη, ιθαγενείς της Αλάσκας). Τα αποτελέσματα που εμφανίζονται είναι πάντα ένας θλιβερός κατάλογος που περιγράφει λεπτομερώς την υπερεκπροσώπηση των αυτοχθόνων σε αρνητικά στοιχεία για την υγεία, την εκπαίδευση, τη φτώχεια και το ποσοστό των φυλακισμένων.

Ο αντίκτυπος των δεδομένων 5D στις ζωές των αυτοχθόνων, επίσης, δεν είναι θετικός. Ως ο κύριος τρόπος με τον οποίο οι αυτόχθονες πληθυσμοί τοποθετούνται στο εθνικό αφήγημα, τα δεδομένα αυτά διαμορφώνουν την αντίληψη του κυρίαρχου, μη αυτόχθονα πληθυσμού για τους αυτόχθονες πληθυσμούς. Οι εν λόγω ιστορίες δεδομένων, που επηρεάζουν αυτά τα αφηγήματα, προβάλλονται συχνά μέσω δημοσιογραφικών ρεπορτάζ. Για παράδειγμα, ο Στόουναμ παρουσιάζει μελέτη όλων των άρθρων που σχετίζονται με την υγεία των Αβορίγινων σε τέσσερα εξέχοντα διαδικτυακά και έντυπα μέσα ενημέρωσης στην Αυστραλία (Stoneham, 2014). Τα τρία τέταρτα αυτών των άρθρων ήταν αρνητικά, εστιάζοντας σε θέματα όπως το αλκοόλ, η παιδική κακοποίηση, τα ναρκωτικά, η βία, η αυτοκτονία και η εγκληματικότητα, σε σύγκριση με μόλις το 15% των άρθρων που θεωρήθηκαν θετικά (το 11% αξιολογήθηκαν ως ουδέτερα) –η αναλογία ήταν επτά αρνητικά άρθρα προς ένα θετικό. Τέτοιες αφηγήσεις είναι, επίσης, ως επί το πλείστον αποσυνδεδεμένες από το κοινωνικό και πολιτισμικό τους πλαίσιο και αναλύονται απλουστευτικά, με τον αυτόχθονα πληθυσμό να συγκρίνεται συστηματικά με το (άρρητο) πρότυπο των μη-αυτοχθόνων (Walter & Andersen, 2013). Το αποτέλεσμα είναι ότι στο εθνικό φαντασιακό οι αυτόχθονες λαοί παρουσιάζονται υποτιμητικά –ως πρόβλημα, και όχι ως λαοί που φέρουν υπέρμετρο το φορτίο της ιστορικής και σύγχρονης ανισότητας.

Υπάρχουν όλο και περισσότερες ενδείξεις ότι οι φυλετικές προκαταλήψεις που ενσωματώνονται στα μεγάλα δεδομένα και οι αλγόριθμοι που αναπτύσσονται για την ανάλυσή τους θα ενισχύσουν, αντί να μειώσουν, τον αντίκτυπο των δεδομένων 5D στους αυτόχθονες πληθυσμούς (Henman, 2018). Έτσι, ενώ στα ιδιαίτερα ανεπτυγμένα αποικιοκρατικά κράτη, όπως η Αοτεαρόα-Νέα Ζηλανδία και η Αυστραλία, τα επιζήμια αποτελέσματα των πολιτικών που συνεπάγονται διακρίσεις έχουν υποχωρήσει, σε κάποιο βαθμό, χάρη στον ακτιβισμό των αυτοχθόνων και στα κινήματα κοινωνικής δικαιοσύνης σε βάθος πολλών ετών, αυτές οι αναδυόμενες πρακτικές δεδομένων μπορεί ακούσια να εδραιώσουν τις υπάρχουσες ανισότητες και να αναζωπυρώσουν παλαιότερα φαινόμενα. Με τον εντοπισμό (και τη βελτίωση) των κοινωνικών προβλημάτων να ανατίθεται πλέον, όλο και περισσότερο, σε αλγορίθμους, η πιθανότητα να επανέλθει η αδικία στο σύστημα με τρόπους που τοποθετούν σε μειονεκτική θέση τους αυτόχθονες πληθυσμούς αυξάνεται εκθετικά. Αναδιατυπώνοντας μια κλασική θεώρηση για τα δεδομένα, θα μπορούσαμε να πούμε ότι εάν οι «κανόνες» των δεδομένων του αλγορίθμου στοχεύουν σε προβλήματα όπου οι αυτόχθονες πληθυσμοί υπερεκπροσωπούνται, τότε ο στόχος θα είναι ο προβληματικός αυτόχθονας.

ID-SOV στην πράξη

Τα κινήματα ID-SOV είναι ενεργά στον Καναδά, στην Αυστραλία, στη Νέα Ζηλανδία και στις Ηνωμένες Πολιτείες και η επιρροή τους ολοένα και αυξάνεται. Οι πρωτοπόροι του ID-SOV είναι τα Πρώτα Έθνη του Καναδά. Έχοντας κουραστεί να αναλαμβάνουν οι μη αυτόχθονες χρήστες δεδομένων τον ρόλο του αμερόληπτου «ειδικού» για τους λαούς των Πρώτων Εθνών, οι ακτιβιστές της κοινότητας ανέπτυξαν ένα νέο μοντέλο που εξασφάλιζε τον συλλογικό έλεγχο των Πρώτων Εθνών στα δικά τους δεδομένα. Με βάση τις τέσσερις αρχές με το κατοχυρωμένο σήμα-κατατεθέν OCAP®[33], διεκδικούν το δικαίωμά τους να διατηρήσουν τη συλλογική ιδιοκτησία, τον έλεγχο, την πρόσβαση και την κατοχή των δεδομένων των Πρώτων Εθνών και, 20

33. ΣτΕ: Αρκτικόλεξο που σημαίνει Ownership (Ιδιοκτησία), Control (Ελεγχος), Access (Πρόσβαση) και Κατοχή (Possession).

χρόνια μετά, έχουν γίνει το ντε φάκτο πρότυπο για τη διεξαγωγή έρευνας με τα Πρώτα Έθνη (First Nations Information Governance Centre, 2016). Στην Αοτεαρόα-Νέα Ζηλανδία, το δίκτυο Te Mana Raraunga (TMR) για την κυριαρχία των Μαορί επί των δεδομένων τους ιδρύθηκε το 2015, συγκεντρώνοντας περισσότερους από 100 ερευνητές, επαγγελματίες και επιχειρηματίες Μαορί από τους τομείς της έρευνας, της πληροφορικής, της κοινότητας και των ΜΚΟ[34]. Το TMR έχει δραστηριοποιηθεί πολύ ενεργά στην προώθηση της ανάγκης για κυριαρχία των δεδομένων των Μαορί και της διακυβέρνησης των δεδομένων σε όλο τον δημόσιο τομέα και το 2018 έθεσε την εθνική στατιστική υπηρεσία προ των ευθυνών της για τον χειρισμό της απογραφής στη Νέα Ζηλανδία (Te Mana Raraunga, 2018b), κάτι που αναφέρθηκε ευρέως τόσο από τα συστημικά μέσα ενημέρωσης όσο και από εκείνα των αυτοχθόνων. Το TMR έχει, επίσης, εκφράσει ανησυχίες σχετικά με την «κοινωνική άδεια» για τη χρήση δεδομένων σε σχέση με τα δεδομένα των Μαορί (Te Mana Raraunga, 2017) και ανέπτυξε το δικό του σύνολο αρχών για την κυριαρχία των Μαορί επί των δεδομένων τους, για να ωθήσει προς την κατεύθυνση μιας ηθικής χρήσης αυτών (Te Mana Raraunga, 2018a). Για τους υποστηρικτές της κυριαρχίας των Μαορί επί των δεδομένων τους, συμπεριλαμβανομένου του TMR, ο στόχος δεν είναι μόνο η προστασία των ατόμων και των κοινοτήτων των Μαορί από μελλοντική βλάβη και στιγματισμό. Είναι και η διασφάλιση της γνώσης και των δικαιωμάτων πνευματικής ιδιοκτησίας των Μαορί, και η εξασφάλιση ότι οι επενδύσεις σε δημόσια δεδομένα δημιουργούν οφέλη και αξία με δίκαιο και ισότιμο τρόπο, στα οποία οι Μαορί μπορούν να συμμετέχουν πλήρως.

Στην Αυστραλία, το 2016 δημιουργήθηκε η συλλογικότητα Maiam nayri Wingara Indigenous Data Sovereignty Collective και το 2018, σε συνεργασία με το Αυστραλιανό Ινστιτούτο Διακυβέρνησης των Αυτοχθόνων, εξέδωσε ανακοίνωση από την εθνική συνάντηση των ηγετών των Αβορίγινων και των νησιωτών του Στενού Τόρες. Στο ανακοινωθέν, αναφέρεται η απαίτηση να ανήκουν στους ιθαγενείς οι αποφάσεις και ο έλεγχος του οικοσυστήματος δεδομένων –συμπεριλαμβανομένης της δημιουργίας, της ανάπτυξης, της διαχείρισης, της ανάλυσης, της διάδοσης και της υποδομής (Maiam nayri Wingara Indigenous Data Sovereignty Collective & Australian Indigenous Governance Institute, 2018).

Η Maiam nayri Wingara, μαζί με άλλους φορείς αυτοχθόνων, υποστηρίζει ενεργά αλλαγές στον τρόπο με τον οποίο τα δεδομένα των αυτοχθόνων στην Αυστραλία νοηματοδοτούνται, σχεδιάζονται, χρησιμοποιούνται, κατασκευάζονται, αναλύονται και ερμηνεύονται. Η φιλοδοξία είναι τα δεδομένα να συμβάλλουν ενεργά στην ευημερία των Αβορίγινων και των νησιωτών του Στενού Τόρες. Για να συμβεί αυτό, προϋπόθεση είναι μια εκ νέου κατασκευή της σχέσης μεταξύ των κατόχων/παραγωγών των δεδομένων των ιθαγενών και των αυτοχθόνων λαών τους οποίους αφορούν αυτά τα δεδομένα, η οποία θα βασίζεται στη διακυβέρνηση των δεδομένων των αυτοχθόνων.

Προς έναν σημαντικότερο ρόλο των πρωτοβουλιών ID-SOV στη δημοσιογραφία δεδομένων

Η δημοσιογραφία δεδομένων είναι σε μεγάλο βαθμό σε θέση να αμφισβητήσει, και όχι απλά να επαναπροσδιορίσει, τα 5D των δεδομένων των αυτοχθόνων. Οι δημοσιογράφοι δεδομένων έχουν πολλές ευκαιρίες να επανεξετάσουν τον τρόπο με τον οποίο χρησιμοποιούν τα δεδομένα, για να εκπροσωπήσουν τους αυτόχθονες λαούς και τις ιστορίες τους και να εκθέσουν τους πολύπλοκους τρόπους με τους οποίους τα δεδομένα των αυτοχθόνων παράγονται, ελέγχονται, διαδίδονται και «αξιοποιούνται» από την κυβέρνηση και τη βιομηχανία. Με τον τρόπο αυτό, οι δημοσιογράφοι δεδομένων δεν θα έπρεπε να βασίζονται σε μη αυτόχθονες παραγωγούς και χρήστες δεδομένων. Η ανάπτυξη των δικτύων ID-SOV σημαίνει ότι υπάρχει ένας αυξανόμενος αριθμός αυτοχθόνων εμπειρογνωμόνων δεδομένων στους οποίους μπορούν να απευθυν-

34. https://www.temanararaunga.maori.nz/tutohinga/

θούν. Πολλοί από όσους συμμετέχουν στο έργο των ID-SOV έχουν στενούς δεσμούς με τις κοινότητές τους και καθοδηγούνται από μια ισχυρή δέσμευση για δικαιοσύνη των δεδομένων και για την εξεύρεση τρόπων, ώστε τα «καλά δεδομένα» να ενισχύουν τα «καλά αποτελέσματα». Τα ερωτήματα που εγείρονται από το ID-SOV, ιδίως σχετικά με την κυριότητα των δεδομένων, τον έλεγχο, τη ζημία και το συλλογικό όφελος, έχουν ευρύτερη εφαρμογή πέρα από τις κοινότητες των αυτοχθόνων. Συμμετέχοντας στις προσεγγίσεις και στις αρχές του ID-SOV, οι δημοσιογράφοι δεδομένων μπορούν να ανοίξουν ουσιαστικούς χώρους για τις οπτικές και τις ανησυχίες των αυτοχθόνων που θα πλαισιώσουν τις αφηγήσεις τους, ενώ παράλληλα θα οξύνουν τη δημοσιογραφική τους αντίληψη, για να θέσουν τους κυβερνώντες προ των ευθυνών τους.

Παραπομπές

Anderson, I. et al. (2016) «Indigenous and tribal peoples' health (The Lancet–Lowitja Institute Global Collaboration): *A population study», The Lancet,* 388(10040), σσ. 131–157. DOI: https://doi.org/10.1016/S0140-6736(16)00345-7

Berda, Y. (2013) «Managing dangerous populations: Colonial legacies of security and surveillance», *Sociological Forum,* 28(3), σσ. 627-630. DOI: https://doi.org/10.1111/socf.12042

Briscoe, G. (2003) *Counting, health and identity: A history of aboriginal health and demography in Western Australia and Queensland 1900-1940.* Κανμπέρα: Aboriginal Studies Press.

Carroll, R. S., Rodriguez-Lonebear, D., & Martinez, A. (2017) *Policy brief (Version 2): Data governance for native nation rebuilding.* Τάκσον, Αριζόνα: Native Nations Institute. Διαθέσιμο σε: usindigenousdata.org/policy-briefs (Προσπελάστηκε: 21 Νοεμβρίου 2021).

Chesterman, J. & Galligan, B. (1997) *Citizens without rights: Aborigines and Australian citizenship.* Κέιμπριτζ: Cambridge University Press.

First Nations Information Governance Centre (2016). «Pathways to First Nations' data and information sovereignty», στο T. Kukutai & J. Taylor (επιμ.), *Indigenous data sovereignty: Toward an agenda,* Κανμπέρα: Australian National University Press, σσ. 139-155. DOI: https://doi.org/10.22459/CAEPR38.11.2016.08

Henman, P. (2018) «Of algorithms, apps and advice: Digital social policy and service delivery», *Journal of Asian Public Policy,* 12(2), σσ. 1-19. DOI: https://doi.org/10.1080/17516234.2018.1495885

Hudson, M., κ.α. (2020). «Rights, interests and expectations: Indigenous perspectives on unrestricted access to genomic data», *Nature Reviews Genetics,* 21(6), σσ. 377-384. DOI: https://doi.org/10.1038/s41576-020-0228-x

Hudson, M., Farrar, D., & McLean, L. (2016) «Tribal data sovereignty: Whakatōhea rights and interests», στο T. Kukutai & J. Taylor (επιμ.), *Indigenous data sovereignty: Toward an agenda,* Κανμπέρα: Australian National University Press, σσ. 157-178. DOI: https://doi.org/10.22459/CAEPR38.11.2016.09

Kukutai, T. & Taylor, J. (επιμ.) (2016) *Indigenous data sovereignty: Toward an agenda.* Κανμπέρα: Australian National University Press.

Maiam nayri Wingara Indigenous Data Sovereignty Collective & Australian Indigenous Governance Institute (2018). *Indigenous Data Sovereignty Summit Communique.* Κανμπέρα, Αυστραλία, 20 Ιουνίου 2018. Διαθέσιμο σε: https://www.aigi.com.au/wp-content/uploads/2019/10/Communique-Indigenous-Data-Sovereignty-Summit-1.pdf (Προσπελάστηκε: 21 Νοεμβρίου 2021).

Mullane-Ronaki, M.-T. T. K. K. (2017). *Indigenising the national census? A global study of the enumeration of indigenous peoples, 1985-2014*. Διπλωματική εργασία. University of Waikato.

Snipp, M. (2016) «What does data sovereignty imply: What does it look like?», στο T. Kukutai & J. Taylor (επιμ.), *Indigenous data sovereignty: Toward an agenda,* Κανμπέρα: Australian National University Press, σσ. 39-56.

Stephens, C., Porter, J., Nettleton, C. & Willis, R. (2006) «Disappearing, displaced, and undervalued: A call to action for Indigenous health worldwide», *The Lancet,* 367(9527), σσ. 2019-2028. DOI: https://doi.org/10.1016/S0140-6736(06)68892-2

Stoneham, M. (2014) «Bad news: Negative Indigenous health coverage reinforces stigma», *The Conversation,* 1 Απριλίου. Διαθέσιμο σε: http://theconversation.com/bad-news-negative-indigenous-health-coverage-reinforces-stigma-24851 (Προσπελάστηκε: 21 Νοεμβρίου 2021).

Te Mana Raraunga (2017). *Statement on social license.* Διαθέσιμο σε: https://www.temanararaunga.maori.nz/nga-panui (Προσπελάστηκε: 21 Νοεμβρίου 2021).

Te Mana Raraunga (2018a). *Principles of Māori data sovereignty.* Διαθέσιμο σε: https://www.temanararaunga.maori.nz/nga-panui (Προσπελάστηκε: 21 Νοεμβρίου 2021).

Te Mana Raraunga (2018b). *Te Mana Raraunga statement on 2018 New Zealand Census of Population and Dwellings: A call for action on Māori census data.* Διαθέσιμο σε: https://www.temanararaunga.maori.nz/nga-panui (Προσπελάστηκε: 21 Νοεμβρίου 2021).

United Nations (2009). *State of the world's indigenous peoples.* Διαθέσιμο σε: https://www.un.org/esa/socdev/unpfii/documents/SOWIP/en/SOWIP_web.pdf (Προσπελάστηκε: 21 Νοεμβρίου 2021).

Van Alsenoy, B., Ballet, J., Kuczerawy, A. & Dumortier, J. (2009) «Social networks and web 2.0: Are users also bound by data protection regulations?», Identity in the Information Society, 2(1), σσ. 65-79. DOI: https://doi.org/10.1007/s12394-009-0017-3

Walter, M. (2016) «Data politics and Indigenous representation in Australian statistics», στο T. Kukutai & J. Taylor (επιμ.), *Indigenous data sovereignty: Toward an agenda,* Κανμπέρα: Australian National University Press, σσ. 79-98.

Walter, M. (2018) «The voice of indigenous data: Beyond the markers of disadvantage», *Griffith Review,* 60, σ. 256.

Walter, M., & Andersen, C. (2013) *Indigenous statistics: A quantitative research methodology.* Γουόλνατ Κρικ, Καλιφόρνια: Left Coast Press.

Οι συγγραφείς

Η Τάχου Κουκουτάι (φυλή Ngāti Tiipa) είναι καθηγήτρια στο Εθνικό Ινστιτούτο Δημογραφικής και Οικονομικής Ανάλυσης του Πανεπιστημίου του Γουάικατο και ιδρυτικό μέλος του δικτύου Te Mana Raraunga για την κυριαρχία των Μαορί επί των δεδομένων τους.
Η Μάγκι Γουόλτερ (PhD) είναι κοινωνιολόγος από το Παλαουάν (Αβορίγινας της Τασμανίας) και ιδρυτικό μέλος της συλλογικότητας Maiam nayri Wingara Australian Data Sovereignty Data Collective.

3.4
Εναλλακτικές Πρακτικές Δεδομένων στην Κίνα

Yolanda Jinxin Ma

Περίληψη: Αυτό το κεφάλαιο παρουσιάζει μια εικόνα εκ των έσω για το τοπίο της δημοσιογραφίας δεδομένων στην Κίνα, τους βασικούς πρωταγωνιστές και την κουλτούρα των δεδομένων, καθώς και ορισμένες πρακτικές συμβουλές.

Λέξεις-κλειδιά: Κίνα, κουλτούρα δεδομένων, συμμετοχή των πολιτών, ανοιχτά δεδομένα, δημοσιογραφία δεδομένων, οπτικοποίηση δεδομένων

Πριν από μερικά χρόνια, έκανα μια παρουσίαση για τη δημοσιογραφία δεδομένων στην Κίνα, στο Google News Summit που διοργάνωνε το Google News Lab. Ήταν μια όμορφη χειμωνιάτικη μέρα στην καρδιά της Σίλικον Βάλεϊ και η αίθουσα ήταν κατάμεστη από περίπου 100 υψηλόβαθμους επαγγελματίες των μέσων ενημέρωσης, κυρίως από δυτικές χώρες. Ξεκίνησα ζητώντας τους να σηκώσουν το χέρι αν πιστεύουν, πρώτον, ότι δεν υπάρχουν καλά δεδομένα στην Κίνα και, δεύτερον, ότι δεν υπάρχει πραγματική δημοσιογραφία στην Κίνα. Και στις δύο περιπτώσεις σηκώθηκαν αρκετά χέρια και ακούστηκαν γέλια.

Αυτές είναι δύο κοινές πεποιθήσεις, εάν όχι προκαταλήψεις, που συναντώ συχνά, όταν παρευρίσκομαι ή μιλάω σε διεθνή δημοσιογραφικά συνέδρια. Εξ όσων έχω διαπιστώσει τα τελευταία έξι χρόνια, σε καμία περίπτωση δεν λείπουν τα δεδομένα. Στην πραγματικότητα, στην Κίνα παράγεται καθημερινά μια τεράστια ποσότητα δεδομένων, τα οποία βελτιώνονται ραγδαία σε ποιότητα και αποκτούν ευρύτερη κοινωνική σημασία. Και αντί της υποτιθέμενης απουσίας «πραγματικής» δημοσιογραφίας, υπάρχουν στην πραγματικότητα πολλοί δημοσιογράφοι που παράγουν καθημερινά σημαντικές ιστορίες, αν και δεν δημοσιεύονται τελικά όλες.

Δημιουργία δεδομένων με γνώμονα το θέμα

Οι ιστορίες δεδομένων παράγονταν ακόμη και πριν από την εισαγωγή του όρου «δημοσιογραφία δεδομένων» («data journalism») στην Κίνα. Ενώ σήμερα χρησιμοποιούμε συνήθως τον όρο «ιστορίες βασισμένες στα δεδομένα» («data-driven stories») στην Κίνα, υπήρξε μια περίοδος που είδαμε να συμβαίνει το αντίθετο: αντί τα δεδομένα να είναι ο οδηγός των ιστοριών, είδαμε ιστορίες ή συγκεκριμένα θέματα να οδηγούν στην παραγωγή δεδομένων. Αυτό συνέβαινε συνήθως σε θέματα που βρίσκουν απήχηση στους απλούς πολίτες, όπως η ατμοσφαιρική ρύπανση.

Από το 2010, το Υπουργείο Περιβάλλοντος δημοσιεύει έναν δείκτη ατμοσφαιρικής ρύπανσης σε πραγματικό χρόνο, αλλά από τις δημοσιεύσεις αρχικά απουσίαζε ένα σημαντικό στοιχείο[35]: δεν δημοσιεύονταν τα στοιχεία για τα αιωρούμενα σωματίδια («particulate matter», PM) ή τους ρύπους με διάμετρο μικρότερη από 2,5 μικρόμετρα ($PM_{2.5}$), οι οποίοι μπορούν να προκαλέσουν μη αναστρέψιμες βλάβες στον ανθρώπινο οργανισμό.

35. http://www.gov.cn/jrzg/2010-11/25/content_1753524.htm (στα κινεζικά, *Πρακτορείο Xinhua News*, «Εθνικό σύστημα για την ποιότητα του αέρα σε πραγματικό χρόνο στο Πεκίνο», 25 Νοεμβρίου 2010)

Δεδομένης της σοβαρότητας της ατμοσφαιρικής ρύπανσης και της έλλειψης επίσημων στοιχείων για τα $PM_{2.5}$, τον Νοέμβριο του 2011 ξεκίνησε μια πανεθνική εκστρατεία με τίτλο *I test the air for the motherland* («Ελέγχω τον αέρα για την πατρίδα»). Η εκστρατεία ενθάρρυνε κάθε πολίτη να συμβάλλει στην παρακολούθηση της ποιότητας του αέρα και να δημοσιεύει τα αποτελέσματα σε πλατφόρμες κοινωνικής δικτύωσης[36]. Η εκστρατεία διοργανώθηκε από μια μη κερδοσκοπική περιβαλλοντική οργάνωση. Ο εξοπλισμός ελέγχου χρηματοδοτήθηκε από τους πολίτες (δηλαδή, μέσω crowdfunding) και ο μη κερδοσκοπικός οργανισμός παρείχε εκπαίδευση στους ενδιαφερόμενους εθελοντές. Αυτή η κινητοποίηση απέκτησε ευρύτερη δυναμική μετά την ένωση των δυνάμεων μερικών διαδικτυακών influencer, συμπεριλαμβανομένου του Παν Σιγάι, ενός γνωστού επιχειρηματία, ο οποίος είχε τότε περισσότερους από επτά εκατομμύρια ακολούθους στο Sina Weibo, μία από τις πιο διαδεδομένες πλατφόρμες κοινωνικής δικτύωσης στην Κίνα (Page, 2011).

Μετά από δύο χρόνια δημόσιας εκστρατείας, τα δεδομένα για τα $PM_{2.5}$ συμπεριλήφθηκαν, τελικά, στις κυβερνητικές ανακοινώσεις. Ήταν μια καλή αρχή, αλλά οι προκλήσεις παρέμεναν. Οι αμφιβολίες σχετικά με την ακρίβεια των δεδομένων προκλήθηκαν από τις αποκλίσεις μεταξύ των δεδομένων που δημοσίευσε η κυβέρνηση και εκείνων που δημοσίευσε η πρεσβεία των ΗΠΑ στην Κίνα (Spegele, 2012).

Επιπλέον, τα δεδομένα δεν ήταν φιλικά προς τους δημοσιογράφους. Παρά τις ενημερώσεις ανά ώρα από περισσότερες από 100 πόλεις, οι πληροφορίες παρέχονταν μόνο σε κυλιόμενη βάση στην ιστοσελίδα, χωρίς δυνατότητα λήψης ενός συνόλου δεδομένων σε οποιοδήποτε φορμάτ. Μολονότι τα δεδομένα παρέχονται πλέον κεντρικά, τα παλαιότερα δεδομένα δεν είναι προσβάσιμα στο κοινό. Με άλλα λόγια, αν δεν μπορεί κανείς να γράψει ένα πρόγραμμα, για να συλλέγει τα δεδομένα ανά ώρα και να τα αποθηκεύει τοπικά, είναι αδύνατο να κάνει οποιαδήποτε ανάλυση των τάσεων με την πάροδο του χρόνου ή να προβεί σε συγκρίσεις μεταξύ πόλεων.

Αυτό δεν είναι το τέλος της ιστορίας. Η παραγωγή δεδομένων με γνώμονα το θέμα συνεχίζεται. Όταν τα δεδομένα δεν είναι καλά δομημένα και όταν οι δημοσιογράφοι δεδομένων δυσκολεύονται λόγω περιορισμένων τεχνικών δεξιοτήτων, τότε η κοινωνία των πολιτών και οι «κομπιουτεράδες» παρεμβαίνουν, για να παράσχουν υποστήριξη.

Ένα πρώτο παράδειγμα από το 2011 ήταν το PM25.in, το οποίο συλλέγει δεδομένα ατμοσφαιρικής ρύπανσης και τα δημοσιεύει σε καθαρό φορμάτ. Ο ιστότοπος ισχυρίζεται ότι έχουν πραγματοποιηθεί περισσότερες από ένα δισεκατομμύριο αναζητήσεις από τότε που ξεκίνησε να λειτουργεί[37]. Ένα άλλο παράδειγμα είναι η Qing Yue, μια μη κυβερνητική οργάνωση που συλλέγει και καθαρίζει περιβαλλοντικά δεδομένα από κυβερνητικούς ιστοτόπους σε όλα τα επίπεδα και, στη συνέχεια, τα δημοσιεύει στο κοινό σε φιλικές προς τον χρήστη μορφές. Τα επεξεργασμένα δεδομένα τους χρησιμοποιούνται ευρέως όχι μόνο από ομάδες δεδομένων σε καθιερωμένα μέσα ενημέρωσης αλλά και από τις ίδιες τις κυβερνητικές υπηρεσίες για την καλύτερη χάραξη πολιτικής.

Η παραγωγή δεδομένων και η αυξανόμενη ευαισθητοποίηση γύρω από ορισμένα ζητήματα βαδίζουν χέρι χέρι. Το 2015, ένα ντοκιμαντέρ που διερευνούσε τη σοβαρότητα της ατμοσφαιρικής ρύπανσης ξεσήκωσε τη χώρα. Η αυτοχρηματοδοτούμενη ταινία, με τίτλο *Under the Dome*, εξέθεσε την περιβαλλοντική κρίση της επιβλαβούς αιθαλομίχλης σε ολόκληρη τη χώρα και εντόπισε τις ρίζες του προβλήματος και τους διάφορους υπεύθυνους (Jing, 2015). Η ταινία έχει συγκριθεί με το *An Inconvenient Truth* του Αλ Γκορ τόσο ως προς το ύφος όσο και ως προς τον αντίκτυπο. Αυτή η αφήγηση περιείχε πολλά επιστημονικά δεδομένα, διαγράμματα που εξηγούσαν τις ετήσιες τάσεις και οπτικοποιήσεις κοινωνικών δικτύων για τη διαφθορά στις βιομηχανίες περιβάλλοντος και ενέργειας. Μόλις κυκλοφόρησε στο Διαδίκτυο, η ταινία έγινε viral και έφτασε τα 200 εκατομμύρια χτυπήματα μέσα σε τρεις ημέρες, πριν να λογοκριθεί και κατέβει

36. http://www.bjep.org.cn/pages/Index/40-1699?rid=2782 (στα κινεζικά)
37. https://chrome.google.com/webstore/detail/pm25by-bestapp-labs/bgjclhmlafjipakmliln-loihemmlfndj

μέσα σε μία εβδομάδα. Είχε, όμως, αυξήσει με επιτυχία την ευαισθητοποίηση του κοινού και είχε πυροδοτήσει μια εθνική συζήτηση για το θέμα, μεταξύ άλλων γύρω από την προσβασιμότητα και την ποιότητα των δεδομένων για την ατμοσφαιρική ρύπανση. Κατάφερε, επίσης, να γνωστοποιήσει με επιτυχία στην ηγεσία της χώρας τη σημασία του θέματος.

Δύο εβδομάδες μετά την κυκλοφορία του ντοκιμαντέρ, σε συνέντευξη Τύπου που παραχώρησε το Εθνικό Λαϊκό Συνέδριο, ο πρωθυπουργός Λι Κετσιάνγκ απάντησε σε ερώτηση σχετικά με την ατμοσφαιρική ρύπανση που αναφερόταν στην ταινία, παραδεχόμενος ότι η κυβέρνηση δεν κατάφερε να ικανοποιήσει τα αιτήματα του κοινού για τον περιορισμό της ρύπανσης. Αναγνώρισε ορισμένα από τα προβλήματα που έθιξε το ντοκιμαντέρ, συμπεριλαμβανομένης της χαλαρής εφαρμογής των κανονισμών για τον περιορισμό της ρύπανσης, και τόνισε ότι η κυβέρνηση θα επιβάλλει βαρύτερες ποινές, προκειμένου να μειωθεί το τοξικό νέφος (Wong & Buckley, 2015). Στα τέλη του Αυγούστου του 2015 εκδόθηκε ο νέος νόμος για την πρόληψη και τον έλεγχο της ατμοσφαιρικής ρύπανσης, ο οποίος τέθηκε σε εφαρμογή τον Ιανουάριο του 2016 (Lijian κ.α., 2015).

Η ατμοσφαιρική ρύπανση είναι μόνο ένα παράδειγμα που δείχνει ότι, ακόμη και όταν η διαθεσιμότητα ή η προσβασιμότητα των δεδομένων αποτελούν πρόκληση, η ανησυχία του κοινού για τα εκάστοτε ζητήματα μπορεί να οδηγήσει σε συνεισφορά των πολιτών στη δημιουργία δεδομένων, καθώς και σε αλλαγές στη στάση της κυβέρνησης και στη διαθεσιμότητα των δεδομένων του δημόσιου τομέα για τα εν λόγω ζητήματα. Σε πιο εδραιωμένα οικοσυστήματα, μπορεί τα δεδομένα να είναι πιο εύκολα διαθέσιμα και εύχρηστα, και η δουλειά του δημοσιογράφου πιο απλή: βρίσκει δεδομένα και τα χρησιμοποιεί ως βάση για θέματα. Στην Κίνα, η διαδικασία είναι ίσως λιγότερο γραμμική και οι πολίτες, η κυβέρνηση, η κοινωνία των πολιτών και τα μέσα ενημέρωσης αλληλεπιδρούν σε πολλαπλά στάδια αυτής της διαδικασίας. Τα δεδομένα, αντί να χρησιμεύουν απλώς ως αφετηρία για θέματα, μπορούν, επίσης, να εμφανιστούν σε μεταγενέστερο στάδιο, για να τροφοδοτήσουν νέου είδους σχέσεις μεταξύ των δημοσιογράφων και του κοινού τους.

Εξελισσόμενη κουλτούρα δεδομένων

Το περιβάλλον δεδομένων στην Κίνα έχει αλλάξει ραγδαία την τελευταία δεκαετία. Αυτό οφείλεται εν μέρει στη δυναμική που περιγράφηκε μέχρι στιγμής σε αυτό το κεφάλαιο και εν μέρει σε άλλους παράγοντες, όπως το παγκόσμιο κίνημα των ανοιχτών δεδομένων, οι ταχέως αναπτυσσόμενες εταιρείες του Διαδικτύου και το εκπληκτικά υψηλό ποσοστό διείσδυσης της κινητής τηλεφωνίας. Η κουλτούρα των δεδομένων, επίσης, εξελίσσεται γύρω από αυτές τις τάσεις.

Η κυβερνητική νομοθεσία αποτελεί τη ραχοκοκαλιά της πολιτικής για τη διαθεσιμότητα των δεδομένων. Προς έκπληξη πολλών, η Κίνα διαθέτει νόμους για την ελευθερία της πληροφόρησης. Οι Κανονισμοί του Κρατικού Συμβουλίου για τη Δημοσιοποίηση Κυβερνητικών Πληροφοριών εγκρίθηκαν το 2007 και τέθηκαν σε ισχύ την 1η Μαΐου 2008. Ο νόμος προβλέπει εντολή δημοσιοποίησης και επιβεβαιώνει τη δέσμευση για κυβερνητική διαφάνεια. Μετά τη θέσπιση του κανονισμού, οι κυβερνητικές υπηρεσίες σε όλα τα επίπεδα δημιούργησαν ειδικές ιστοσελίδες για τη δημοσιοποίηση των πληροφοριών που έχουν στη διάθεσή τους, συμπεριλαμβανομένων των συνόλων δεδομένων.

Ωστόσο, αν και ο νόμος έδωσε στους δημοσιογράφους το δικαίωμα να ζητούν ορισμένα δεδομένα ή πληροφορίες από τις Αρχές, μέσα στα πρώτα τρία χρόνια από την εφαρμογή του, δεν υπάρχουν γνωστές στο κοινό περιπτώσεις μέσων ενημέρωσης ή δημοσιογράφων που να έχουν ζητήσει την κοινοποίηση δεδομένων, σύμφωνα με μελέτη του 2011 που δημοσιεύθηκε από το Caixin, έναν όμιλο μέσων ενημέρωσης που εδρεύει στο Πεκίνο και φημίζεται για την άσκηση ερευνητικής δημοσιογραφίας[38]. Η μελέτη αποκάλυψε ότι, το 2010, η *Southern Weekly*, μια κορυφαία εφημερίδα, έλαβε ποσοστό ανταπόκρισης μόλις 44% σε αίτημα που είχε στείλει σε 29

38. http://finance.ifeng.com/leadership/gdsp/20110901/4512444.shtml (στα κινεζικά)

περιβαλλοντικά γραφεία, για να ελέγξει τον βαθμό συμμόρφωσής τους με τον νόμο. Οι οργανισμοί μέσων ενημέρωσης συνήθως δεν διαθέτουν νομική ομάδα ή άλλα συστήματα που να υποστηρίζουν τους δημοσιογράφους, για να συνεχίσουν τις έρευνές τους και να προωθήσουν τα αιτήματά τους για πληροφορίες. Σε μια άλλη περίπτωση, ένας δημοσιογράφος που, με ίδια πρωτοβουλία, προσέφυγε στο δικαστήριο ενάντια στην κυβέρνηση για τη μη δημοσιοποίηση πληροφοριών, κατέληξε να χάσει τη δουλειά του. Οι δυσκολίες και οι κίνδυνοι που αντιμετωπίζουν οι Κινέζοι δημοσιογράφοι, όταν επιχειρούν να αξιοποιήσουν τα νομικά εργαλεία, μπορεί να είναι πολύ μεγαλύτεροι από εκείνους που αντιμετωπίζουν οι δυτικοί συνάδελφοί τους.

Η Κίνα ανταποκρίνεται, επίσης, στο παγκόσμιο κίνημα των ανοιχτών δεδομένων και στο αυξανόμενο ενδιαφέρον για τα μεγάλα δεδομένα. Το 2012, τόσο η Σαγκάη όσο και το Πεκίνο εγκαινίασαν τις δικές τους πύλες ανοιχτών δεδομένων. Η καθεμία από αυτές διαθέτει εκατοντάδες σύνολα δεδομένων για θέματα όπως η χρήση γης, οι μεταφορές, η εκπαίδευση και η παρακολούθηση της ρύπανσης. Τα επόμενα χρόνια, δημιουργήθηκαν περισσότερες από δώδεκα πύλες ανοιχτών δεδομένων, όχι μόνο στις μεγαλύτερες πόλεις, αλλά και σε τοπικές περιφέρειες και σε λιγότερο ανεπτυγμένες επαρχίες. Η ανάπτυξη ήταν μάλλον «από κάτω προς τα πάνω», χωρίς πρότυπο ή τυποποιημένη δομή για την κοινοποίηση δεδομένων σε τοπικό επίπεδο, γεγονός που δεν συνέβαλε στην ευρύτερη συγκρισιμότητα ή χρηστικότητα αυτών των δεδομένων.

Μέχρι το 2015, το Κρατικό Συμβούλιο είχε δημοσιεύσει το Σχέδιο Δράσης για την Ανάπτυξη Μεγάλων Δεδομένων («Big Data Development Action Plan»), όπου τα ανοιχτά δεδομένα αναγνωρίστηκαν επίσημα ως ένα από τα δέκα βασικά εθνικά έργα και παρουσιάστηκε ένα συγκεκριμένο χρονοδιάγραμμα για το άνοιγμα των κυβερνητικών δεδομένων[39]. Ωστόσο, οι δημοσιογράφοι δεν ξεκινούν πάντοτε από τα επίσημα δεδομένα, τα οποία επίσης δεν ευθυγραμμίζονται πάντοτε με τα συμφέροντα και τις ανησυχίες του κοινού.

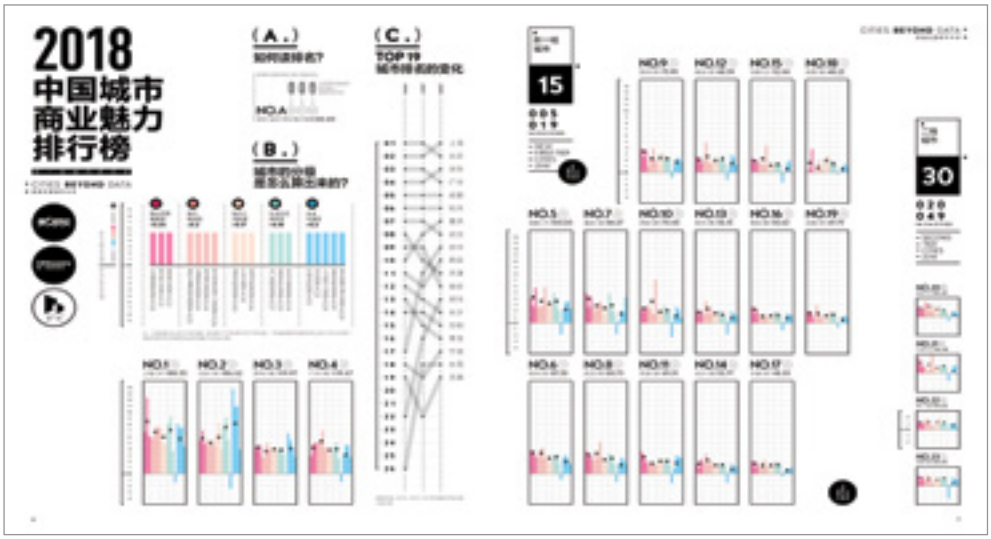
Από την άλλη πλευρά, ο ιδιωτικός τομέας, ιδίως οι τεχνολογικοί γίγαντες όπως η Alibaba ή η Tencent, έχουν συσσωρεύσει με την πάροδο των ετών τεράστιες ποσότητες δεδομένων. Σύμφωνα με τα τελευταία επίσημα στοιχεία της, οι ετήσιοι ενεργοί καταναλωτές της Alibaba είχαν φτάσει τα 601 εκατομμύρια έως τις 30 Σεπτεμβρίου 2018 (Alibaba Group Announces, 2018). Τα δεδομένα ηλεκτρονικού εμπορίου από μια τόσο ισχυρή βάση χρηστών –που ισοδυναμεί με ολόκληρο τον πληθυσμό της Νοτιοανατολικής Ασίας– μπορούν να αποκαλύψουν πολλές εμπορικές τάσεις, δημογραφικές μετατοπίσεις, κατευθύνσεις αστικής μετανάστευσης, αλλαγές στις καταναλωτικές συνήθειες και ούτω καθεξής. Υπάρχουν, επίσης, ιστότοποι με αξιολογήσεις χρηστών για θέματα ειδικού ενδιαφέροντος όπου είναι διαθέσιμα πιο συγκεκριμένα δεδομένα, όπως το Dianping, το κινεζικό αντίστοιχο του Yelp. Παρά τις ανησυχίες σχετικά με την προστασία της ιδιωτικότητας και της ασφάλειας, εάν χρησιμοποιηθούν σωστά, οι πλατφόρμες αυτές παρέχουν πλούσιους πόρους για εξόρυξη δεδομένων από δημοσιογράφους.

Ένα εξαιρετικό παράδειγμα αξιοποίησης των μεγάλων δεδομένων είναι το *Rising Lab*, μια ομάδα του Shanghai Media Group, που ειδικεύεται σε ιστορίες δεδομένων για την αστική ζωή[40]. Το Lab δημιουργήθηκε ως απάντηση στην αναδυόμενη τάση της αστικοποίησης: σύμφωνα με κυβερνητική έκθεση του 2016 (Gov't Report: China's Urbanization, 2016), η Κίνα διαθέτει σήμερα περισσότερες από 600 πόλεις, έναντι 193 το 1978, ενώ το 56% του πληθυσμού ζει σε αστικές περιοχές. Μαζί με την ταχεία αστικοποίηση αυξάνεται και η χρήση του Διαδικτύου και των κινητών τηλεφώνων, και επέρχονται αλλαγές στον τρόπο ζωής, όπως η ταχεία υιοθέτηση μοντέλων οικονομίας διαμοιρασμού. Αυτές οι τάσεις έχουν μεγάλο αντίκτυπο στη συγκέντρωση δεδομένων.

Με συμφωνίες συνεργασίας και τεχνική υποστήριξη από εταιρείες τεχνολογίας, το Lab συνέλεξε δεδομένα από ιστοτόπους και εφαρμογές που χρησιμοποιούν συχνά οι κάτοικοι της πόλης. Τα δεδο-

39. http://www.gov.cn/zhengce/content/2015-09/05/content_10137.htm (στα κινεζικά)
40. https://zhuanlan.zhihu.com/therisinglab (στα κινεζικά)

μένα αυτά αντανακλούσαν διάφορες πτυχές της αστικής ζωής, συμπεριλαμβανομένων των τιμών των ακινήτων, του αριθμού των καφετεριών και των μπαρ, του αριθμού των χώρων συνεργασίας και της ποιότητας των δημόσιων συγκοινωνιών. Παράλληλα με την αρχική του μεθοδολογία, το Lab δημιούργησε μια σειρά από κατατάξεις πόλεων που λαμβάνουν υπόψη πτυχές όπως η εμπορική ελκυστικότητα, το επίπεδο καινοτομίας και η ποικιλομορφία της ζωής (Εικόνα 3.4.1). Η κατάταξη και οι ιστορίες επικαιροποιούνται κάθε χρόνο με βάση τα νέα δεδομένα, αλλά ακολουθούν την ίδια μεθοδολογία, για να διασφαλίζεται η συνοχή. Η ιδέα και οι ιστορίες έχουν τύχει καλής υποδοχής από το κοινό και έχουν αρχίσει να επηρεάζουν τις πολιτικές αστικού σχεδιασμού και τις επιχειρηματικές αποφάσεις των εταιρειών, σύμφωνα με τον Σεν Κονγκλ, διευθυντή του Lab (Shen, 2018).

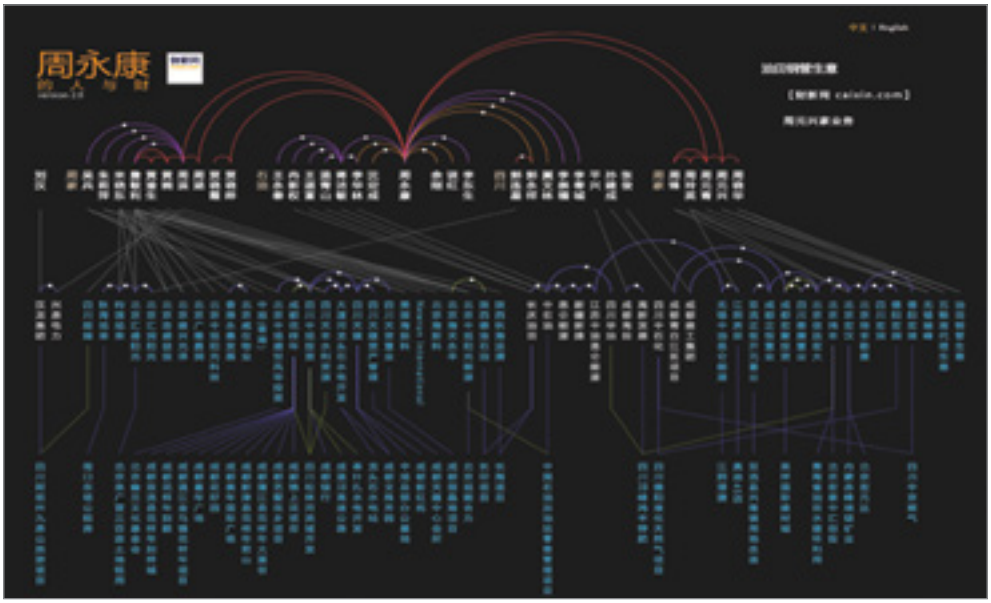


Εικόνα 3.4.1: Οι «Πόλεις πέρα από τα δεδομένα» το 2018 (*Cities beyond data*): κατάταξη της ελκυστικότητας των πόλεων για τις επιχειρήσεις στην Κίνα. Πηγή: *The Rising Lab*.

Η επιτυχία του Lab καταδεικνύει τις νέες δυναμικές που αναπτύσσονται ανάμεσα στους παρόχους δεδομένων, στους δημοσιογράφους και στους πολίτες. Δείχνει πώς πιο χαλαρά θέματα έχουν, επίσης, γίνει πεδίο δράσης για τη δημοσιογραφία δεδομένων, παράλληλα με άλλα πιεστικά ζητήματα, όπως η περιβαλλοντική κρίση, η διαφθορά, η δικαστική αδικία, η δημόσια υγεία και το ξέπλυμα χρήματος. Διερευνά, επίσης, νέα πιθανά επιχειρηματικά μοντέλα για τη δημοσιογραφία δεδομένων, καθώς και τον τρόπο με τον οποίο τα προϊόντα που βασίζονται σε δεδομένα μπορούν να προσδώσουν αξία σε κυβερνήσεις και επιχειρήσεις.

Οι συνήθειες των αναγνωστών ως προς την κατανάλωση ειδήσεων είχαν, επίσης, αντίκτυπο στην ανάπτυξη της δημοσιογραφίας δεδομένων. Δύο πτυχές χρήζουν προσοχής εδώ: η μία είναι η οπτική κατανάλωση ειδήσεων και η άλλη είναι η κατανάλωση ειδήσεων μέσω κινητών τηλεφώνων. Από το 2011, τα «ενημερωτικά γραφικά» («infographics») έχουν γίνει δημοφιλή χάρη στις προσπάθειες μερικών μεγάλων ειδησεογραφικών πόρταλ να δημιουργήσουν ειδικούς κόμβους με ιστορίες με infographics, κυρίως με βάση τα δεδομένα. Το 2014, η ιστορία της πτώσης του πρώην επικεφαλής ασφαλείας Ζου Γιονγκάνγκ, ενός από τους εννέα πιο υψηλόβαθμους πολιτικούς στην Κίνα, ήταν η μεγαλύτερη είδηση της χρονιάς. Μαζί με την είδηση, το Caixin παρήγαγε μια διαδραστική οπτικοποίηση για τα μέσα κοινωνικής δικτύωσης (Εικόνα 3.4.2), προκειμένου να απεικονίσει το πολύπλοκο δίκτυο γύρω από τον Ζου, συμπεριλαμβανομένων 37 ατόμων και 105 εταιρειών ή έργων που συνδέονται με αυτόν, καθώς και τη σχέση

μεταξύ αυτών των οντοτήτων –όλα βασισμένα στο ερευνητικό κομμάτι των 60.000 λέξεων που παρήγαγε η δημοσιογραφική του ομάδα. Αυτό το διαδραστικό θέμα έλαβε τέσσερα εκατομμύρια χτυπήματα μέσα σε μία εβδομάδα και άλλα 20 εκατομμύρια προβολές στα μέσα κοινωνικής δικτύωσης, σύμφωνα με το Caixin[41]. Η ευρεία κυκλοφορία αυτού του έργου έφερε νέα είδη αφήγησης δεδομένων σε νέες ομάδες κοινού και άνοιξε την όρεξη για οπτικές ιστορίες.



Εικόνα 3.4.2: Ένα διαδραστικό κομμάτι που αποτυπώνει τα δίκτυα του Ζου Γιονγκάνγκ. Πηγή: Caixin.

Σχεδόν την ίδια στιγμή, η βιομηχανία των μέσων ενημέρωσης καλωσόριζε την εποχή της κινητής τηλεφωνίας. Όλο και περισσότερες ιστορίες δεδομένων, όπως και κάθε άλλο διαδικτυακό περιεχόμενο στην Κίνα, πλέον διαδίδονται κυρίως μέσω κινητών τηλεφώνων. Σύμφωνα με το Κέντρο Διαδικτυακών Πληροφοριών της Κίνας («China Internet Network Information Center», CNNIC), άνω του 95% των χρηστών του Διαδικτύου στη χώρα χρησιμοποίησαν κάποια κινητή συσκευή για πρόσβαση στο Διαδίκτυο το 2016 (Chung, 2017). Το WeChat, η εγχώρια δημοφιλής εφαρμογή ανταλλαγής μηνυμάτων και πλατφόρμα κοινωνικής δικτύωσης, έφτασε το ένα δισεκατομμύριο χρήστες τον Μάρτιο του 2018 (Jao, 2018). Η κυριαρχία των κινητών συσκευών σημαίνει ότι οι ιστορίες δεδομένων στην Κίνα όχι απλά προσανατολίζονται πλέον κυρίως προς τις κινητές συσκευές, αλλά, σε αρκετές περιπτώσεις, προσανατολίζονται αποκλειστικά εκεί. Αυτή η ζήτηση της αγοράς οδήγησε σε πολλές λιτές, απλές και μερικές φορές δημιουργικές διαδραστικές εφαρμογές που είναι φιλικές προς τις κινητές συσκευές.

Εν ολίγοις, η κουλτούρα των δεδομένων στην Κίνα εξελίσσεται, καθοδηγούμενη από διάφορους παράγοντες, από τα παγκόσμια κινήματα έως την κυβερνητική νομοθεσία, από τη δημόσια ζήτηση έως τα αιτήματα των μέσων ενημέρωσης, από τις νέες γενιές παρόχων δεδομένων έως τις νέες γενιές καταναλωτών ειδήσεων. Οι αλληλεξαρτώμενες σχέσεις ανάμεσα στους διάφορους παράγοντες έχουν δημιουργήσει μια πολύ σύνθετη δυναμική, όπου συνυπάρχουν περιορισμοί και ευκαιρίες. Η δημοσιογραφία δεδομένων άνθισε και ακολούθησε το δικό της μονοπάτι στην Κίνα.

41. http://bit.ly/storybench-caixin (στα κινεζικά· *Storybench*, «Πώς δημιουργήθηκε ένα προϊόν δημοσιογραφίας δεδομένων με τέσσερα εκατομμύρια κλικ», 10 Οκτωβρίου 2017)

Πρακτικές συμβουλές

Αυτή η τελευταία ενότητα απευθύνεται σε αναγνώστες αυτού του βιβλίου οι οποίοι επιθυμούν να ασχοληθούν με ιστορίες σχετικά με την Κίνα και αναρωτιούνται από πού να ξεκινήσουν. Δεν θα είναι εύκολο. Εάν δεν μιλάτε κινεζικά, θα αντιμετωπίσετε γλωσσικά εμπόδια, καθώς οι περισσότερες πηγές δεδομένων είναι διαθέσιμες μόνο στα κινεζικά. Στη συνέχεια, θα αντιμετωπίσετε συνήθη προβλήματα που σχετίζονται με την εργασία με δεδομένα: ακρίβεια δεδομένων, πληρότητα δεδομένων, ασυνέπεια δεδομένων κλπ, αλλά θα υποθέσουμε ότι, ως αναγνώστης αυτού του βιβλίου, διαθέτετε τις δεξιότητες για να αντιμετωπίσετε αυτά τα ζητήματα, ή τουλάχιστον την προθυμία να μάθετε πώς να το κάνετε. Ένας καλός τρόπος για να ξεκινήσετε θα ήταν να εντοπίσετε τους σημαντικότερους πρωταγωνιστές στη δημοσιογραφία δεδομένων στην Κίνα. Αρκετά από τα κορυφαία μέσα ενημέρωσης διαθέτουν ομάδες δεδομένων και είναι καλό να παρακολουθήσετε τις ιστορίες τους και να μιλήσετε στους δημοσιογράφους για συμβουλές. Να ορισμένα μέσα που θα πρέπει να γνωρίζετε: το *Data Visualisation Lab* (Caixin), το *Beautiful Data Channel* (The Paper), το *Rising Lab* (Shanghai Media Group) και το *DT Finance*[42].

Η δεύτερη ερώτηση αφορά το πού θα βρείτε τα δεδομένα. Ένας ολοκληρωμένος κατάλογος πηγών δεδομένων θα αποτελούσε ξεχωριστό βιβλίο. Επομένως, παραθέτουμε μερικές μόνο προτάσεις για να ξεκινήσετε. Αρχίστε με κυβερνητικούς ιστοτόπους, τόσο από κεντρικά υπουργεία όσο και περιφερειακές υπηρεσίες. Θα πρέπει να γνωρίζετε ποια υπηρεσία είναι η κατάλληλη για τα δεδομένα που αναζητάτε και θα πρέπει να ελέγξετε τόσο τους θεματικούς τομείς των υπουργείων (για παράδειγμα, το Υπουργείο Προστασίας του Περιβάλλοντος) όσο και τον ειδικό δικτυακό τόπο δεδομένων σε τοπικό επίπεδο, εφόσον υπάρχει κάτι τέτοιο.

Θα βρείτε τα πιο απρόσμενα δεδομένα –για παράδειγμα, θα περιμένατε ότι η κινεζική κυβέρνηση δημοσιεύει το πλήρες κείμενο εκατομμυρίων δικαστικών αποφάσεων μετά το 2014; Τα νομικά έγγραφα είναι σχετικά διαφανή στις Ηνωμένες Πολιτείες αλλά όχι στην Κίνα. Ωστόσο, το Ανώτατο Λαϊκό Δικαστήριο («Supreme People's Court», SPC) ξεκίνησε μια βάση δεδομένων που ονομάζεται China Judgments Online και κάνει ακριβώς αυτό.

Μόλις βρείτε στο Διαδίκτυο κάποια δεδομένα που θα μπορούσαν να σας φανούν χρήσιμα, φροντίστε να τα «κατεβάσετε» σε ένα τοπικό αντίγραφο. Παραμένει σύνηθες φαινόμενο τα δεδομένα να μην είναι διαθέσιμα στο Διαδίκτυο. Μερικές φορές τα δεδομένα δημοσιεύονται με τη μορφή ετήσιων κυβερνητικών εκθέσεων τις οποίες μπορείτε να παραγγείλετε ηλεκτρονικά ή είναι διαθέσιμα μόνο σε έντυπα αρχεία. Για παράδειγμα, ορισμένες κυβερνητικές υπηρεσίες διαθέτουν τα αρχεία ιδιωτικών εταιρειών, αλλά δεν είναι όλα διαθέσιμα στο Διαδίκτυο.

Εάν τα δεδομένα δεν δημοσιεύονται καθόλου από την κυβέρνηση, ελέγξτε αν υπάρχει διαθέσιμο περιεχόμενο που δημιουργούν οι χρήστες. Για παράδειγμα, τα δεδομένα για τη δημόσια υγεία είναι πολύ περιορισμένα, αλλά υπάρχουν, μεταξύ άλλων, ειδικοί ιστότοποι με πληροφορίες σχετικά με τις εισαγωγές στα νοσοκομεία ή στα κέντρα ηλικιωμένων. Η αναζήτηση και ο καθαρισμός αυτών των δεδομένων θα σας βοηθούσε να αποκτήσετε μια καλή εικόνα για το θέμα.

Συνιστάται, επίσης, η χρήση βάσεων δεδομένων του Χονγκ Κονγκ –από τις επίσημες, όπως το μητρώο εταιρειών του Χονγκ Κονγκ, μέχρι τις ανεξάρτητες, όπως τα Webb-site Reports. Καθώς η ηπειρωτική Κίνα και το Χονγκ Κονγκ έχουν έρθει πιο κοντά πολιτικά και οικονομικά, υπάρχουν περισσότερες πληροφορίες εκεί, χάρη στο διαφανές περιβάλλον του Χονγκ Κονγκ και την εφαρμογή της νομοθεσίας, οι οποίες μπορεί να είναι πολύτιμες για την παρακολούθηση του χρήματος.

Υπάρχουν, επίσης, στοιχεία για την Κίνα που δεν βρίσκονται απαραίτητα στην Κίνα. Υπάρχουν διεθνείς οργανισμοί ή ακαδημαϊκά ιδρύματα που διαθέτουν πλούσια σύνολα δεδομένων σχε-

42. http://datanews.caixin.com/, https://www.thepaper.cn/list_25635, https://www.cbnweek.com/topics/10, https://www.dtcj.com/ (όλα στα κινεζικά)

τικά με την Κίνα. Για παράδειγμα, ο ιστότοπος The Paper χρησιμοποίησε δεδομένα από τη NASA και το Πανεπιστήμιο του Χάρβαρντ σε ένα από τα πιο πρόσφατα θέματά του.

Τέλος, αν και ορισμένες προκλήσεις και εμπειρίες είναι μοναδικές για την Κίνα, πολλές από αυτές θα μπορούσαν ενδεχομένως να αποτελέσουν χρήσιμα μαθήματα για τους δημοσιογράφους άλλων χωρών, όπου οι κοινωνικές, πολιτιστικές και πολιτικές συνθήκες έχουν διαφορετική μορφή αλλά παρόμοιους περιορισμούς.

Παραπομπές

Business Wire (2018) «Alibaba Group announces September quarter 2018 results», 2 Νοεμβρίου. Διαθέσιμο σε: https://www.businesswire.com/news/home/20181102005230/en/Alibaba-Group-Announces-September-Quarter-2018-Results/ (Προσπελάστηκε: 23 Νοεμβρίου 2021).

Chung, M. C. (2017) «More than 95% of Internet users in China use mobile devices to go online», *EMarketer*, 2 Φεβρουαρίου. Διαθέσιμο σε: https://www.emarketer.com/Article/More-than-95-of-Internet-Users-China-Use-Mobile-Devices-Go-Online/1015155 (Προσπελάστηκε: 23 Νοεμβρίου 2021).

Gov't report: China's urbanization level reached 56.1% (2016), CNTV, 20 Απριλίου. Διαθέσιμο σε: http://english.www.gov.cn/news/video/2016/04/20/content_281475331447793.htm (Προσπελάστηκε: 23 Νοεμβρίου 2021).

Jao, N. (2018) «WeChat now has over 1 billion active monthly users worldwide», *TechNode*, 5 Μαρτίου. Διαθέσιμο σε: https://technode.com/2018/03/05/wechat-1-billion-users/ (Προσπελάστηκε: 23 Νοεμβρίου 2021).

Jing, C. (2015) *Chai Jing's review: Under the dome—investigating China's smog [YouTube],* 28 Φεβρουαρίου. Διαθέσιμο σε: https://www.youtube.com/watch?v=T6X2uwlQGQM (Προσπελάστηκε: 23 Νοεμβρίου 2021).

Lijian, Z., Xie, T., & Tang, J. (2015) «How China's new air law aims to curb pollution», *China Dialogue,* 30 Δεκεμβρίου. Διαθέσιμο σε: https://www.chinadialogue.net/article/show/single/en/8512-How-China-s-new-air-law-aims-to-curb-pollution (Προσπελάστηκε: 23 Νοεμβρίου 2021).

Page, J. (2011). «Microbloggers pressure Beijing to improve air pollution monitoring», *The Wall Street Journal,* 8 Νοεμβρίου. Διαθέσιμο σε: https://blogs.wsj.com/chinarealtime/2011/11/08/internet-puts-pressure-on-beijing-to-improve-air-pollution-monitoring/ (Προσπελάστηκε: 23 Νοεμβρίου 2021)

Shen, J. (2018) *Data journalism in China panel, Uncovering Asia, Investigative Journalism Conference* 2018. Σεούλ, Κορέα, 5-7 Οκτωβρίου. Διαθέσιμο σε: https://2018.uncoveringasia.org/schedule/ (Προσπελάστηκε: 23 Νοεμβρίου 2021).

Spegele, B. (2012), «Comparing pollution data: Beijing vs. U.S. embassy on PM2.5», *The Wall Street Journal,* 23 Ιανουαρίου. Διαθέσιμο σε: https://blogs.wsj.com/chinarealtime/2012/01/23/comparing-pollution-data-beijing-vs-u-s-embassy-on-pm2-5/ (Προσπελάστηκε 23 Νοεμβρίου 2021).

Wong, E., & Buckley, C. (2015) «Chinese premier vows tougher regulation on air pollution», *The New York Times,* 15 Μαρτίου. Διαθέσιμο σε: https://www.nytimes.com/2015/03/16/world/asia/chinese-premier-li-keqiang-vows-tougher-regulation-on-air-pollution.html (Προσπελάστηκε: 23 Νοεμβρίου 2021).

Η συγγραφέας

Η Γιολάντα Τζίνξιν Μα είναι ειδική σε θέματα ψηφιακού μετασχηματισμού στον ΟΗΕ και νωρίτερα ήταν επαγγελματίας και υπέρμαχος της δημοσιογραφίας δεδομένων την οποία εισήγαγε στην Κίνα και σε ολόκληρη την Ασία.

3.5

Οικοδόμηση μιας Βάσης Δεδομένων για την Καταγραφή των Συγκρούσεων για τη Γη στην Ινδία

Kumar Sambhav Shrivastava, Ankur Paliwal

Περίληψη: Καταγραφή των συγκρούσεων για τη γη προκειμένου να αφηγηθούμε βαθύτερες, ουσιαστικότερες και σημαντικότερες ιστορίες για τη γη και τη σχέση της με την ποικιλόμορφη κοινωνία της Ινδίας.

Λέξεις-κλειδιά: σύγκρουση για τη γη, Ινδία, δημοσιογραφία δεδομένων, βάσεις δεδομένων, συνεργασία

Η γη είναι ένας ανεπαρκής πόρος στην Ινδία. Η χώρα διαθέτει μόλις το 2,4% της παγκόσμιας έκτασης, αλλά φιλοξενεί πάνω από το 17% του παγκόσμιου πληθυσμού. Ως μία από τις ταχύτερα αναπτυσσόμενες οικονομίες του κόσμου, χρειάζεται μεγάλες εκτάσεις γης για να τροφοδοτήσει τη φιλόδοξη ατζέντα τής βιομηχανικής ανάπτυξης και των υποδομών της. Τουλάχιστον 110 εκατομμύρια στρέμματα γης απαιτούνται για αναπτυξιακά έργα τα επόμενα 15 χρόνια. Όμως, ένα τεράστιο τμήμα του πληθυσμού της Ινδίας –κυρίως περιθωριοποιημένες κοινότητες– εξαρτώνται από τη γη για τη συντήρηση και τον βιοπορισμό τους. Περισσότεροι από 200 εκατομμύρια άνθρωποι εξαρτώνται από τα δάση, ενώ 118,9 εκατομμύρια εξαρτώνται από γεωργικές εκτάσεις στην Ινδία.

Οι αντικρουόμενες αυτές ανάγκες προκαλούν συγκρούσεις. Σε πολλές περιπτώσεις, η γη αποκτάται με τη βία ή με δόλο από το κράτος ή από ιδιωτικά συμφέροντα. Οι διαφωνούντες συλλαμβάνονται από κρατικές υπηρεσίες με ψευδείς κατηγορίες, οι αποζημιώσεις καταβάλλονται μερικώς, κοινότητες εκτοπίζονται, σπίτια πυρπολούνται και άνθρωποι σκοτώνονται. Οι κοινωνικές ανισότητες γύρω από την κάστα, την τάξη και το φύλο πυροδοτούν, επίσης, τους αγώνες για τη γη. Οι καταστροφές που προκαλούνται από την κλιματική αλλαγή καθιστούν τις κοινότητες που εξαρτώνται από τη γη ακόμη πιο ευάλωτες στους εκτοπισμούς. Όλα αυτά αντικατοπτρίζονται στις πολλές μάχες που λαμβάνουν χώρα για τη γη σε ολόκληρη την Ινδία.

Ως δημοσιογράφοι που γράφουμε για θέματα ανάπτυξης στην Ινδία, συναντάμε πολλές τέτοιες συγκρούσεις. Ωστόσο, συνειδητοποιήσαμε ότι δεν ήταν εύκολο να πουλήσουμε αυτές τις ιστορίες, οι οποίες διαδραματίζονται σε απομακρυσμένες γωνιές της Ινδίας, στους εκδότες στο Νέο Δελχί. Τα κυρίαρχα μέσα ενημέρωσης δεν ανέφεραν συγκρούσεις για τη γη, πέρα από όσες κατέληγαν σε θανάσιμη βία ή όσες εκδικάστηκαν στα εθνικά δικαστήρια. Τα σποραδικά ρεπορτάζ από λίγους δημοσιογράφους είχαν μικρό αντίκτυπο. Τις φωνές των ανθρώπων που επλήγησαν από τέτοιες συγκρούσεις δεν τις άκουγε κανείς. Οι ανησυχίες τους παρέμεναν χωρίς ανταπόκριση.

Σκεφτήκαμε ότι ο λόγος ήταν πως οι δημοσιογράφοι και οι εκδότες αντιμετώπιζαν τις συγκρούσεις ως μεμονωμένα περιστατικά. Γνωρίζαμε ότι οι συγκρούσεις για τη γη ήταν ένα από τα πιο σημαντικά θέματα για την πολιτική οικονομία της Ινδίας. Το ερώτημα, όμως, ήταν πώς θα γινόταν να πουλήσουμε αυτά τα θέματα στους εκδότες και, κατ' επέκταση, στους αναγνώστες.

Σκεφτήκαμε ότι αν οι δημοσιογράφοι μπορούσαν να αυξήσουν τις ανταποκρίσεις τους για μεμονωμένες περιπτώσεις συγκρούσεων, ώστε να εξετάσουν ευρύτερες τάσεις, οι ιστορίες τους όχι μόνο θα είχαν μεγαλύτερη απήχηση, αλλά θα μπορούσαν, επίσης, να δείξουν την ένταση των διαφόρων ειδών συγκρούσεων και τον αντίκτυπό τους στους ανθρώπους, στην οικονομία και στο περιβάλλον. Η μεγαλύτερη πρόκληση για την επίτευξη αυτού του στόχου ήταν η έλλειψη μιας βάσης δεδομένων την οποία οι δημοσιογράφοι θα μπορούσαν να εξερευνήσουν, για να δουν ποιες τάσεις αναδύονται γύρω από συγκεκριμένα είδη συγκρούσεων, όπως αυτές που αφορούν δρόμους, δήμους, εξορύξεις ή περιοχές προστασίας της άγριας πανίδας. Δεν υπήρχε τέτοια βάση δεδομένων για τις συνεχιζόμενες συγκρούσεις για τη γη στην Ινδία. Αποφασίσαμε, λοιπόν, να φτιάξουμε μία.

Τον Νοέμβριο του 2016, ξεκινήσαμε το Land Conflict Watch, ένα ερευνητικό έργο δημοσιογραφίας δεδομένων που στόχο έχει να χαρτογραφήσει και να τεκμηριώσει όλες τις συνεχιζόμενες συγκρούσεις για τη γη στην Ινδία. Αναπτύξαμε μια μεθοδολογία τεκμηρίωσης, σε συνεννόηση με ακαδημαϊκούς οι οποίοι ασχολούνται με τη διαχείριση της γης. Συγκεντρώσαμε ένα δίκτυο ερευνητών και δημοσιογράφων, από όλη τη χώρα, για να καταγράψουν τις συγκρούσεις στις περιοχές τους, ακολουθώντας αυτή τη μεθοδολογία.

Για τους σκοπούς αυτού του έργου, ορίσαμε τις συγκρούσεις για τη γη ως οποιαδήποτε κατάσταση που περιλαμβάνει αντικρουόμενες απαιτήσεις ή διεκδικήσεις σχετικά με τη χρήση ή την ιδιοκτησία της γης και όπου οι κοινότητες είναι ένα από τα αντιμαχόμενα μέρη. Περιλαμβάνονται οι συνεχιζόμενες συγκρούσεις στις οποίες τα αιτήματα ή οι διεκδικήσεις έχουν ήδη καταγραφεί σε γραπτή ή σε οπτικοακουστική μορφή σε οποιοδήποτε μέρος, από ένα χωριό έως σε εθνικό επίπεδο. Αυτές οι καταγραφές μπορεί να είναι ειδησεογραφικές αναφορές, ψηφίσματα συνελεύσεων χωριών, αρχεία δημόσιων διαβουλεύσεων για αναπτυξιακά έργα, καταγγελίες που υποβάλλονται από τους πολίτες στις κυβερνητικές αρχές, αστυνομικά αρχεία ή δικαστικά έγγραφα. Συγκρούσεις όπως περιουσιακές διαφορές μεταξύ δύο ιδιωτών ή μεταξύ ενός ιδιώτη και της κυβέρνησης αποκλείονται, εκτός εάν επηρεάζουν άμεσα το ευρύτερο κοινό.

Οι ερευνητές και οι δημοσιογράφοι παρακολουθούν την κάλυψη των εθνικών και τοπικών μέσων ενημέρωσης για τις περιοχές τους και αλληλεπιδρούν με τοπικούς ακτιβιστές, κοινοτικές οργανώσεις και δικηγόρους, προκειμένου να εντοπίσουν περιπτώσεις συγκρούσεων. Στη συνέχεια, συλλέγουν και επαληθεύουν πληροφορίες από δημοσίως διαθέσιμα κυβερνητικά έγγραφα και ανεξάρτητες μελέτες, καθώς και από συζητήσεις με τα ενδιαφερόμενα μέρη. Καταγράφονται δεδομένα όπως η τοποθεσία της σύγκρουσης, οι λόγοι της σύγκρουσης, ο αριθμός των πληγέντων ατόμων, η πληγείσα περιοχή, ο τύπος γης –ιδιωτική, κοινόχρηστη ή δασική–, τα ονόματα των εμπλεκόμενων κυβερνητικών και εταιρικών φορέων και μια αφηγηματική περίληψη της σύγκρουσης.

Οι ερευνητές αρχειοθετούν όλα τα δεδομένα στο λογισμικό αναφοράς και αξιολόγησης που είναι ενσωματωμένο στον ιστότοπο του Land Conflict Watch. Τα δεδομένα εξετάζονται και επαληθεύονται από ειδικούς αξιολογητές. Το λογισμικό επιτρέπει τη ροή εργασίας μεταξύ των ερευνητών και των αξιολογητών πριν από τη δημοσίευση των δεδομένων. Το dashboard του πόρταλ δεν παρουσιάζει μόνο τη μακροσκοπική εικόνα των συνεχιζόμενων συγκρούσεων σε εθνικό επίπεδο, αλλά μεγεθύνεται για να δώσει λεπτομέρειες για κάθε σύγκρουση, μαζί με τα συνοδευτικά έγγραφα, σε μικροσκοπικό επίπεδο. Παρέχει, επίσης, την κατά προσέγγιση γεωγραφική θέση της σύγκρουσης σε έναν διαδραστικό χάρτη.

Αυτή τη στιγμή στην πλατφόρμα συνεισφέρουν περίπου 35 δημοσιογράφοι και ερευνητές. Μέχρι τον Σεπτέμβριο του 2018, το έργο είχε καταγράψει πάνω από 640 περιστατικά. Αυτές οι συγκρούσεις επηρεάζουν περίπου 7,3 εκατομμύρια ανθρώπους και εκτείνονται σε 24 εκατομμύρια στρέμματα γης. Επενδύσεις αξίας 186 δισεκατομμυρίων δολαρίων συνδέονται με έργα και προγράμματα που επηρεάζονται από αυτές τις συγκρούσεις.

Καθώς μια σύγκρουση καταγράφεται, περιγράφεται στο πόρταλ και στα μέσα κοινωνικής δικτύωσης, για να ενημερωθούν οι δημοσιογράφοι και οι ερευνητές σε εθνικό επίπεδο. Στη συνέχεια, η ομάδα του έργου συνεργάζεται με δημοσιογράφους για τη δημιουργία εμπεριστατωμένων, ερευνητικών θεμάτων που αφορούν την αλληλεπίδραση των δικαιωμάτων γης, των συγκρούσεων για τη γη, της πολιτικής, της οικονομίας, της τάξης, του φύλου και του περιβάλλοντος, χρησιμοποιώντας αυτά τα δεδομένα. Συνεργαζόμαστε, επίσης, με εθνικά και διεθνή μέσα ενημέρωσης για να δημοσιεύσουμε αυτές τις ιστορίες. Πολλές από αυτές έχουν αναδημοσιευθεί από άλλα κυρίαρχα μέσα ενημέρωσης. Έχουμε, επίσης, πραγματοποιήσει εκπαιδευτικές δράσεις, για να υποστηρίξουμε τους δημοσιογράφους στη χρήση της βάσης δεδομένων για την εύρεση και ανάπτυξη ιστοριών γύρω από τη διαχείριση της γης.

Το Land Conflict Watch είναι ένα έργο που βρίσκεται σε εξέλιξη. Εκτός από τον σχεδιασμό ιστοριών, συνεργαζόμαστε επίσης με ακαδημαϊκούς, ερευνητές και φοιτητές για να δρομολογήσουμε δημόσιες συζητήσεις. Τα δεδομένα του Land Conflict Watch έχουν παρατεθεί από πολιτικές ομάδες προβληματισμού στις εκθέσεις τους. Ειδικοί σε θέματα διαχείρισης της γης έχουν γράψει άρθρα γνώμης σε εθνικής εμβέλειας εφημερίδες, χρησιμοποιώντας αυτά τα δεδομένα. Λαμβάνουμε τακτικά αιτήματα από φοιτητές ινδικών και ξένων πανεπιστημίων για τη χρήση των δεδομένων μας στην έρευνά τους. Διάφορες μη κερδοσκοπικές οργανώσεις χρησιμοποιούν δεδομένα, έγγραφα και υποθέσεις σχετικά με τις συγκρούσεις για τη γη, προκειμένου να ενισχύσουν τις εκστρατείες τους για την υπεράσπιση των εδαφικών δικαιωμάτων των κοινοτήτων που πλήττονται από αυτές τις συγκρούσεις. Οι ιστορίες ενημερώνουν τους ανθρώπους και συμβάλλουν στη διαμόρφωση της συζήτησης γύρω από τα δικαιώματα γης και τα ζητήματα που σχετίζονται με τη διακυβέρνηση στην Ινδία.

Οι συγγραφείς

Ο Κουμάρ Σάμπαβ Σριβαστάβα είναι Ινδός δημοσιογράφος που ασχολείται με την αλληλεπίδραση της δημόσιας πολιτικής με τις επιχειρήσεις και την κοινωνική δικαιοσύνη.
Ο Άνκουρ Παλιγουάλ είναι δημοσιογράφος που γράφει για την αλληλεπίδραση της επιστήμης με την ανθρώπινη κατάσταση.

3.6

Ανασύνθεση Δημόσιων Δεδομένων στην Κούβα: Συνεργασίες όταν οι Πληροφορίες Απουσιάζουν, Είναι Παρωχημένες ή Ανεπαρκείς

Saimi Reyes Carmona, Yudivián Almeida Cruz, Ernesto Guerra

Περίληψη: Πώς μια μικρή ομάδα δημοσιογραφίας δεδομένων στην Κούβα μάχεται κατά της έλλειψης δεδομένων.

Λέξεις-κλειδιά: Κούβα, έλλειψη δεδομένων, εκμάθηση δεδομένων, τεχνητή νοημοσύνη, δημοσιογραφία δεδομένων, συνεργασίες ερευνητών-δημοσιογράφων

Το Postdata.club είναι μια μικρή ομάδα. Ξεκινήσαμε τέσσερις δημοσιογράφοι και ένας ειδικός στα μαθηματικά και στην πληροφορική που, το 2014, αποφασίσαμε να ασχοληθούμε από κοινού με τη δημοσιογραφία δεδομένων στην Κούβα. Ως εκείνη τη στιγμή, δεν υπήρχε κανένα μέσο ενημέρωσης στην Κούβα το οποίο να είναι αφιερωμένο αποκλειστικά στη δημοσιογραφία δεδομένων και μας ενδιέφερε να κατανοήσουμε τι συνεπάγεται αυτή η πρακτική.

Σήμερα, είμαστε δύο δημοσιογράφοι και ένας επιστήμονας δεδομένων που εργαζόμαστε στον ελεύθερο χρόνο μας πάνω σε ιστορίες δεδομένων για το Postdata.club. Η δημοσιογραφία δεδομένων δεν περιλαμβάνεται στις καθημερινές μας εργασίες. Η Σάιμι Ρέγιες είναι αρχισυντάκτρια ενός πολιτιστικού ιστοτόπου, ο Γιουντίβιαν Αλμέιντα είναι καθηγητής στη Σχολή Μαθηματικών και Πληροφορικής στο Πανεπιστήμιο της Αβάνας και ο Ερνέστο Γκιέρα είναι δημοσιογράφος σε ένα περιοδικό εκλαϊκευμένης επιστήμης και τεχνολογίας. Σκοπός μας δεν είναι να είμαστε απλώς ένας οργανισμός μέσων ενημέρωσης, αλλά να αποτελέσουμε έναν πειραματικό χώρο όπου θα είναι δυνατόν να εξερευνήσουμε και να μάθουμε πράγματα, με και μέσω των δεδομένων, για το έθνος στο οποίο ζούμε.

Το Postdata.club ζει στο GitHub[43]. Αυτό συμβαίνει, επειδή θέλουμε να μοιραστούμε όχι μόνο ιστορίες αλλά και τον τρόπο με τον οποίο διεξάγουμε έρευνες και μελέτες. Ανάλογα με τις απαιτήσεις της ιστορίας που θέλουμε να πούμε, αποφασίζουμε για τα μέσα που θα χρησιμοποιήσουμε, είτε πρόκειται για γραφικά, είτε για εικόνες, βίντεο ή ήχο. Επικεντρωνόμαστε στη δημοσιογραφία με κοινωνικό αντίκτυπο, άλλοτε σε μεγάλη και άλλοτε σε μικρότερη φόρμα. Ενδιαφερόμαστε για όλα τα θέματα που μπορούμε να προσεγγίσουμε με δεδομένα, αλλά, πάνω από όλα, ενδιαφερόμαστε για εκείνα που σχετίζονται με την Κούβα ή τους ανθρώπους της.

Ο τρόπος με τον οποίο προσεγγίζουμε τις έρευνές μας εξαρτάται από τα δεδομένα στα οποία

43. ΣτΕ: Πλατφόρμα δημοσίευσης, διαμοιρασμού, ανταλλαγής και συνεργατικής συγγραφής κώδικα. Το περιεχόμενο δομείται και δημοσιεύεται με τη μορφή (ανοικτών) αποθετηρίων, δηλαδή συλλογών προγραμμάτων, συνόλων δεδομένων, έργων κλπ. Τα αποθετήρια στο GitHub μπορούν να περιέχουν οποιουδήποτε είδους κώδικα, περιλαμβανομένων αρχείων .html και CSS. Έτσι, μέσω των GitHub Pages, είναι δυνατή και η δημιουργία/φιλοξενία ιστοτόπου.

έχουμε πρόσβαση. Ορισμένες φορές έχουμε πρόσβαση σε δημόσιες και ανοιχτές βάσεις δεδομένων. Με αυτές, προχωρούμε σε αναλύσεις δεδομένων, για να διαπιστώσουμε εάν υπάρχει κάποια ιστορία που μπορεί να ειπωθεί. Μερικές φορές έχουμε ερωτήσεις και ψάχνουμε στα δεδομένα, για να βρούμε απαντήσεις που θα μπορούσαν να αποτελέσουν μια ιστορία. Σε άλλες περιπτώσεις, διερευνούμε τα δεδομένα και, στην πορεία, βρίσκουμε ενδιαφέροντα στοιχεία ή ερωτήματα που μπορεί να απαντηθούν από άλλα δεδομένα τα οποία δεν έχουμε ακόμη στην κατοχή μας.

Άλλες φορές –και μάλιστα όχι σπάνια–, για να υποστηρίξουμε τις αναλύσεις και τις έρευνές μας, πρέπει να δημιουργήσουμε μόνοι μας βάσεις δεδομένων οι οποίες εξαρτώνται από πληροφορίες που είναι μεν δημόσιες, αλλά δεν είναι σωστά δομημένες. Για παράδειγμα, για να καλύψουμε τις εκλογές στην Κούβα, χρειάστηκε να δημιουργήσουμε βάσεις δεδομένων, συνδυάζοντας πληροφορίες από διαφορετικές πηγές. Ξεκινήσαμε με δεδομένα που έχουν δημοσιευθεί στον ιστότοπο του Κοινοβουλίου της Κούβας. Τα δεδομένα αυτά, ωστόσο, δεν ήταν πλήρη. Επομένως, τα συμπληρώσαμε με δημοσιεύματα του Τύπου και πληροφορίες από ιστοσελίδες του Κομμουνιστικού Κόμματος της Κούβας.

Για να καλύψουμε το πρόσφατα διορισμένο Υπουργικό Συμβούλιο, ήταν, επίσης, απαραίτητο να δημιουργήσουμε μια βάση δεδομένων. Στην περίπτωση αυτή, οι πληροφορίες που παρείχε η Εθνοσυνέλευση δεν ήταν πλήρεις και χρησιμοποιήσαμε δημοσιεύματα του Τύπου, την Εφημερίδα της Κυβερνήσεως και άλλους ιστοτόπους, για να αποκτήσουμε μια πιο ολοκληρωμένη εικόνα. Και στις δύο περιπτώσεις, δημιουργήσαμε βάσεις δεδομένων σε .json φορμάτ, οι οποίες αναλύθηκαν και χρησιμοποιήθηκαν για τα περισσότερα άρθρα που γράψαμε σχετικά με τις εκλογές και την εκτελεστική και νομοθετική εξουσία στην Κούβα.

Στις περισσότερες περιπτώσεις, κοινοποιούμε τέτοιες βάσεις δεδομένων στον ιστότοπό μας με μια επεξήγηση για τις μεθόδους μας. Ωστόσο, η δουλειά μας ορισμένες φορές περιπλέκεται από την έλλειψη δεδομένων που θα έπρεπε να είναι δημόσια και προσβάσιμα. Πολλές από τις πληροφορίες που χρησιμοποιούμε παρέχονται από κυβερνητικούς φορείς, αλλά στη χώρα μας πολλοί φορείς δεν έχουν διαδικτυακή παρουσία ή δεν δημοσιοποιούν όλες τις πληροφορίες που θα έπρεπε. Σε ορισμένες περιπτώσεις, απευθυνθήκαμε απευθείας σε αυτούς, για να ζητήσουμε πρόσβαση σε ορισμένες πληροφορίες –μια διαδικασία που συχνά είναι βραδυκίνητη αλλά σημαντική.

Για εμάς, ένα από τα μεγαλύτερα προβλήματα με τα δεδομένα που μπορούμε να αποκτήσουμε στην Κούβα είναι η παλαιότητά τους. Όταν τελικά αποκτήσουμε πρόσβαση στις πληροφορίες που αναζητούμε, αυτές είναι συχνά ελλιπείς ή πολύ παρωχημένες. Έτσι, τα δεδομένα μπορεί να είναι διαθέσιμα για ανάγνωση και λήψη σε έναν δικτυακό τόπο, αλλά να έχουν σταματήσει να ενημερώνονται πέντε χρόνια πριν. Σε αυτές τις περιπτώσεις, εντοπίζουμε άλλους αξιόπιστους δικτυακούς τόπους που παρέχουν επικαιροποιημένες πληροφορίες ή καταφεύγουμε σε έντυπα ή σκαναρισμένα έγγραφα, ή σε ανθρώπινες πηγές.

Οι συνεργασίες με φοιτητές και ερευνητές είναι ένας από τους τρόπους με τους οποίους προσεγγίζουμε καταστάσεις όπου οι πληροφορίες λείπουν, είναι παρωχημένες ή λιγοστές. Από το 2017, διδάσκουμε ένα μάθημα δημοσιογραφίας δεδομένων σε φοιτητές δημοσιογραφίας στη Σχολή Επικοινωνίας του Πανεπιστημίου της Αβάνας. Με τις επαφές μας με αυτούς τους μελλοντικούς δημοσιογράφους και επαγγελματίες της επικοινωνίας μάθαμε νέους τρόπους εργασίας και ανακαλύψαμε νέους τρόπους πρόσβασης στις πληροφορίες.

Ένα από τα πράγματα που κάνουμε σε αυτά τα μαθήματα είναι να συμπεριλάβουμε τους φοιτητές στη δημιουργία μιας βάσης δεδομένων. Για παράδειγμα, στην Κούβα δεν υπήρχε μία και μοναδική πηγή με τα ονόματα των ατόμων που έχουν λάβει εθνικά βραβεία, με βάση το έργο που έχουν κάνει στη ζωή τους, σε διαφορετικούς τομείς. Μαζί με φοιτητές και καθηγητές, συγκεντρώσαμε και δομήσαμε μια βάση δεδομένων με όσους έλαβαν περισσότερα από 27 βραβεία

από τότε που αυτά άρχισαν να απονέμονται μέχρι σήμερα. Οι πληροφορίες αυτές μας επέτρεψαν να αποκαλύψουμε ότι υπήρχε χάσμα μεταξύ των δύο φύλων στην απονομή των βραβείων. Οι γυναίκες είχαν λάβει τα εν λόγω βραβεία μόνο στο 25% των περιπτώσεων. Με αυτήν την ανακάλυψη μπορέσαμε, από κοινού, να γράψουμε ένα θέμα που ενθάρρυνε τον προβληματισμό για θέματα φύλου σε σχέση με την εθνική αναγνώριση των διαφόρων εργασιών.

Το 2017 είχαμε μια άλλη αποκαλυπτική εμπειρία. Αυτή η εμπειρία μάς βοήθησε να κατανοήσουμε ότι, σε πολλές περιπτώσεις, δεν πρέπει να αρκεστούμε σε υπάρχουσες δημοσιευμένες βάσεις δεδομένων και δεν πρέπει να κάνουμε πολλές πρώιμες υποθέσεις σχετικά με το τι είναι και τι δεν είναι εφικτό. Στο πλαίσιο της τελικής τους εργασίας, ζητήσαμε από τους φοιτητές να σχηματίσουν μικρές ομάδες για να διεξαγάγουν μια έρευνα. Οι ομάδες αυτές αποτελούνταν, σε κάθε περίπτωση, από ένα από τα τέσσερα μέλη της ομάδας του Postdata.club, δύο φοιτητές δημοσιογραφίας και έναν φοιτητή πληροφορικής. Μία από τις ομάδες πρότεινε να ασχοληθεί με νέες πρωτοβουλίες αυτοαπασχόλησης στην Κούβα. Εδώ, αυτοί οι άνθρωποι ονομάζονται *cuentapropistas*. Αυτό που πριν από λίγα χρόνια αποτελούσε μια πολύ περιορισμένη πρακτική, σήμερα αναπτύσσεται ραγδαία λόγω της σταδιακής αποδοχής αυτής της μορφής απασχόλησης από την κοινωνία.

Θέλαμε να διερευνήσουμε το φαινόμενο της αυτοαπασχόλησης στην Κούβα. Παρόλο που το θέμα έχει απασχολήσει συχνά, δεν υπήρχε σχεδόν τίποτα σχετικά με τις ιδιαιτερότητες της αυτοαπασχόλησης ανά επαρχία, τον αριθμό των αδειών που χορηγούνται ανά τομέα δραστηριότητας ή τις τάσεις με την πάροδο του χρόνου. Μαζί με τους φοιτητές, συζητήσαμε ποια ερωτήματα έπρεπε να εξετάσουμε και καταλήξαμε στο συμπέρασμα ότι μας έλειπαν οι καλές πηγές δεδομένων. Εκεί όπου αυτές οι πληροφορίες θα έπρεπε να έχουν αναρτηθεί δημόσια, δεν υπήρχε κανένα ίχνος τους. Πέρα από ορισμένες συνεντεύξεις και μεμονωμένα νούμερα, δεν υπήρχαν πολλές πληροφορίες για το θέμα αυτό στον εθνικής εμβέλειας Τύπο.

Πιστεύαμε ότι θα ήταν δύσκολο να συγκεντρωθούν τα δεδομένα. Ωστόσο, φοιτητές δημοσιογραφίας από το πρόγραμμά μας προσέγγισαν το Υπουργείο Εργασίας και Κοινωνικών Ασφαλίσεων και ζήτησαν πληροφορίες σχετικά με την αυτοαπασχόληση στην Κούβα. Σε μερικές μέρες οι φοιτητές είχαν μια βάση δεδομένων στα χέρια τους. Ξαφνικά, είχαμε πληροφορίες που θα ενδιέφεραν πολλούς Κουβανούς και θα μπορούσαμε να τις μοιραστούμε μαζί με τις ιστορίες μας. Είχαμε λανθασμένα υποθέσει ότι τα δεδομένα δεν προορίζονταν για το κοινό, ενώ το υπουργείο απλώς δεν διέθετε ενημερωμένη διαδικτυακή πύλη.

Συμπτωματικά, οι πληροφορίες έφτασαν στα χέρια μας σε μια ιδιαίτερα βολική στιγμή: εκείνη την περίοδο, το Υπουργείο Εργασίας και Κοινωνικών Ασφαλίσεων αποφάσισε να περιορίσει την έκδοση αδειών για 28 από τις δραστηριότητες που είχαν εγκριθεί για μη κρατική απασχόληση. Έτσι, βρεθήκαμε σε θέση να χρησιμοποιήσουμε γρήγορα τα δεδομένα που είχαμε αποκτήσει, για να αναλύσουμε πώς αυτά τα νέα μέτρα θα επηρέαζαν την οικονομία της χώρας και τη ζωή των αυτοαπασχολούμενων εργαζομένων.

Οι περισσότεροι αναγνώστες μας εξεπλάγησαν από το γεγονός ότι μπορέσαμε να αποκτήσουμε τα δεδομένα και ότι ήταν σχετικά εύκολο να τα αποκτήσουμε. Στην πραγματικότητα, η πρόσβαση σε αυτά τα δεδομένα κατέστη δυνατή, επειδή οι φοιτητές μας τα είχαν ζητήσει από το υπουργείο και μέχρι σήμερα το Postdata.club είναι ο διαδικτυακός τόπος που καθιστά αυτές τις πληροφορίες δημόσια προσβάσιμες.

Το να κάνεις δημοσιογραφία δεδομένων στην Κούβα εξακολουθεί να αποτελεί πρόκληση. Μεταξύ άλλων, η δυναμική της δημιουργίας και της πρόσβασης σε δεδομένα και οι πολιτικές και θεσμικές κουλτούρες διαφέρουν από άλλες χώρες όπου τα δεδομένα είναι πιο εύκολα διαθέσιμα. Ως εκ τούτου, πρέπει πάντα να είμαστε δημιουργικοί στην αναζήτηση νέων τρόπων πρόσβασης σε πληροφορίες, για να αφηγηθούμε ιστορίες που έχουν σημασία. Είναι εφικτό μόνο αν συνεχίσουμε να προσπαθούμε, και στο Postdata.club θα προσπαθούμε πάντα να αποτε-

λούμε παράδειγμα για το πώς η δημοσιογραφία δεδομένων είναι εφικτή ακόμη και σε περιοχές όπου τα δεδομένα είναι πιο δύσκολο να βρεθούν.

Οι συγγραφείς

Η Σάιμι Ρέγιες είναι δημοσιογράφος, ιδρύτρια και αρχισυντάκτρια του Postdata.club.
Ο Γιουντίβιαν Αλμέιντα είναι δημοσιογράφος, ιδρυτής και υπεύθυνος επεξεργασίας δεδομένων στο Postdata.club και Καθηγητής Πληροφορικής και Δημοσιογραφίας Δεδομένων στο Πανεπιστήμιο της Αβάνας.
Ο Ερνέστο Γκιέρα είναι δημοσιογράφος και ένας από τους ιδρυτές των πρώτων ιστοτόπων δημοσιογραφίας δεδομένων στην Κούβα.

3.7
Δημιουργώντας Δεδομένα με τους Αναγνώστες της *La Nación*

Flor Coelho

Περίληψη: Διοργάνωση μαραθωνίων με τη συμμετοχή πολιτών και χρήση της πλατφόρμας ανοιχτού κώδικα Vozdata για τη συνεργασία με αναγνώστες, πανεπιστήμια και ΜΚΟ γύρω από μεγάλες έρευνες βασισμένες σε δεδομένα.

Λέξεις-κλειδιά: μαραθώνιοι πολιτών, crowdsourcing, ερευνητική δημοσιογραφία, ανοιχτός κώδικας, συνεργασίες ερευνητών-δημοσιογράφων, ανοιχτά δεδομένα

Στη *La Nación* έχουμε πραγματοποιήσει μεγάλες έρευνες βασισμένες σε δεδομένα, σε συνεργασία με τους αναγνώστες μας. Σε αυτό το κεφάλαιο ρίχνουμε μια ματιά στα «παρασκήνια», στον τρόπο με τον οποίο έχουμε οργανώσει τη συμμετοχή των αναγνωστών γύρω από ορισμένα από αυτά τα έργα –μεταξύ άλλων μέσω του καθορισμού στόχων, της υποστήριξης κοινοτήτων έρευνας και της καλλιέργειας μακροχρόνιων συνεργασιών με τους αναγνώστες μας και άλλους εξωτερικούς οργανισμούς και φορείς με τους οποίους συμπράττουμε.

Σε τέτοια έργα, στόχος μας είναι συχνά να επιτύχουμε το «ακατόρθωτο», χρησιμοποιώντας την τεχνολογία για να διευκολύνουμε συνεργασίες μεγάλης κλίμακας, δίνοντας τη δυνατότητα στους χρήστες να ασχοληθούν με τη δημοσιογραφική έρευνα και τη διαδικασία δημοσιοποίησης επίσημων δεδομένων.

Για παράδειγμα, χρειαστήκαμε περίπου πέντε χρόνια για την αποδελτίωση 10.000 αρχείων .pdf με τα έξοδα της Γερουσίας, δύο χρόνια για την ακρόαση 40.000 τηλεφωνικών κλήσεων που είχαν υποκλαπεί και μερικούς μήνες για την ψηφιοποίηση περισσότερων από 20.000 χειρόγραφων εκλογικών εντύπων[44].

Για τέτοιου είδους πρωτοβουλίες πληθοπορισμού («crowdsourcing»), βασιστήκαμε στη διαδικτυακή πλατφόρμα συνεργασίας Vozdata. Η πλατφόρμα δημιουργήθηκε με έμπνευση από τις εκστρατείες crowdsourcing «Free the Files» της ProPublica για τις εκλογικές δαπάνες αλλά και του Guardian για τα έξοδα των βουλευτών· αναπτύχθηκε με την υποστήριξη του προγράμματος υποτροφιών Knight-Mozilla του OpenNews και της CIVICUS, μιας παγκόσμιας συμμαχίας οργανώσεων και ακτιβιστών της κοινωνίας των πολιτών. Το λογισμικό πίσω από το Vozdata ήταν ανοιχτού κώδικα όπως, για παράδειγμα, το Crowdata[45].

44. http://blogs.lanacion.com.ar/projects/data/argentina%C2%B4s-senate-expens-es-2004-2013/, http://blogs.lanacion.com.ar/projects/data/prosecutor-nisman-phone-interceptions-mapped-in-playlists/, http://blogs.lanacion.com.ar/projects/data/vozdata-2015-unveiling-argentina%c2%b4s-elections-system-failures-with-impact/

45. http://blogs.lanacion.com.ar/projects/data/vozdata-ii-civic-participation-for-investigative-reporting-and-educational-platform/, http://www.theguardian.com/news/datablog/2009/jun/18/mps-expenses-houseofcommons, https://www.propublica.org/series/free-the-files , https://github.com/crowdata/crowdata

Οργανώνοντας τη συμμετοχή

Για τα έργα αυτά οι συνεργάτες μας ήταν κυρίως φοιτητές δημοσιογραφίας, εθελοντές πολίτες, ΜΚΟ που ασχολούνται με τη διαφάνεια και συνταξιούχοι. Ο καθένας τους έχει διαφορετικά κίνητρα συμμετοχής ανάλογα με το έργο. Αυτά μπορεί να περιλαμβάνουν τη συμβολή σε έργα δημοσίου συμφέροντος, τη συνεργασία με την ομάδα δεδομένων μας και τη γνωριμία με άλλους ανθρώπους στις συναντήσεις μας.

Το Vozdata διαθέτει ομάδες και λειτουργίες κατάταξης σε πραγματικό χρόνο. Εξετάσαμε πώς αυτά μπορούν να βελτιώσουν τη συμμετοχή μέσω της «παιχνιδοποίησης» («gamification»)[46]. Είχαμε εξαιρετικά αποτελέσματα στην προώθηση της συμμετοχής των πολιτών με αυτόν τον τρόπο κατά τη διάρκεια των εθνικών εορτών της Αργεντινής. Η συμμετοχή στη δημιουργία συνεργατικών βάσεων δεδομένων γίνεται ως επί το πλείστον εξ αποστάσεως (διαδικτυακά). Ωστόσο, ενθαρρύναμε, επίσης, τους χρήστες να συμμετάσχουν σε «offline» δράσεις πολιτών που πραγματοποιήθηκαν στη *La Nación* ή κατά τη διάρκεια «hackathon»[47] σε διάφορες εκδηλώσεις. Σε κάποιες περιπτώσεις, έχουμε δημιουργήσει ανοιχτές (δηλαδή ελεύθερα επαναχρησιμοποιήσιμες) βάσεις δεδομένων με φοιτητές δημοσιογραφίας σε συνεργαζόμενα πανεπιστήμια.

Ενώ τα hackathon είναι εκδηλώσεις που συνήθως διαρκούν μία ή δύο ημέρες, οι διαδικτυακοί μας μαραθώνιοι μπορούν να κρατήσουν για μήνες. Η γραμμή προόδου δείχνει πόσα έγγραφα έχουν ολοκληρωθεί και το ποσοστό που απομένει.

Θέτοντας υψηλούς στόχους

Ο κύριος ρόλος των συμμετεχόντων στα έργα *Senate Expenses* και *Electoral Telegrams* ήταν η συλλογή συγκεκριμένων δομημένων δεδομένων από τα παρεχόμενα έγγραφα. Αυτό περιελάμβανε περισσότερους από χίλιους μοναδικούς χρήστες. Εκτός από την εξαγωγή αυτών των λεπτομερειών, οι αναγνώστες είχαν επίσης τη δυνατότητα να επισημάνουν δεδομένα ως ύποπτα ή μη αποδεκτά και να αφήσουν ένα σχόλιο με πρόσθετες πληροφορίες. Η τελευταία λειτουργία χρησιμοποιήθηκε σπάνια. Όταν υπάρχει προθεσμία για την ολοκλήρωση ενός έργου crowdsourcing, ο στόχος μπορεί να μην επιτευχθεί. Αυτό μας συνέβη στο έργο *Electoral Telegrams*. Η ημέρα των εκλογών πλησίαζε και έπρεπε να δημοσιεύσουμε ορισμένα συμπεράσματα. Ενώ κάποιες επαρχίες έφτασαν στο 100%, πολλές είχαν ολοκληρώσει μόνο το 10% έως 15% των αρχείων, γεγονός που επισημάναμε κατά τη δημοσίευση.

Υποστήριξη ερευνητικών κοινοτήτων

Για την έρευνα 40.000 αρχείων του Εισαγγελέα Νίσμαν, δουλέψαμε με ένα αξιόπιστο δίκτυο 100 συνεργατών. Πολλά αρχεία ήχου που αφορούσαν ιδιωτικές συνομιλίες (π.χ. οικογενειακούς διαλόγους) βρίσκονταν στην κατοχή του Ιρανού πράκτορα του οποίου το τηλέφωνο παγιδεύτηκε από ομοσπονδιακό δικαστήριο. Μια ομάδα έξι εθελοντών μπήκε πολύ βαθιά στην έρευνα. Δημιουργήσαμε ένα γκρουπ στο WhatsApp όπου ο καθένας μπορούσε να προτείνει στοιχεία και ενδιαφέρουσες πληροφορίες.

Ένας από τους εθελοντές μας έλυσε ένα μυστήριο που μας απασχολούσε επί δύο μήνες. Είχαμε επισημάνει αρκετές συνομιλίες όπου δύο άτομα μιλούσαν κωδικοποιημένα, χρησιμοποιώντας

46. ΣτΜ: Η παιχνιδοποίηση συνεπάγεται τη στρατηγική υλοποίηση μιας διαδικασίας που δεν είναι παιχνίδι, προσδίδοντάς της παιγνιώδη χαρακτήρα. Συνήθως χρησιμοποιεί μηχανισμούς και σχεδιαστικά στοιχεία παιχνιδιού για να μετρήσει, να επηρεάσει και να ανταμείψει συμπεριφορές χρηστών-συμμετεχόντων σε μια διαδικασία.

47. ΣτΜ: Μαραθώνιοι ανάπτυξης εφαρμογών, με στόχο τη δημιουργία λειτουργικού λογισμικού εντός του πλαισίου μιας διοργάνωσης και μέχρι το τέλος αυτής.

ψευδώνυμα και αριθμούς (π.χ. «Mr. Dragon, 2000»). Πολλοί εθελοντές είχαν ακούσει και απομαγνητοφωνήσει τέτοιες ηχογραφήσεις. Σκεφτήκαμε να δημιουργήσουμε μια ξεχωριστή βάση δεδομένων για να αναλύσουμε τον κώδικα σε αυτές. Μια μέρα, ένας εθελοντής ανακάλυψε ότι οι συζητήσεις αφορούσαν στοιχήματα στον ιππόδρομο! Μια γρήγορη αναζήτηση στο Google επιβεβαίωσε ότι πολλά από τα ονόματα ανήκαν σε άλογα αγώνων.

Εικόνα 3.7.1: Έναρξη του μαραθωνίου με τη συμμετοχή πολιτών του Vozdata για το έργο *Senate Expenses* στη *La Nación*. Πηγή: *La Nación*.

Υπάρχουν πάντα και οι έμπειροι χρήστες. Ωστόσο, ανάλογα με την κλίμακα του έργου, πολλοί εθελοντές-συνεργάτες, με λίγα έγγραφα ο καθένας, συνήθως υπερβαίνουν τη «συνεισφορά του υπερχρήστη».

Καλλιεργώντας συνεργασίες

Η συμβουλή μας για τους δημοσιογράφους και τους οργανισμούς που θέλουν να εμπλέξουν τους αναγνώστες τους στις έρευνες δεδομένων είναι να ορίσουν έναν ειδικό διαχειριστή της κοινότητας για την οργάνωση και την επικοινωνία μέσω συνεργατικών υπολογιστικών φύλλων (π.χ. Google Sheets), λιστών αλληλογραφίας και μέσων κοινωνικής δικτύωσης.

Οι μεγάλες συλλογές εγγράφων μπορούν να αποτελέσουν ένα καλό σημείο εκκίνησης: η καμπύλη μάθησης είναι μικρή και οι συμμετέχοντες αισθάνονται ότι γίνονται μέρος μιας σημαντικής διαδικασίας. Είναι, επίσης, χρήσιμο να υποστηρίζονται οι συνεργάτες με βιντεοσκοπημένα σεμινάρια ή σχετικές επεξηγηματικές παρουσιάσεις στον χώρο του οργανισμού σας ή σε ειδικές εκδηλώσεις.

Όταν κερδίσαμε ένα βραβείο που σχετιζόταν με αυτά τα συνεργατικά έργα, διοργανώσαμε ένα πρόγευμα, για να μοιραστούμε το βραβείο με τους εθελοντές. Αυτές είναι μακροχρόνιες

σχέσεις με τους αναγνώστες μας. Επομένως, φροντίσαμε να αφιερώσουμε χρόνο και ενέργεια, προκειμένου να συναντιόμαστε σε εκδηλώσεις, να επισκεπτόμαστε πανεπιστήμια, να δίνουμε συνεντεύξεις για φοιτητικές εργασίες και ούτω καθεξής.

Όσον αφορά στις συνεργασίες με πανεπιστήμια, οι καθηγητές λειτουργούν συνήθως ως κόμβοι. Κάθε χρόνο έχουν μια νέα τάξη φοιτητών που συνήθως είναι πρόθυμοι να δουλέψουν μαζί μας σε συνεργατικά έργα (προγραμματίστε το εκ των προτέρων!).

Οι ΜΚΟ που ασχολούνται με τη διαφάνεια μπορούν, επίσης, να καταδείξουν τα οφέλη αυτών των δραστηριοτήτων. Στην πλατφόρμα μας, η κάθε εργασία μπορεί να καταχωρηθεί, έτσι ώστε τα έργα να παρουσιάζονται εύκολα και να επιτυγχάνεται η αναγνώριση των εθελοντών στα μέσα ενημέρωσης.

Κατά τη δημοσίευση των αποτελεσμάτων και των ιστοριών, συνιστούμε να γίνεται αναφορά στη διαδικασία συνεργασίας και στους συμμετέχοντες οργανισμούς σε κάθε πλατφόρμα (έντυπα, διαδικτυακά μέσα αλλά και μέσα κοινωνικής δικτύωσης) και στις ταχυδρομικές αποστολές. Η ανάδειξη του συλλογικού χαρακτήρα τέτοιων έργων μπορεί να στείλει ένα ισχυρότερο μήνυμα σε αυτούς που θέλουμε να καταστήσουμε υπόλογους.

Συμπέρασμα

Προκειμένου να δημιουργήσετε δεδομένα μαζί με τους αναγνώστες, είναι ζωτικής σημασίας να διαθέσετε χρόνο και πόρους, ώστε να αλληλεπιδράσετε με την κοινότητά σας, να ασχοληθείτε με τα αιτήματα, να αναλύσετε τα αποτελέσματα, να απολαύσετε τις επαφές και να συμμετέχετε σε εκδηλώσεις.

Οι εθελοντές ταξινομούν τα έγγραφα, επειδή πιστεύουν ότι ένα έργο έχει σημασία. Για τις κυβερνήσεις και εκείνους που αναφέρονται σε αυτό είναι ένα σημάδι ότι το έργο δεν αφορά μόνο τον Τύπο, αλλά επηρεάζει, επίσης, την κοινωνία των πολιτών. Μέσω τέτοιων έργων, οι συμμετέχοντες μπορούν να γίνουν παθιασμένοι υποστηρικτές και διαδικτυακοί διανομείς του περιεχομένου.

Η συγγραφέας

Η Φλορ Κοέλο είναι υπεύθυνη έρευνας για τα νέα μέσα ενημέρωσης και υπεύθυνη επιμόρφωσης στη *La Nación* στην Αργεντινή.

3.8

Διεξαγωγή Έρευνας Κοινού στο πλαίσιο Δημοσιογραφικής Έρευνας

Crina-Gabriela Boroş

Περίληψη: Διαβάστε αυτό το κείμενο πριν από τη διεξαγωγή έρευνας κοινού για τη δημοσιογραφία της λογοδοσίας: τι πρέπει, τι δεν πρέπει να κάνετε και πώς να χειριστείτε τις προβληματικές περιστάσεις.

Λέξεις-κλειδιά: στατιστική, δημοσιογραφία δεδομένων, έρευνες κοινού, λογοδοσία

Είναι ένα ζήτημα περιπτωσιολογικό ή συστηματικό; Προσπαθείτε να το εξακριβώσετε, όταν συνειδητοποιείτε ότι δεν υπάρχουν πινακοποιημένα δεδομένα –μια φανταχτερή φράση που αναφέρεται απλώς σε πληροφορίες που παρατίθενται σε γραμμές και στήλες. Τι πρέπει να κάνετε;

Και τι είναι τα δεδομένα, εν πάση περιπτώσει; Υπάρχουν πολλοί «ξερολίστικοι» ορισμοί που κυκλοφορούν, μερικοί από τους οποίους είναι τρομακτικοί[48]. Ας αντικαταστήσουμε, λοιπόν, τον όρο «δεδομένα» με την απλή έννοια της «πληροφορίας». Και καθώς συλλέγετε αυτά τα δεδομένα, σε οποιοδήποτε σχήμα ή μορφή, πρέπει να μπορείτε να βρείτε μοτίβα και ακραίες τιμές. Αυτό σημαίνει ότι πρέπει να έχετε μια σημαντική ποσότητα συστηματικά συγκεντρωμένου πρωτογενούς υλικού που τεκμηριώνει ένα θέμα, σύμφωνα με μια συγκεκριμένη μέθοδο (σκεφτείτε τα ερωτηματολόγια). Είτε χρησιμοποιείτε ένα υπολογιστικό φύλλο, είτε περιβάλλον προγραμματισμού, είτε μια εφαρμογή, είτε στιλό και χαρτί, δεν έχει σημασία.

Μερικές φορές, οι σκέψεις, τα συναισθήματα ή οι προηγούμενες προσωπικές εμπειρίες που είναι παγιδευμένες στις καρδιές και στο μυαλό των ανθρώπων μπορούν να διατυπωθούν ως δεδομένα. Μια μέθοδος συλλογής αυτών των πολύτιμων πληροφοριών είναι ο σχεδιασμός μιας έρευνας κοινού που θα συγκεντρώνει και θα ταξινομεί αυτά τα συναισθήματα και τις εμπειρίες σε έναν πίνακα, ένα αρχείο ή μια βάση δεδομένων στην οποία δεν έχει πρόσβαση κανείς άλλος εκτός από εσάς.

Για παράδειγμα, το Ίδρυμα Thomson Reuters («Thomson Reuters Foundation», TRF) ανέλαβε ένα έργο σχετικά με το πώς αντιλαμβάνονται οι γυναίκες στις μεγαλύτερες πρωτεύουσες του κόσμου τις επιπτώσεις της σεξουαλικής βίας στα μέσα μαζικής μεταφοράς[49]. Ήταν μια προσπάθεια που βασίστηκε σε έρευνα κοινού, με στόχο την ευαισθητοποίηση για το θέμα αλλά και τη σύγκριση και την αντιπαραβολή (αυτά κάνουν οι στατιστικές).

Για να παρουσιάσουμε αυτό το θέμα, περάσαμε από κάμποσους κύκλους της κολάσεως, καθώς υπάρχουν αυστηρές συμβάσεις που οφείλουν να πληρούν οι μέθοδοι των κοινωνικών επιστημών, όπως είναι οι έρευνες κοινού, ακόμη και όταν χρησιμοποιούνται από τους δημοσιογράφους στην άσκηση της εργασίας τους.

48. Βλ. 130 ορισμούς των δεδομένων, των πληροφοριών και της γνώσης στο Zins, C. (2007) «Conceptual approaches for defining data, information, and knowledge», *Journal of the American Society for Information Science and Technology*, 58, σσ. 479-493.

49. http://news.trust.org/spotlight/most-dangerous-transport-systems-for-women/?tab=stories

Ακολουθούν μερικοί βασικοί κανόνες διεξαγωγής δημοσκοπήσεων, τους οποίους οι δημοσιογράφοι θα ήταν χρήσιμο να γνωρίζουν, αλλά συχνά δεν τους εκπαιδεύει κανείς σχετικά.

- *Δεν επιτρέπεται να επιλέγονται οι ερωτώμενοι.* Προκειμένου να θεωρηθεί «αντιπροσωπευτική», η ομάδα των ερωτώμενων θα πρέπει συμβατικά να περιλαμβάνει άτομα από όλες τις κοινωνικές κατηγορίες, τις ηλικιακές ομάδες, τα επίπεδα εκπαίδευσης και τις γεωγραφικές περιοχές για τις οποίες γίνεται το ρεπορτάζ. Σύμφωνα με καθιερωμένες μεθόδους, τα δείγματα του υπό μελέτη πληθυσμού πρέπει να είναι αντιπροσωπευτικά.
- *Η επιλογή των ερωτώμενων πρέπει να είναι τυχαιοποιημένη*, δηλαδή όλοι να έχουν την ίδια πιθανότητα κάποιος να τραβήξει το όνομά τους από ένα καπέλο. Εάν διεξάγετε μια δημοσκόπηση και μιλάτε σε όποιον είναι πιο κοντά σας χωρίς κανένα κριτήριο ή μέθοδο, θεωρείται ότι υπάρχει κίνδυνος παραγωγής παραπλανητικών δεδομένων, ιδίως εάν στοχεύετε σε γενικεύσιμους ισχυρισμούς.
- *Ο αριθμός των ατόμων που συμμετέχουν σε μια έρευνα κοινού πρέπει, επίσης, να είναι από ένα ορισμένο όριο και άνω, για να είναι αντιπροσωπευτικός.* Για να υπολογίσετε το μέγεθος του δείγματος, υπάρχουν χρήσιμες διαδικτυακές αριθμομηχανές, όπως αυτές που παρέχονται από τις εταιρείες Raosoft, Survey Monkey ή Survey Systems[50]. Κατά κανόνα: διατηρήστε το επίπεδο εμπιστοσύνης στο 95% και το περιθώριο σφάλματος έως 5%.
- *Οι επιλογές απάντησης πρέπει να επιτρέπουν στους ερωτώμενους να μη γνωρίζουν ή να μην είναι σίγουροι.* Όταν οι δημοσιογράφοι ακολουθούν αυτούς τους βασικούς κανόνες, τα ευρήματά τους είναι σχεδόν απρόσβλητα. Κατά τη διάρκεια της διεξαγωγής της έρευνας για την ασφάλεια των δημόσιων συγκοινωνιών από το TRF, η μεθοδολογία της δημοσκόπησής μας ακολουθούσε τους συντηρητικούς κανόνες των κοινωνικών επιστημών. Το θέμα μας αφορούσε μια τόσο κοινή ανθρώπινη εμπειρία, που λέει πολλά για τη λειτουργία των σύγχρονων κοινωνιών το γεγονός ότι ένας οργανισμός του ΟΗΕ προσφέρθηκε να συμμετάσχει στην προσπάθειά μας. Ήταν μια ιδιαίτερη τιμή, την οποία, ως δημοσιογράφοι, έπρεπε, όμως, να αρνηθούμε.

Εάν σας αρέσει αυτό που σας λέμε, ήρθε η ώρα να παρακολουθήσετε ένα μάθημα στατιστικής.

Μερικές φορές η αυστηρή δημοσκόπηση δεν είναι ρεαλιστική. Αυτό δεν σημαίνει πάντα ότι δεν πρέπει να κάνετε δημοσκοπήσεις.

Ενώ υπάρχουν καθιερωμένες μέθοδοι για τις έρευνες κοινού, αυτές δεν εξαντλούν τι είναι δυνατό, νόμιμο ή ενδιαφέρον. Μπορεί να υπάρχουν άλλοι τρόποι διεξαγωγής δημοσκοπήσεων, ανάλογα με τις ανησυχίες, τους περιορισμούς και τους πόρους σας.

Για παράδειγμα, όταν το *openDemocracy* θέλησε να πάρει συνέντευξη από δημοσιογράφους σε 47 κράτη-μέλη του Ευρωπαϊκού Συμβουλίου σχετικά με την εμπορική πίεση μέσα στις αίθουσες σύνταξης, υπήρχε μικρή πιθανότητα να προκύψει υλικό στατιστικής σημαντικότητας.

«Μα γιατί;» μπορεί να ρωτήσετε.

Όλοι οι ερωτώμενοι έγιναν καταγγέλλοντες. Οι καταγγέλλοντες χρειάζονται προστασία, συμπεριλαμβανομένης της μη αποκάλυψης σημαντικών πραγματικών δημογραφικών δεδομένων, όπως η ηλικία ή το φύλο. Για παράδειγμα, περιμέναμε υλικό από χώρες όπου η άσκηση της ελευθερίας του λόγου μπορεί να οδηγήσει σε σοβαρές προσωπικές συνέπειες. Αποφασίσαμε ότι η παροχή προσωπικών δεδομένων δεν θα πρέπει να είναι υποχρεωτική· ούτε, αν παρέχονταν τα δεδομένα αυτά, θα έπρεπε να βρεθούν στον διακομιστή μιας εταιρείας που θα ήταν συνιδιοκτήτης των πληροφοριών μας.

50. http://www.raosoft.com/samplesize.html, https://www.surveymonkey.com/mp/sample-size-calculator/, https://www.surveysystem.com/sscalc.htm

Η ΕΕ είχε εντελώς διαφορετικές και ελλιπείς μετρήσεις των δημοσιογράφων στην περιοχή, πράγμα που σημαίνει ότι η δημιουργία ενός αντιπροσωπευτικού δείγματος σε επίπεδο χώρας ήταν δύσκολη υπόθεση.

Δεν μπορούσαμε να παρατάξουμε όλα τα συνδικάτα και τις ενώσεις του Τύπου και να τυχαιοποιήσουμε τους ερωτώμενους, επειδή οι κατάλογοι μελών είναι απόρρητοι. Επίσης, δεν τους συμπεριλαμβάνουν όλους, αν και θα ήταν μια αποδεκτή βάση, εφόσον ήμαστε ειλικρινείς σχετικά με τους περιορισμούς μας. Επιπλέον, σε ορισμένες χώρες, τα έργα διαφάνειας οδηγούν σε καταπίεση και λάβαμε συμβουλές εμπειρογνωμόνων σχετικά με χώρες στις οποίες δεν θα μπορούσαμε να ζητήσουμε την υποστήριξη των συνδικάτων χωρίς να προσελκύσουμε είτε εποπτεία είτε τιμωρητικές συνέπειες.

Σε τέτοιες περιπτώσεις, ωστόσο, δεν χρειάζεται να καούν και τα χλωρά μαζί με τα ξερά. Δεν τα κάψαμε.

Αντίθετα, προσδιορίσαμε τι είχε σημασία για το ρεπορτάζ μας και πώς οι δημοκοπικές μέθοδοι θα μπορούσαν να προσαρμοστούν, για να προκύψουν ιστορίες.

Αποφασίσαμε ότι το κύριο σημείο εστίασής μας ήταν τα παραδείγματα πίεσης της αγοράς μέσα στις αίθουσες σύνταξης ανά χώρα –εάν υπήρχε ένα μοτίβο για το πώς ασκούνταν αυτές οι πιέσεις και αν τα μοτίβα ταίριαζαν σε όλη την ευρύτερη περιοχή. Μας ενδιέφερε, επίσης, το είδος των φορέων που κατηγορούνταν για δραστηριότητες ξεπλύματος χρήματος στον Τύπο.

Προχωρήσαμε και δημιουργήσαμε ένα ερωτηματολόγιο για την έρευνα κοινού, βασισμένο σε συνεντεύξεις με δημοσιογράφους, εκθέσεις για την ελευθερία των μέσων ενημέρωσης και ανατροφοδότηση από ομάδες εστίασης («focus group»). Συμπεριλάβαμε ενότητες για ανοιχτές απαντήσεις.

Προωθήσαμε το ερωτηματολόγιο μέσω όλων των ελεγχόμενων καναλιών δημοσιογραφικής οργάνωσης. Στην ουσία, δεν κάναμε τυχαία επιλογή, αλλά δεν είχαμε, επίσης, κανέναν έλεγχο για το ποιος από τον Τύπο έλαβε το ερωτηματολόγιο. Είχαμε, επίσης, συνεργάτες –συμπεριλαμβανομένων των Δημοσιογράφων Χωρίς Σύνορα, της Εθνικής Ένωσης Δημοσιογράφων και της Ευρωπαϊκής Ομοσπονδίας Δημοσιογράφων– οι οποίοι βοήθησαν στη διάδοση του ερωτηματολογίου.

Η ανατροφοδότηση που λάβαμε μέσω της έρευνας κοινού προστέθηκε σε μια μοναδική βάση δεδομένων, αποδόθηκε βαθμολογία στις απαντήσεις και καταμετρήθηκαν οι ερωτώμενοι ανά χώρα, έγιναν συγκρίσεις μεταξύ περιπτωσιολογικών στοιχείων (θέματα που αναφέρθηκαν σποραδικά) και συστηματικών ζητημάτων (προβλήματα που αναφέρθηκαν καθολικά).

Τα πεδία ανοιχτών ερωτήσεων αποδείχθηκαν ιδιαίτερα χρήσιμα: οι ερωτώμενοι τα χρησιμοποίησαν, για να μας δώσουν συμβουλές. Ερευνήσαμε τα σχόλιά τους, με γνώμονα τα μοτίβα οικονομικής λογοκρισίας και τους τύπους των φερόμενων ως υπαιτίων. Αυτό μας έδωσε χρήσιμες πληροφορίες σχετικά με τα επόμενα ρεπορτάζ μας για την ελευθερία του Τύπου[51].

Παρόλο που δημοσιεύσαμε μια επισκόπηση των ευρημάτων, δεν δημοσιεύσαμε ποτέ την ανάλυση των δεδομένων για τον απλούστατο λόγο ότι η επιλογή δεν μπορούσε να γίνει με τυχαίο τρόπο και τα μεγέθη του δείγματος σε επίπεδο χώρας δεν είχαν επιτευχθεί πάντα[52]. Ωστόσο, καταφέραμε να κατανοήσουμε αρκετά καλά πόσο ελεύθερος είναι ο Τύπος, σύμφωνα με το ίδιο του το προσωπικό, πώς δημιουργείται η διαφθορά στα μέσα ενημέρωσης, πώς εξελίσσεται και, δυστυχώς, πόσο ευάλωτοι είναι οι δημοσιογράφοι και η αλήθεια[53].

Υπάρχουν, λοιπόν, κανόνες για την παραβίαση των κανόνων;

51. https://www.opendemocracy.net/en/author/crina-boros-2/
52. https://www.opendemocracy.net/openmedia/mary-fitzgerald/welcome-to-openmedia
53. https://www.opendemocracy.net/en/welcome-to-openmedia/

Μόνο μερικοί. Πάντα να περιγράφετε με ακρίβεια τις προσπάθειές σας. Εάν ρωτήσατε τρεις κορυφαίους οικονομικούς συμβούλους της κυβέρνησης με μια ερώτηση «ναι»-«όχι», πείτε το. Εάν πήρατε συνέντευξη από δέκα θύματα εκφοβισμού, περιγράψτε πώς τα διαλέξατε και γιατί επιλέξατε τα συγκεκριμένα. Μη χαρακτηρίζετε ελαφρά τη καρδία τις συνεντεύξεις ως έρευνες κοινού.

Εάν διεξαγάγετε μια στατιστικά σημαντική μελέτη, να έχετε την ευγένεια να δημοσιεύσετε τη μεθοδολογία της[54]. Αυτό επιτρέπει τον απαραίτητο έλεγχο, ώστε το κοινό σας και οι ειδικοί να εμπιστεύονται τα δημοσιεύματά σας. Χωρίς μεθοδολογία, δεν υπάρχει εμπιστοσύνη.

Μην είστε ο επόμενος μεγάλος συγγραφέας «ψευδών ειδήσεων». Εάν ένας αρχισυντάκτης σάς πιέζει να κάνετε συσχετισμούς που βασίζονται σε συμπεράσματα και όχι σε ακριβή συλλογή δεδομένων, χρησιμοποιήστε γλώσσα που δεν υποδηλώνει αιτιότητα ή επιστημονική ισχύ. Η δουλειά μας είναι να δημοσιεύουμε την αλήθεια, όχι μόνο τα γεγονότα. Μη χρησιμοποιείτε τα γεγονότα, για να καλύψετε την έλλειψη βεβαιότητας για το ποια είναι η αλήθεια.

Πού βασίζεται η ιστορία σας; Σε ένα μοτίβο; Σε μια ακραία τιμή; Αποφασίστε ποια δεδομένα πρέπει να συλλέξετε βάσει αυτής της απάντησης. Ανακαλύψτε πού και πώς μπορούν να ληφθούν τα δεδομένα προτού αποφασίσετε για τις καταλληλότερες μεθόδους. Η μέθοδος δεν είναι ποτέ το θέμα, το θέμα είναι η ιστορία.

Εάν διεξαγάγετε μια έρευνα κοινού, ελέγξτε τα ευρήματά σας επιτόπου και προστατέψτε τα δημοσιεύματά σας από δυνητικά προβληματικούς ισχυρισμούς. Για παράδειγμα, ας πούμε ότι μια έρευνα κοινού δείχνει ότι το τμήμα της πόλης στο οποίο ζείτε έχει το υψηλότερο ποσοστό εγκληματικότητας. Εσείς, όμως, αισθάνεστε ασφαλείς. Ωστόσο, είχατε βιώσει, σχεδόν σε εβδομαδιαία βάση, τη βία στους δρόμους μιας άλλης γειτονιάς στην οποία ζούσατε επί έναν χρόνο. Επομένως, μπορεί να μην εμπιστεύεστε ακόμη τα δεδομένα. Για να ελέγξετε αν μπορείτε να εμπιστευτείτε τα δεδομένα σας, επισκεφθείτε τα μέρη που συγκρίνετε και αντιπαραβάλλετε· μιλήστε με ανθρώπους στους δρόμους, στα καταστήματα, στις τράπεζες, στις παμπ και στα σχολεία· εξετάστε ποια δεδομένα συλλέχθηκαν· είναι πιθανότερο οι κάτοικοι μιας περιοχής να κάνουν παράπονα σε σχέση με τους κατοίκους μιας άλλης περιοχής; Για τι είδους εγκλήματα μιλάμε;

Έχουν σταθμιστεί τα είδη εγκλήματος που εξετάστηκαν στην ανάλυση ή μήπως η κλοπή παρουσιάζεται ισοδύναμη με φόνο; Τέτοιες προσπάθειες «μελέτης πεδίου» θα σας επιτρέψουν να αξιολογήσετε τα δεδομένα σας και να αποφασίσετε σε ποιον βαθμό μπορείτε να εμπιστευτείτε τα αποτελέσματα της περαιτέρω ανάλυσης.

Η συγγραφέας

Η Κρίνα-Γκαμπριέλα Μπόρος είναι βραβευμένη ερευνήτρια δημοσιογράφος με έδρα το Ηνωμένο Βασίλειο, η οποία διδάσκει δημοσιογραφία με γνώμονα τα δεδομένα διεθνώς και έχει δημιουργήσει ένα πρόγραμμα κατάρτισης για τη στρατηγική δημοσιογραφία δεδομένων στο ετήσιο συνέδριο Dataharvest.

54. http://news.trust.org/spotlight/most-dangerous-transport-systems-for-women/?tab=methodology

4.

Δουλεύοντας με Δεδομένα

4.1

Δημοσιογραφία Δεδομένων: Πώς Σχετίζεται ο Φεμινισμός με την Τεχνολογία Πληροφοριών

Catherine D'Ignazio

Περίληψη: Η υιοθέτηση μιας φεμινιστικής προσέγγισης στη δημοσιογραφία δεδομένων σημαίνει ότι πρέπει να κατανοήσουμε τους τρόπους με τους οποίους η ανισότητα υπεισέρχεται στις βάσεις δεδομένων και στους αλγορίθμους, καθώς και να αναπτύξουμε στρατηγικές για τον μετριασμό αυτών των αποκλίσεων.

Λέξεις-κλειδιά: δημοσιογραφία δεδομένων, φεμινισμός, φύλο, ηθική, ανισότητα, βάσεις δεδομένων

Η πρόοδος της τεχνολογίας τα τελευταία 70 χρόνια έχει δώσει στους ανθρώπους τη δυνατότητα να αποθηκεύουν και να επεξεργάζονται περισσότερες πληροφορίες παρά ποτέ. Τα χρήματα που κερδίζουν οι πιο επιτυχημένες εταιρείες τεχνολογίας στον κόσμο –η Google, η Facebook[55], η Amazon, η Microsoft, η Apple– προέρχονται από τη συγκέντρωση δεδομένων. Τόσο σε επιχειρηματικό όσο και σε κυβερνητικό επίπεδο, θεωρείται ολοένα και πιο σημαντική η λήψη αποφάσεων «βάσει δεδομένων». Τα δεδομένα είναι ισχυρά –επειδή είναι οικονομικά προσοδοφόρα και τα εκτιμούν οι ισχυροί–, αλλά δεν κατανέμονται ισότιμα, ούτε αυτά, ούτε οι δεξιότητες που απαιτούνται για να εργαστεί κάποιος με δεδομένα, ούτε οι τεχνολογικοί πόροι οι οποίοι είναι απαραίτητοι για την αποθήκευση και την επεξεργασία τους. Οι άνθρωποι που εργάζονται με δεδομένα δεν είναι αντιπροσωπευτικό δείγμα του γενικού πληθυσμού –είναι δυσανάλογα πολλοί άνδρες, λευκοί, συνήθως από το βόρειο ημισφαίριο και με υψηλό μορφωτικό επίπεδο.

Ακριβώς λόγω αυτών των βασικών ανισοτήτων στο οικοσύστημα των δεδομένων, η υιοθέτηση μιας φεμινιστικής προσέγγισης στη δημοσιογραφία δεδομένων μπορεί να είναι χρήσιμη για την αποκάλυψη μιας κρυφής μεροληψίας στη ροή της πληροφορίας. Ο φεμινισμός μπορεί να οριστεί απλά ως η πίστη στην κοινωνική, πολιτική και οικονομική ισότητα των δύο φύλων και η οργανωμένη δραστηριότητα για λογαριασμό αυτής της πεποίθησης. Οι φεμινιστικές έννοιες και τα φεμινιστικά εργαλεία μπορούν να είναι χρήσιμα, για να εξετάσουμε την κοινωνική ισχύ θέτοντας το φύλο ως κεντρική (αλλά όχι τη μοναδική) διάσταση της ανάλυσης. Ένα από τα καθοριστικά χαρακτηριστικά του σύγχρονου φεμινισμού είναι η επιμονή του στη διατομεακότητα –την ιδέα ότι πρέπει να λαμβάνουμε υπόψη όχι μόνο τον σεξισμό, αλλά και τον ρατσισμό, τον ταξισμό, τον ικανισμό και άλλες δομικές επιρροές, όταν συζητάμε πώς η ανισότητα στη διανομή της εξουσίας σκιάζει την αλήθεια[56]. Για τους δημοσιογράφους που συντάσσονται με

55. ΣτΕ: Νυν «Meta» μετά την αλλαγή της επωνυμίας της εταιρείας που ανακοινώθηκε τον Οκτώβριο 2021.
56. Πράγματι, ο φεμινισμός που δεν λαμβάνει υπόψη πώς αλληλεπιδρούν με το φύλο άλλοι παράγοντες ταυτότητας θα πρέπει να χαρακτηριστεί ως «λευκός φεμινισμός». Η διατομεακότητα ορίστηκε για πρώτη φορά από τη νομικό Κίμπερλι Κρένσο και προέρχεται από την πνευματική κληρονομιά των μαύρων φεμινιστριών, που υποστήριξαν ότι η ανισότητα των φύλων δεν μπορεί να εξεταστεί απομονωμένη από την ανισότητα που σχετίζεται με τη φυλή και την κοινωνική τάξη.

την επαγγελματική στάση «λέμε την αλήθεια στην εξουσία»[57], η φεμινιστική προσέγγιση μπορεί να μοιάζει αρκετά οικεία.

Αυτό το κεφάλαιο εξετάζει διάφορα στάδια στη διαδικασία επεξεργασίας δεδομένων –συλλογή δεδομένων, πλαίσιο των δεδομένων και επικοινωνία δεδομένων– και επισημαίνει τα σημεία όπου η προκατάληψη μπορεί να καιροφυλακτεί σαν παγίδα, καθώς και τις ευκαιρίες για τη χρήση μιας φεμινιστικής οπτικής. Ας σημειωθεί ότι η φεμινιστική προσέγγιση δεν είναι χρήσιμη μόνο για δεδομένα που αφορούν γυναίκες ή θέματα φύλου, αλλά είναι κατάλληλη για κάθε έργο που αφορά ανθρώπους ή ανθρώπινους θεσμούς (δηλαδή, σχεδόν για οποιοδήποτε έργο), διότι όπου υπάρχουν άνθρωποι υπάρχει κοινωνική ανισότητα[58].

Συλλογή δεδομένων

Ο έλεγχος της εξουσίας –πώς λειτουργεί αυτή και ποιος επωφελείται– αποτελούσε πάντα κεντρικό στοιχείο των φεμινιστικών εγχειρημάτων. Η θεωρία της κοινωνιολόγου Πατρίσια Χιλ Κόλινς για το *πλέγμα της κυριαρχίας* («matrix of domination») μας βοηθά να κατανοήσουμε ότι η εξουσία είναι περίπλοκη και ότι «ελάχιστα είναι τα ατόφια θύματα και οι ατόφιοι καταπιεστές» (Hill Collins, 2000, σ. 332). Ενώ τείνουμε να σκεφτόμαστε την αδικία στον διαπροσωπικό τομέα (π.χ. ένα σεξιστικό σχόλιο), υπάρχουν συστημικές δυνάμεις που πρέπει να κατανοήσουμε και να αναδείξουμε (π.χ. ο σεξισμός στα θεσμικά όργανα που συλλέγουν δεδομένα), προκειμένου να επέλθει αλλαγή.

Η διαρθρωτική ανισότητα εμφανίζεται στη συλλογή δεδομένων με δύο τρόπους. Πρώτον, συγκεκριμένοι φορείς υπερκαταμετρώνται στη διαδικασία συλλογής δεδομένων. Η υπερκαταμέτρηση σχετίζεται συνήθως με την επιτήρηση που ασκούν οι έχοντες εξουσία σε όσους έχουν λιγότερη εξουσία. Για παράδειγμα, η αστυνομία της Βοστώνης δημοσίευσε στοιχεία σχετικά με το πρόγραμμα «stop-and-frisk»[59] το 2015. Τα στοιχεία δείχνουν ότι η αστυνομία περιπολεί δυσανάλογα συχνά σε γειτονιές μαύρων, μεταναστών και ισπανόφωνων και ελέγχει δυσανάλογα συχνά νεαρούς μαύρους άνδρες. Σε τέτοιες περιπτώσεις υπερκαταμέτρησης, είναι σημαντικό να γνωρίζουμε ποιες ομάδες κατέχουν εξουσία και ποιες ομάδες είναι πιθανό να αποτελέσουν στόχο της επιτήρησης. Δουλειά του δημοσιογράφου δεδομένων είναι να αναγνωρίσει και να ποσοτικοποιήσει την ανισότητα, καθώς και να κατονομάσει τις διαρθρωτικές δυνάμεις που δρουν –στην προκείμενη περίπτωση, τον ρατσισμό.

Ο δεύτερος τρόπος με τον οποίο η διαρθρωτική ανισότητα εμφανίζεται στη συλλογή δεδομένων είναι η υποκαταμέτρηση ή η μη καταμέτρηση. Για παράδειγμα, γιατί η πιο ολοκληρωμένη βάση δεδομένων για τις γυναικοκτονίες (δολοφονίες με βάση το φύλο) στο Μεξικό έχει συγκροτηθεί από μια γυναίκα με το ψευδώνυμο Πρινσέσα;[60] Παρά το γεγονός ότι οι θάνατοι γυναικών στη Σιουδάδ Χουάρες και σε ολόκληρη τη χώρα συνεχίζουν να αυξάνονται, παρά τη σύσταση ειδικής επιτροπής για τις γυναικοκτονίες το 2006, παρά την απόφαση του Διαμερικανικού Δικαστηρίου Ανθρωπίνων Δικαιωμάτων κατά του μεξικανικού κράτους το 2009, η πολιτεία δεν καταγράφει συστηματικά τις γυναικοκτονίες. Η υποκαταμέτρηση παρατηρείται σε πολλά ζητήματα που αφορούν τις γυναίκες και τα έγχρωμα άτομα, στα οποία οι θεσμοί καταμέτρησης παραλείπουν συστηματικά να λογοδοτήσουν για βλάβες για τις οποίες είναι οι ίδιοι υπεύθυ-

57. ΣτΜ: Πασιφιστικό σύνθημα που έχει τις ρίζες του στο Κίνημα για τα Πολιτικά Δικαιώματα και στόχο την επίτευξη κοινωνικής δικαιοσύνης.
58. Για παράδειγμα, βλ. το κεφάλαιο των Κουκουτάι και Γουόλτερ για την κυριαρχία των αυτοχθόνων επί των δεδομένων τους. (ΣτΕ: Κεφάλαιο 3.3, «Κυριαρχία των Αυτοχθόνων επί των Δεδομένων τους: Επιπτώσεις για τη Δημοσιογραφία Δεδομένων» στο παρόν έργο)
59. ΣτΜ: Μια πρακτική προσωρινής κράτησης, ανάκρισης και ενίοτε έρευνας υπόπτων στον δρόμο για όπλα και άλλα προϊόντα παράνομης διακίνησης.
60. https://feminicidiosmx.crowdmap.com/

νοι. Αυτό σημαίνει ότι το σύστημα διαχείρισης είναι έκθετο. Σε περιπτώσεις υποκαταμέτρησης, οι δημοσιογράφοι δεδομένων μπορούν να κάνουν ακριβώς αυτό που έκανε η Πρινσέσα: να μετρήσουν οι ίδιοι, στο μέτρο των δυνατοτήτων τους. Συνήθως, αυτό περιλαμβάνει έναν συνδυασμό πληθοπορισμού («crowdsourcing»), συλλογής ιστοριών και στατιστικών συμπερασμάτων. Στις ΗΠΑ, άλλα παραδείγματα υποκαταμέτρησης περιλαμβάνουν τις δολοφονίες από την αστυνομία και τη μητρική θνησιμότητα. Δημοσιογράφοι έχουν αναλάβει να εκπονήσουν έργα συλλογής δεδομένων και για τα δυο αυτά ζητήματα.

Πλαίσιο των δεδομένων

Ενώ το κίνημα των ανοιχτών δεδομένων και ο πολλαπλασιασμός των API[61] φαντάζει ως κάτι καλό για τους δημοσιογράφους δεδομένων, τα δεδομένα που αποκτώνται «στην άγρια φύση» συνοδεύονται από ένα σύνολο προβληματισμών, ιδίως όταν πρόκειται για ανθρώπινα και κοινωνικά φαινόμενα. Η φεμινίστρια φιλόσοφος Ντόνα Χαραγουέι αναφέρει ότι η κάθε γνώση είναι «τοποθετημένη» κάπου, πράγμα που σημαίνει ότι βρίσκεται πάντα σε ένα κοινωνικό, πολιτισμικό, ιστορικό και υλικό πλαίσιο. Η διαλεύκανση και η διερεύνηση του τρόπου με τον οποίο τα σύνολα δεδομένων είναι προϊόντα αυτών των πλαισίων μπορεί να μας βοηθήσει να κατανοήσουμε τους τρόπους με τους οποίους η εξουσία και τα προνόμια μπορεί να σκιάζουν την αλήθεια.

Για παράδειγμα, οι φοιτητές μου στο μάθημα οπτικοποίησης δεδομένων ήθελαν να γράψουν μια τελική εργασία σχετικά με τις σεξουαλικές επιθέσεις στις πανεπιστημιουπόλεις[62]. Τα κολέγια και τα πανεπιστήμια στις Ηνωμένες Πολιτείες υποχρεούνται να αναφέρουν ετησίως τις σεξουαλικές επιθέσεις και άλλα εγκλήματα στον χώρο της πανεπιστημιούπολης, σύμφωνα με τον Νόμο Κλέρι. Επομένως, θα έπρεπε να υπάρχει μια ολοκληρωμένη εθνική βάση δεδομένων για το ζήτημα. Ωστόσο, τα δεδομένα του Νόμου Κλέρι προβληματίσαν τους φοιτητές –το Κολέγιο Γουίλιαμς, για παράδειγμα, είχε εξαιρετικά υψηλούς αριθμούς σε σύγκριση με άλλα αστικά κολέγια. Οι φοιτητές διαπίστωσαν, διερευνώντας το πλαίσιο και εξετάζοντας τον τρόπο με τον οποίο ήταν δομημένο το περιβάλλον συλλογής, ότι οι αριθμοί έλεγαν μια ιστορία που κατά πάσα πιθανότητα ήταν το αντίθετο της αλήθειας. Η σεξουαλική επίθεση είναι ένα ζήτημα που στιγματίζει και οι επιζώντες συχνά φοβούνται την ενοχοποίηση του θύματος και τα αντίποινα και δεν προβαίνουν σε καταγγελίες. Έτσι, τα κολέγια που κατέγραφαν υψηλούς αριθμούς ήταν εκείνα που είχαν αφιερώσει περισσότερους πόρους για τη δημιουργία ενός περιβάλλοντος στο οποίο οι επιζώντες θα αισθάνονταν ασφαλείς, ώστε να υποβάλουν καταγγελία. Αντιθέτως, εκείνα που ανέφεραν χαμηλό αριθμό σεξουαλικών επιθέσεων είχαν ένα εχθρικό κλίμα που δεν υποστήριζε τους επιζώντες να σπάσουν τη σιωπή τους.

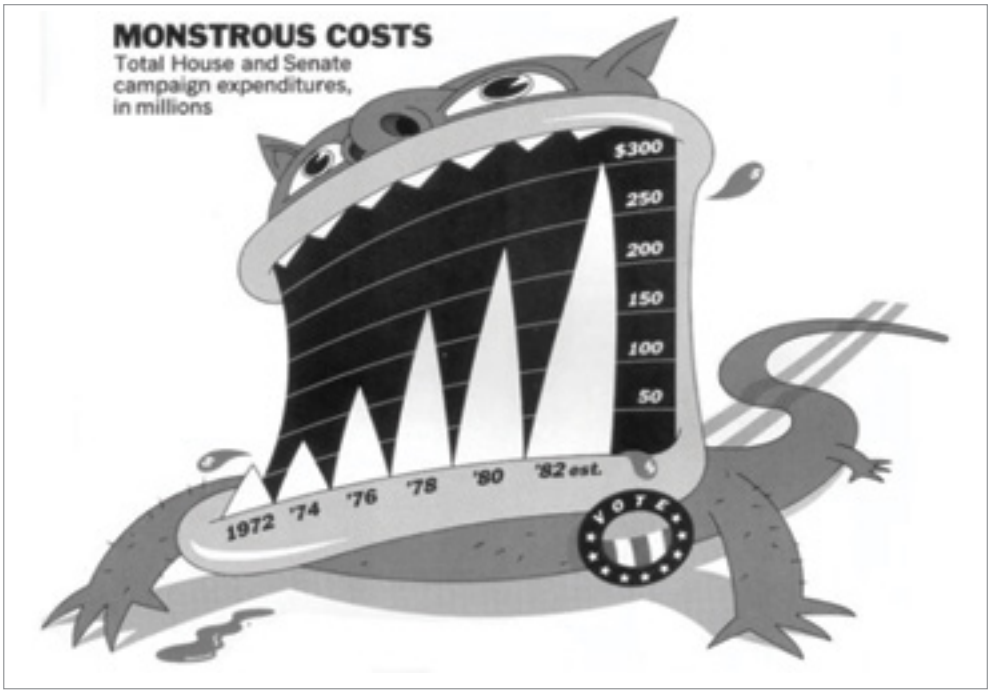
Εδώ, υπάρχει μια παγίδα και μια ευκαιρία. Η παγίδα είναι ότι οι δημοσιογράφοι παίρνουν τους αριθμούς που κατεβάζουν από το Διαδίκτυο τοις μετρητοίς, χωρίς να κατανοούν τις αποχρώσεις του περιβάλλοντος συλλογής, συμπεριλαμβανομένων των ανισοτήτων εξουσίας, του κοινωνικού στίγματος και των πολιτισμικών κανόνων σχετικά με την πιθανότητα να εμφανιστεί κάποιος στα θεσμικά όργανα (π.χ. ομάδες όπως οι γυναίκες, οι μετανάστες και οι έγχρωμοι αισθάνονται γενικά λιγότερη εμπιστοσύνη απέναντι στους θεσμούς καταμέτρησης, και μάλιστα με εξαιρετικά σοβαρό λόγο). Η ευκαιρία είναι ότι υπάρχουν ακόμη πολλές ιστορίες σχετικά με το πλαίσιο των δεδομένων που πρέπει να αφηγηθούμε. Αντί να χρησιμοποιούν πάντα τους αριθμούς αναζητώντας νέες αναλύσεις, οι δημοσιογράφοι δεδομένων μπορούν να χρησιμοποιήσουν τους αριθμούς, για να εξετάσουν το περιβάλλον συλλογής, να επισημάνουν τις εσφαλμένες πρακτικές και τις ανισορροπίες εξουσίας και να αλλάξουν τις πρακτικές καταμέτρησης, ώστε τα θεσμικά όργανα να καταγράφουν αυτό που πραγματικά έχει σημασία.

61. ΣτΜ: Από το «Application Programming Interface» (διεπαφή προγραμματισμού εφαρμογών)
62. Το τελευταίο άρθρο των Πάτρικ Τόρφι, Μισέλ Γκανιόν και Τζίλιαν Μίχαν είναι δημοσιευμένο εδώ: https://cleryactfallsshort.atavist.com/reporting-sexual-assault-what-the-clery-act-doesnt-tell-us

Επικοινωνία των δεδομένων

Η σύγχρονη δυτική σκέψη για τα δεδομένα έχει προκύψει από ένα «κυρίαρχο στερεότυπο» όπου αυτό που θεωρείται ορθολογικό και αντικειμενικό εκτιμάται περισσότερο από αυτό που θεωρείται συναισθηματικό και υποκειμενικό. (Σκεφτείτε ποιο φύλο προσδιορίζεται ως «λογικό» και ποιο ως «συναισθηματικό».) Το κυρίαρχο στερεότυπο λέει ότι τα συναισθήματα θολώνουν την κρίση και η απόσταση ενισχύει την αντικειμενικότητα. Μια φεμινιστική θεώρηση, όμως, αμφισβητεί τα πάντα σχετικά με αυτό το κυρίαρχο στερεότυπο. Τα συναισθήματα δεν θολώνουν την κρίση –καλλιεργούν την περιέργεια, τη συμμετοχή και το κίνητρο για να μάθουμε περισσότερα. Η Πατρίσια Χιλ Κόλινς, για παράδειγμα, περιγράφει μια ιδανική κατάσταση γνώσης ως μια κατάσταση στην οποία «ούτε η ηθική ούτε τα συναισθήματα υποτάσσονται στη λογική» (Χιλ Κόλινς, 2000, σ. 266).

Τι σημαίνει αυτό για την επικοινωνία των δεδομένων; Ενώ οι προηγούμενες πρακτικές στην οπτικοποίηση δεδομένων ευνοούσαν τα μινιμαλιστικά διαγράμματα και γραφικά ως πιο ορθολογικά, τόσο οι ερευνητές όσο και οι δημοσιογράφοι μαθαίνουν τώρα ότι η αξιοποίηση των μοναδικών ιδιοτήτων της οπτικοποίησης ως μιας μορφής δημιουργικής ρητορικής αποδίδει πιο αξιομνημόνευτα γραφικά προς διαμοιρασμό. Πάρτε, για παράδειγμα, το γράφημα Monstrous Costs («Τερατώδες Κόστος») που δημιουργήθηκε από τον Νάιτζελ Χολμς το 1984, για να απεικονίσει το αυξανόμενο κόστος των πολιτικών εκστρατειών. Αν και στο παρελθόν χλευάστηκε ως μια περίπτωση «σκουπιδοδιαγράμματος», οι ερευνητές έχουν πλέον αποδείξει αυτό που οι περισσότεροι από εμάς γνωρίζουμε διαισθητικά: σε ορισμένους αναγνώστες, αρέσουν τα τέρατα περισσότερο από τα βαρετά ραβδογράμματα (Bateman κ.α., 2010).



Εικόνα 4.1.1: Γράφημα του Νάιτζελ Χολμς από το βιβλίο *Designer's Guide to Creating Charts and Diagrams* (Χολμς, 1984).

Όπως συμβαίνει με κάθε άλλο μέσο επικοινωνίας, η αξιοποίηση του συναισθήματος στα δεδομένα συνεπάγεται ηθικές ευθύνες. Οι ερευνητές κατέδειξαν, επίσης πρόσφατα, τη σημασία του τίτλου

μιας οπτικοποίησης στον τρόπο με τον οποίο οι άνθρωποι ερμηνεύουν το διάγραμμα (Borkin κ.α., 2016). Οι συνήθεις πρακτικές απόδοσης τίτλου τείνουν προς το «ορθολογικό», δηλαδή παρουσιάζουν τα δεδομένα ως ουδέτερα και αντικειμενικά –ας πούμε, «Αναφορές σεξουαλικών επιθέσεων σε πανεπιστημιουπόλεις 2012-2014». Υπάρχουν, όμως, πολλές περιπτώσεις –που και πάλι συνήθως έχουν να κάνουν με γυναίκες και άλλες περιθωριοποιημένες ομάδες– στις οποίες ένας ουδέτερος τίτλος στην πραγματικότητα βλάπτει την ομάδα που προβάλλεται από τα δεδομένα.

Στην περίπτωση των σεξουαλικών επιθέσεων, για παράδειγμα, ένας ουδέτερος τίτλος δηλώνει σιωπηρά ότι τα δεδομένα που διαθέτουμε είναι αληθή και πλήρη, ενώ στην πραγματικότητα γνωρίζουμε ότι αυτό είναι εντελώς ψευδές. Σε άλλες περιπτώσεις, ένας ουδέτερος τίτλος, όπως «Ψυχικά ασθενείς γυναίκες που σκοτώθηκαν σε αντιπαραθέσεις με την αστυνομία 2012-2014», ανοίγει την πόρτα στη διαιώνιση επιβλαβών στερεοτύπων, ακριβώς επειδή δεν κατονομάζει τις δομικές δυνάμεις που δρουν, συμπεριλαμβανομένου του ικανισμού και του σεξισμού, και οι οποίες καθιστούν τις ψυχικά ασθενείς γυναίκες δυσανάλογα συχνά θύματα αστυνομικής βίας στις Ηνωμένες Πολιτείες.

Συμπέρασμα

Η υιοθέτηση μιας φεμινιστικής προσέγγισης στη δημοσιογραφία δεδομένων σημαίνει ότι πρέπει να συντονιστούμε με τους τρόπους με τους οποίους οι υπάρχοντες θεσμοί και πρακτικές ευνοούν ένα status quo στο οποίο οι άνδρες της ελίτ βρίσκονται στην κορυφή και οι άλλοι τοποθετούνται σε διάφορες διασταυρώσεις στο πλέγμα κυριαρχίας της Κόλινς. Η πατριαρχία, η λευκή υπεροχή και η αποικιοκρατία των εποίκων είναι δομικές δυνάμεις. Επομένως, προσφέρονται ιδιαίτερα για συστημική έρευνα και οπτικοποίηση με βάση τα δεδομένα. Πρέπει να αμφισβητήσουμε αρκετά την παραδεδομένη σοφία στη δημοσιογραφία δεδομένων, για να διασφαλίσουμε ότι δεν διαιωνίζουμε ακούσια αυτό το status quo και, ταυτόχρονα, να χρησιμοποιήσουμε τα εργαλεία μας, για να αποκαλύψουμε και να καταργήσουμε την αδικία. Όσοι επιθυμούν να προχωρήσουν περισσότερο προς αυτήν την κατεύθυνση μπορούν να ανατρέξουν στο βιβλίο μου *Data Feminism* (2020), σε συνεργασία με τη Λόρεν Φρ. Κλάιν, που παρουσιάζει με περισσότερες λεπτομέρειες τον τρόπο με τον οποίο οι φεμινιστικές έννοιες μπορούν να εφαρμοστούν στην επιστήμη των δεδομένων και στην επικοινωνία των δεδομένων.

Παραπομπές

Bateman, S., Mandryk, R. L., Gutwin, C., Genest, A., McDine, D., & Brooks, C. (2010) «Useful junk? The effects of visual embellishment on comprehension and memorability of charts», στο *Proceedings of the 28th International Conference on Human Factors in Computing Systems—CHI '10*, σσ. 2573-2582. Ατλάντα, Τζόρτζια, ΗΠΑ, 10-15 Απριλίου. DOI: https://doi.org/10.1145/1753326.1753716

Borkin, M. A., Bylinskii, Z., Kim, N. W., Bainbridge, C. M., Yeh, C. S., Borkin, D., Pfister, H., & Oliva, A. (2016) «Beyond memorability: Visualization recognition and recall», IEEE *Transactions on Visualization and Computer Graphics,* 22(1), σσ. 519-528. DOI: https://doi.org/10.1109/TVCG.2015.2467732

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism.* Κέιμπριτζ, Μασαχουσέτη: MIT Press.

Hill Collins, P. (2000) *Black feminist thought: Knowledge, consciousness and the politics of empowerment.* Νέα Υόρκη: Routledge. Διαθέσιμο σε: http://www.hartford-hwp.com/archives/45a/252.html (Προσπελάστηκε: 25 Νοεμβρίου 2021).

Η συγγραφέας

Η Κάθριν Ντ' Ιγκνάζιο είναι Επίκουρη Καθηγήτρια Πολεοδομίας και Χωροταξίας και διευθύντρια του Εργαστηρίου Δεδομένων και Φεμινισμού στο Ινστιτούτο Τεχνολογίας της Μασαχουσέτης (MIT).

4.2

Οικοδόμηση Συνεργασιών για τα *Panama* και *Paradise Papers*

Emilia Díaz-Struck, Cécile Schilis-Gallego, Pierre Romera

Περίληψη: Πώς η Διεθνής Σύμπραξη Ερευνητών Δημοσιογράφων (International Consortium of Investigative Journalists, ICIJ) καθιστά πιο αποτελεσματική την έρευνα σε γιγάντιους όγκους εγγράφων και δεδομένων.

Λέξεις-κλειδιά: διαρροές δεδομένων, εξαγωγή κειμένου, ριζοσπαστικός διαμοιρασμός, διασυνοριακή έρευνα, δημοσιογραφία δεδομένων, Διεθνής Σύμπραξη Ερευνητών Δημοσιογράφων (ICIJ)

Η Διεθνής Σύμπραξη Ερευνητών Δημοσιογράφων («International Consortium of Investigative Journalists», ICIJ) είναι ένα διεθνές δίκτυο δημοσιογράφων που ιδρύθηκε το 1997. Οι δημοσιογράφοι που συμμετέχουν στις μεγάλες συνεργασίες του ICIJ έχουν διαφορετικά υπόβαθρα και προφίλ. Είναι ένα ευρύ φάσμα δημοσιογράφων με διαφορετικές δεξιότητες, κάποιοι με ισχυρές δεξιότητες στα δεδομένα και στον προγραμματισμό, άλλοι που διαθέτουν τις καλύτερες πηγές και δεξιότητες στο παραδοσιακό ρεπορτάζ. Όλους τους ενώνει το ενδιαφέρον για τη δημοσιογραφία, τη συνεργασία και τα δεδομένα.

Όταν ο διευθυντής του ICIJ, Τζέραρντ Ράιλ, έλαβε στην Αυστραλία έναν σκληρό δίσκο με εταιρικά δεδομένα που αφορούσαν φορολογικούς παραδείσους και ανθρώπους σε ολόκληρο τον κόσμο, ως αποτέλεσμα της τριετούς έρευνάς του για το σκάνδαλο Firepower στην Αυστραλία, δεν μπορούσε εκείνη τη στιγμή να φανταστεί πώς αυτό θα μεταμόρφωνε την ιστορία των συνεργασιών στη δημοσιογραφία. Ο ίδιος έφτασε στο ICIJ το 2011 με περισσότερα από 260GB δεδομένων σχετικά με υπεράκτιες εταιρείες –περίπου 2,5 εκατομμύρια αρχεία, που κατέληξαν σε μια συνεργασία περισσότερων από 86 δημοσιογράφων από 46 χώρες, γνωστή ως *Offshore Leaks* (δημοσιεύθηκε το 2013)[63].

Μετά τα *Offshore Leaks* ακολούθησαν περισσότερα ερευνητικά έργα με μεγάλα σύνολα δεδομένων και εκατομμύρια αρχεία, περισσότερες εξειδικευμένες τεχνολογίες για τη διερεύνησή τους και περισσότερα δίκτυα δημοσιογράφων για τη δημοσιοποίησή τους. Για παράδειγμα, πρόσφατα κοινοποιήσαμε στους συνεργάτες μας ένα νέο εύρημα 1,2 εκατομμυρίων εγγράφων που διέρρευσαν από την ίδια δικηγορική εταιρεία η οποία βρίσκεται στο επίκεντρο της έρευνας για τα *Panama Papers*, τη Mossack Fonseca[64]. Το εύρημα αυτό προστέθηκε στα 11,5 εκατομμύρια αρχεία των *Panama Papers* που μας έφερε το 2015 η γερμανική εφημερίδα *Süddeutsche Zeitung* και στα 13,6 εκατομμύρια έγγραφα που αποτέλεσαν τη βάση για την επακόλουθη έρευνα των *Paradise Papers*[65].

Εάν ένας δημοσιογράφος μόνος του αφιέρωνε ένα λεπτό για να διαβάσει κάθε αρχείο των *Paradise Papers*, θα χρειαζόταν 26 χρόνια για να τα διατρέξει όλα. Προφανώς, αυτό δεν είναι

63. https://www.icij.org/investigations/offshore/how-icijs-project-team-analyzed-offshore-files/
64. https://www.icij.org/investigations/panama-papers/
65. https://www.icij.org/investigations/paradise-papers/

εφικτό. Έτσι, αναρωτηθήκαμε, πώς θα μπορούσαμε να συντομεύσουμε τη διαδικασία. Πώς θα μπορούσαμε να κάνουμε την έρευνα πιο αποτελεσματική και λιγότερο χρονοβόρα; Πώς θα μπορούσε η τεχνολογία να μας βοηθήσει να βρούμε νέα στοιχεία σε τούτο το γιγαντιαίο απόθεμα εγγράφων και να υποστηρίξουμε το συνεργατικό μας μοντέλο;

Σε αυτό το κεφάλαιο δείχνουμε πώς αντιμετωπίζουμε μεγάλες συλλογές εγγράφων που έχουν διαρρεύσει όχι μόνο μέσω εξελιγμένων τεχνολογιών «μεγάλων δεδομένων» («big data»), αλλά μάλλον μέσω ενός εξειδικευμένου αναλυτικού μηχανισμού που περιλαμβάνει: (α) διεθνή συνεργατικά δίκτυα, (β) ασφαλείς πρακτικές και υποδομές επικοινωνίας, (γ) διαδικασίες και συστήματα για τη δημιουργία δομημένων δεδομένων από αδόμητα έγγραφα και (δ) βάσεις δεδομένων γραφημάτων και διερευνητικές οπτικοποιήσεις για τη μελέτη συνδέσεων μεταξύ τους.

Αλληλεπίδραση με τους συνεργάτες

Το μοντέλο του ICIJ είναι να ερευνά το διεθνές φορολογικό σύστημα με ένα παγκόσμιο δίκτυο δημοσιογράφων. Συγκεντρώνουμε κορυφαίους δημοσιογράφους σε πέντε ηπείρους, για να βελτιώσουμε τις ερευνητικές προσπάθειες και να ενώσουμε τις κουκίδες των δεδομένων από τη μία χώρα στην άλλη[66].

Οι φορολογικές ιστορίες είναι σαν παζλ με κομμάτια που λείπουν: ένας ρεπόρτερ στην Εσθονία μπορεί να καταλάβει ένα τμήμα της ιστορίας, ένας ρεπόρτερ στη Βραζιλία μπορεί να κατανοήσει ένα άλλο τμήμα. Εάν φέρουμε αυτούς τους δυο σε επαφή, τότε έχουμε μια πληρέστερη εικόνα. Η δουλειά του ICIJ είναι να συνδέει αυτούς τους δημοσιογράφους και να διασφαλίζει ότι θα μοιραστούν οτιδήποτε ανακαλύπτουν στα δεδομένα.

Ονομάζουμε αυτή μας τη φιλοσοφία «ριζοσπαστικό διαμοιρασμό» («radical sharing»): οι συνεργάτες του ICIJ κοινοποιούν τα ευρήματά τους κατά τη διάρκεια της εργασίας τους, όχι μόνο στους άμεσους συναδέλφους τους, αλλά και σε δημοσιογράφους στην άλλη άκρη του κόσμου.

Για την προώθηση της συνεργασίας, το ICIJ παρέχει μια πλατφόρμα επικοινωνίας που ονομάζεται Global I-Hub και βασίζεται σε στοιχεία λογισμικού ανοιχτού κώδικα[67]. Έχει περιγραφεί από τους χρήστες της ως ένα «ιδιωτικό Facebook» και επιτρέπει το ίδιο είδος άμεσης ανταλλαγής πληροφοριών που λαμβάνει χώρα σε μια αίθουσα σύνταξης στον φυσικό κόσμο. Οι δημοσιογράφοι συμμετέχουν σε ομάδες που παρακολουθούν συγκεκριμένα θέματα –χώρες, αθλήματα, τέχνες, δικαστικές υποθέσεις ή οποιοδήποτε άλλο θέμα ενδιαφέροντος. Μέσα σε αυτές τις ομάδες, μπορούν να δημοσιεύουν ακόμα πιο συγκεκριμένα θέματα, όπως στοιχεία που εντόπισαν στα δεδομένα για έναν πολιτικό ή μια συγκεκριμένη συναλλαγή που εξετάζουν. Εδώ λαμβάνει χώρα το μεγαλύτερο μέρος της συζήτησης, όπου οι δημοσιογράφοι διασταυρώνουν πληροφορίες και μοιράζονται σημειώσεις και ενδιαφέροντα έγγραφα.

Το ICIJ χρειάστηκε να υλοποιήσει αρκετά έργα μέχρι να εξοικειωθούν οι δημοσιογράφοι με την πλατφόρμα I-Hub. Οι συντονιστές του ICIJ που είναι υπεύθυνοι για κάθε περιοχή προσφέρουν υποστήριξη, προκειμένου να διευκολύνουν την είσοδο των δημοσιογράφων στην πλατφόρμα και να αντιμετωπιστούν οποιαδήποτε τεχνικά ζητήματα. Αυτό είναι το κλειδί για να εξασφαλιστεί ότι οι δημοσιογράφοι ακολουθούν τα απαιτούμενα πρότυπα ασφαλείας.

66. http://www.icij.org/journalists

67. Βλ. https://www.icij.org/blog/2014/07/icij-build-global-i-hub-new-secure-collaboration-tool/. Για μια διαφορετική οπτική σχετικά με τις δημοσιογραφικές πλατφόρμες όπως το I-Hub, βλ. το κεφάλαιο του Καντέα σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 7.6, «Φεουδαρχία Δεδομένων: Πώς οι Πλατφόρμες Διαμορφώνουν Διασυνοριακά Ερευνητικά Δίκτυα»).

Κρυπτογραφώντας τα πάντα

Όταν διεξάγεται μια έρευνα στην οποία συμμετέχουν 396 δημοσιογράφοι, πρέπει να είστε ρεαλιστές όσον αφορά στην ασφάλεια: κάθε άτομο αποτελεί πιθανό στόχο επίθεσης και ο κίνδυνος παραβίασης είναι υψηλός. Για να μετριάσει αυτόν τον κίνδυνο, το ICIJ χρησιμοποιεί πολλαπλές αμυντικές μεθόδους.

Για τη συμμετοχή σε μια έρευνα του ICIJ είναι υποχρεωτική η δημιουργία ενός ζεύγους κλειδιών PGP για την κρυπτογράφηση των μηνυμάτων ηλεκτρονικού ταχυδρομείου. Η αρχή του PGP είναι απλή[68]. Υπάρχουν δύο κλειδιά: το ένα είναι δημόσιο και κοινοποιείται σε κάθε πιθανό συνομιλητή, ο οποίος μπορεί να το χρησιμοποιήσει για να σας στείλει κρυπτογραφημένα μηνύματα ηλεκτρονικού ταχυδρομείου. Το δεύτερο κλειδί είναι ιδιωτικό και δεν πρέπει ποτέ να φύγει από τον υπολογιστή σας. Το ιδιωτικό κλειδί εξυπηρετεί μόνο έναν σκοπό: την αποκρυπτογράφηση των μηνυμάτων ηλεκτρονικού ταχυδρομείου που έχουν κρυπτογραφηθεί με το δημόσιο κλειδί σας.

Σκεφτείτε το PGP ως μια θυρίδα ασφαλείας όπου οι άνθρωποι μπορούν να αποθηκεύουν μηνύματα για εσάς. Μόνο εσείς έχετε το κλειδί για να την ανοίξετε και να διαβάσετε τα μηνύματα. Όπως κάθε μέτρο ασφαλείας, το PGP έχει τα τρωτά του σημεία. Για παράδειγμα, μπορεί εύκολα να παραβιαστεί, εάν στον υπολογιστή σας εκτελείται λογισμικό spyware, το οποίο καταγράφει τις λέξεις που πληκτρολογείτε ή κατασκοπεύει κάθε αρχείο στον δίσκο σας. Αυτό υπογραμμίζει τη σημασία της ύπαρξης πολλαπλών επιπέδων ασφάλειας. Εάν ένα από αυτά τα επίπεδα αποτύχει, ελπίζουμε ότι τα άλλα επίπεδα θα περιορίσουν τον αντίκτυπο της παραβίασης.

Για να πιστοποιήσει την ταυτότητα των συνεργατών του, το ICIJ εφαρμόζει έλεγχο ταυτότητας δύο παραγόντων σε όλες τις πλατφόρμες του. Αυτή η τεχνική είναι πολύ δημοφιλής σε μεγάλους ιστοτόπους, όπως το Google, το Twitter και το Facebook. Παρέχει στον χρήστη έναν δεύτερο, προσωρινό κωδικό που απαιτείται για τη σύνδεση, ο οποίος συνήθως δημιουργείται σε διαφορετική συσκευή (π.χ. στο τηλέφωνό σας) και έχει μικρή χρονική διάρκεια. Σε ορισμένες ευαίσθητες πλατφόρμες, προσθέτουμε ακόμη και έλεγχο ταυτότητας τρίτου παράγοντα: το πιστοποιητικό προγράμματος-πελάτη. Βασικά, πρόκειται για ένα μικρό αρχείο που οι δημοσιογράφοι αποθηκεύουν και διαμορφώνουν στους φορητούς υπολογιστές τους. Το σύστημα του δικτύου μας αρνείται την πρόσβαση σε κάθε συσκευή που δεν διαθέτει αυτό το πιστοποιητικό. Ένας άλλος αξιοσημείωτος μηχανισμός που χρησιμοποιεί το ICIJ, για να βελτιώσει την ασφάλειά του, είναι το Ciphermail. Αυτό το λογισμικό εκτελείται ανάμεσα στις πλατφόρμες μας και τα γραμματοκιβώτια των χρηστών, ώστε να διασφαλίζεται ότι κάθε μήνυμα ηλεκτρονικού ταχυδρομείου που λαμβάνουν οι δημοσιογράφοι από το ICIJ είναι κρυπτογραφημένο.

Διαχείριση μη δομημένων δεδομένων

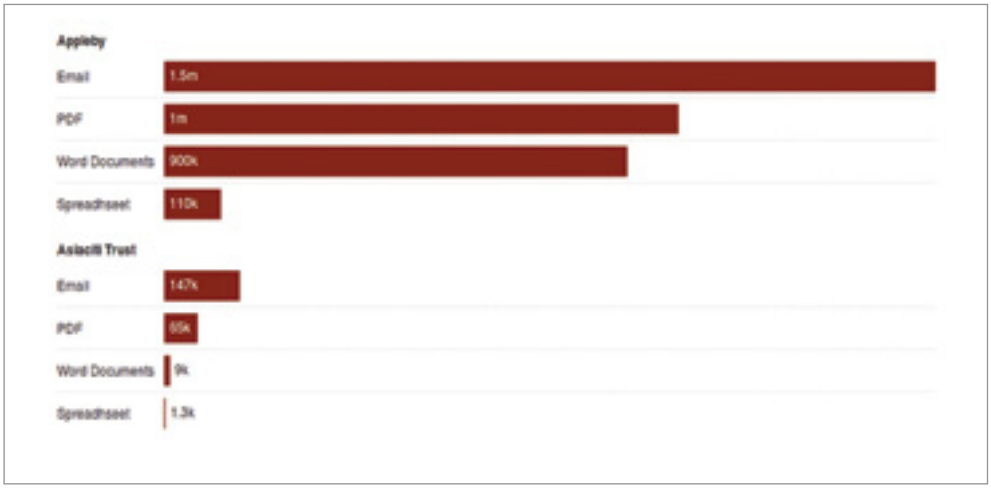
Τα *Paradise Papers* ήταν ένα απόθεμα 13,6 εκατομμυρίων εγγράφων. Μία από τις κύριες προκλήσεις κατά τη διερεύνησή τους ήταν το γεγονός ότι η διαρροή προερχόταν από διάφορες πηγές: την Appleby, την Asiaciti Trust και 19 εθνικά εταιρικά μητρώα[69]. Εάν ρίξετε μια πιο προσεκτική ματιά στα έγγραφα, θα παρατηρήσετε γρήγορα το ποικιλόμορφο περιεχόμενο και χαρακτήρα τους και την εκτεταμένη παρουσία αρχείων σε φορμάτ που δεν είναι μηχαναγνώσιμα, όπως μηνύματα ηλεκτρονικού ταχυδρομείου, έγγραφα .pdf και έγγραφα Word, που δεν μπορούν να εξεταστούν άμεσα από λογισμικό ανάλυσης δομημένων δεδομένων. Τα έγγραφα αυτά αντανακλούν τις εσωτερικές δραστηριότητες των δύο υπεράκτιων δικηγορικών εταιρειών που ερεύνησε το ICIJ.

68. https://www.gnupg.org/
69. https://www.icij.org/investigations/paradise-papers/paradise-papers-exposes-donald-trump-russia-links-and-piggy-banks-of-the-wealthiest-1-percent/, https://www.icij.org/investigations/paradise-papers/appleby-offshore-magic-circle-law-firm-record-of-compliance-failures-icij/, https://www.icij.org/investigations/paradise-papers/roll-roll-asiacitis-u-s-marketing-tour/

Οι προγραμματιστές του ICIJ δημιούργησαν ένα πολύπλοκο και ισχυρό πλαίσιο που επιτρέπει στους δημοσιογράφους να πραγματοποιούν αναζητήσεις σε αυτά τα έγγραφα. Χρησιμοποιώντας την επεκτάσιμη χωρητικότητα του υπολογιστικού νέφους («cloud computing»), τα έγγραφα αποθηκεύτηκαν σε έναν κρυπτογραφημένο δίσκο ο οποίος υποβλήθηκε σε ένα σύστημα εξαγωγής δεδομένων, μια σειρά λογισμικών που παίρνουν κείμενο από έγγραφα και το μετατρέπουν σε δεδομένα τα οποία μπορεί να εξερευνήσει η μηχανή αναζήτησης.

Τα περισσότερα αρχεία ήταν αρχεία .pdf, εικόνες, μηνύματα ηλεκτρονικού ταχυδρομείου, τιμολόγια και άλλα παρόμοια, τα οποία δεν ήταν εύκολα αναζητήσιμα. Χρησιμοποιώντας τεχνολογίες όπως το Apache Tika (για την εξαγωγή μεταδεδομένων και κειμένων), το Apache Solr (για τη δημιουργία μηχανών αναζήτησης) ή το Tesseract (για τη μετατροπή εικόνων σε κείμενο), η ομάδα δημιούργησε ένα λογισμικό ανοιχτού κώδικα που ονομάζεται Extract με μοναδική αποστολή να μετατρέψει αυτά τα έγγραφα σε περιεχόμενο αναζητήσιμο και μηχαναγνώσιμο[70]. Αυτό το εργαλείο ήταν ιδιαίτερα χρήσιμο για τη διανομή αυτών των προσβάσιμων, πλέον, δεδομένων σε έως και 30 διακομιστές.

Το ICIJ δημιούργησε, επίσης, μια διεπαφή χρήστη («user interface») που επιτρέπει στους δημοσιογράφους να εξερευνούν τις επεξεργασμένες πληροφορίες οι οποίες εξάγονται από τα «αδόμητα δεδομένα»: το συνονθύλευμα διαφορετικών τύπων εγγράφων από διάφορες πηγές. Για άλλη μία φορά, επιλέχθηκε η επαναχρησιμοποίηση ενός εργαλείου ανοιχτού κώδικα με την ονομασία Blacklight, το οποίο προσφέρει μια φιλική προς τον χρήστη διαδικτυακή πύλη, όπου οι δημοσιογράφοι μπορούν να εξετάζουν έγγραφα και να κάνουν προηγμένες αναζητήσεις (όπως η κατά προσέγγιση αντιστοίχιση συμβολοσειρών), για να εντοπίζουν πληροφορίες που είναι κρυμμένες μέσα στα έγγραφα που έχουν διαρρεύσει[71].



Εικόνα 4.2.1: Μέσα στη διαρροή.

Χρησιμοποιώντας γραφήματα, για να βρούμε μαζί κρυμμένα διαμαντάκια

Το ICIJ δημοσίευσε την πρώτη έκδοση της βάσης δεδομένων *Offshore Leaks* το 2013, χρησιμοποιώντας βάσεις δεδομένων με γραφήματα, για να επιτρέψει στους αναγνώστες να διερευνήσουν τις συνδέσεις μεταξύ αξιωματούχων και περισσότερων από 100.000 υπεράκτιων εταιρειών. Ο αριθμός αυτός έχει αυξηθεί σε πάνω από 785.000 υπεράκτιες εταιρείες τη στιγμή

70. https://github.com/ICIJ/extract/
71. https://github.com/projectblacklight/blacklight, https://en.wikipedia.org/wiki/Approximate_string_matching

που γράφονται αυτές οι γραμμές, συμπεριλαμβανομένων των διαρροών που ακολούθησαν[72], όπως τα *Panama* και *Paradise Papers*.

Το ICIJ επιχείρησε για πρώτη φορά να χρησιμοποιήσει βάσεις δεδομένων γραφημάτων με τα *Swiss Leaks*, αλλά ήταν με τα *Panama Papers* που οι βάσεις δεδομένων γραφημάτων άρχισαν να παίζουν βασικό ρόλο κατά τη φάση της έρευνας και της δημοσιοποίησης. Η διερεύνηση 11,5 εκατομμυρίων σύνθετων οικονομικών και νομικών αρχείων που αντιστοιχούν σε 2,6TB δεδομένων δεν ήταν εύκολη υπόθεση. Με τη χρήση εργαλείων για διαγράμματα δικτύου, όπως το Neo4J και το Linkurious, το ICIJ μπόρεσε να επιτρέψει στους συνεργάτες του να διερευνήσουν γρήγορα τις συνδέσεις μεταξύ προσώπων και υπεράκτιων εταιρειών.

Οι ομάδες μας που ασχολούνται με τα δεδομένα και την έρευνα εξήγαγαν πληροφορίες από τα αρχεία, τις δόμησαν και κατέστησαν τα δεδομένα αναζητήσιμα μέσω του Linkurious. Ξαφνικά, οι συνεργάτες μας βρέθηκαν σε θέση να μπορούν να αναζητήσουν τα ονόματα ατόμων δημόσιου ενδιαφέροντος και να ανακαλύψουν, για παράδειγμα, ότι ο τότε πρωθυπουργός της Ισλανδίας, Σίγκμουντουρ Γκουνλάουγκσον, ήταν μέτοχος μιας εταιρείας με την επωνυμία Wintris. Η οπτικοποίηση με αυτό το εύρημα θα μπορούσε να αποθηκευτεί και να κοινοποιηθεί σε άλλους συναδέλφους που εργάζονται για την έρευνα σε άλλα μέρη του κόσμου.

Στη συνέχεια, θα μπορούσε κανείς να επιστρέψει στην πλατφόρμα εγγράφων Blacklight, για να κάνει πιο προηγμένες αναζητήσεις και να εξερευνήσει τα αρχεία που σχετίζονται με τη Wintris. Το Blacklight εξελίχθηκε αργότερα στο Knowledge Center με τα *Paradise Papers*. Τα βασικά ευρήματα που προέκυψαν από τη διερεύνηση των δεδομένων και των εγγράφων κοινοποιήθηκαν μέσω της πλατφόρμας Global I-Hub, όπως και τα ευρήματα που προέκυψαν από το παραδοσιακό ρεπορτάζ.

Οι βάσεις δεδομένων και οι τεχνολογίες γραφημάτων τροφοδότησαν το μοντέλο ριζοσπαστικού διαμοιρασμού του ICIJ. «Είναι μαγεία!» είπαν αρκετοί συνεργάτες. Δεν απαιτούνταν δεξιότητες προγραμματισμού για τη διερεύνηση των δεδομένων. Το ICIJ πραγματοποίησε εκπαίδευση σχετικά με τη χρήση των τεχνολογιών μας για την έρευνα και την ασφάλεια και ξαφνικά περισσότεροι από 380 δημοσιογράφοι έκαναν εξόρυξη εκατομμυρίων εγγράφων, χρησιμοποιούσαν βάσεις δεδομένων με γραφήματα, έκαναν προηγμένες αναζητήσεις (συμπεριλαμβανομένων των μαζικών αναζητήσεων) και κοινοποιούσαν όχι μόνο τα ευρήματα και τα αποτελέσματα των ερευνών, αλλά και χρήσιμες συμβουλές σχετικά με τις στρατηγικές αναζήτησης.

Για το έργο *Panama Papers*, οι βάσεις δεδομένων γραφημάτων και άλλες εξειδικευμένες τεχνολογίες, όπως το Knowledge Center και η πλατφόρμα Global I-Hub, συνέδεσαν δημοσιογράφους από σχεδόν 80 χώρες που εργάζονταν σε 25 γλώσσες μέσω μιας παγκόσμιας εικονικής αίθουσας σύνταξης.

Το γεγονός ότι τα δομημένα δεδομένα που συνδέονται με τον μεγάλο αριθμό εγγράφων κοινοποιήθηκαν στο κοινό μέσω της βάσης δεδομένων *Offshore Leaks* επέτρεψε σε νέους δημοσιογράφους να εξερευνήσουν νέες πηγές και να προχωρήσουν σε νέες συνεργασίες, όπως τα έργα *Alma Mater* και *West Africa Leaks*. Επέτρεψε, επίσης, στους πολίτες και στους δημόσιους φορείς να τα χρησιμοποιήσουν ανεξάρτητα για τις δικές τους έρευνες και ελέγχους. Από τον Απρίλιο του 2019, οι κυβερνήσεις σε ολόκληρο τον κόσμο έχουν εισπράξει πάνω από 1,2 δισεκατομμύρια δολάρια σε πρόστιμα και αναδρομικούς φόρους ως αποτέλεσμα της έρευνας για τα *Panama Papers*.

Από την πρώτη δημοσίευση των *Panama Papers* το 2016, οι ομάδες των δημοσιογράφων

72. ΣτΕ: Τον Οκτώβριο 2021, όσο η προετοιμασία της ελληνικής έκδοσης του παρόντος τόμου ήταν σε εξέλιξη, το ICIJ δημοσίευσε, επίσης, την έρευνά του για τα *Pandora Papers*, τη μεγαλύτερη έρευνα στην ιστορία της δημοσιογραφίας, που βασίστηκε σε διαρροή 11,9 εκατομμυρίων εμπιστευτικών εγγράφων και στην οποία συμμετείχαν περισσότεροι από 600 δημοσιογράφοι, από 150 μέσα ενημέρωσης. Η βάση δεδομένων για τα *Pandora Papers*, μεγέθους 2,94TB, είναι μεγαλύτερη από οποιαδήποτε άλλη διαρροή έχει ερευνήσει το ICIJ.

που χρησιμοποιούν τις τεχνολογίες του ICIJ έχουν αυξηθεί και περισσότερες από 500 τέτοιες ομάδες έχουν μπορέσει να διερευνήσουν οικονομικά έγγραφα που έχουν διαρρεύσει και συνεχίζουν να δημοσιεύουν ιστορίες δημοσίου ενδιαφέροντος που συνδέονται με αυτά τα εκατομμύρια αρχεία.

Οι συγγραφείς

Η Εμίλια Ντιάζ-Στρακ είναι ερευνήτρια συντάκτρια και συντονίστρια για τη Λατινική Αμερική στο ICIJ.
Η Σεσίλ Σίλις-Γκαλιέγο είναι δημοσιογράφος στο *Forbidden Stories* και εργάστηκε στο παρελθόν ως δημοσιογράφος δεδομένων στο ICIJ.
Ο Πιέρ Ρομέρα είναι Διευθυντής Τεχνολογίας στο ICIJ.

4.3

Το Κείμενο ως Δεδομένα: Εντοπίζοντας Θέματα σε Συλλογές Κειμένων

Barbara Maseda

Περίληψη: Πώς να βρείτε ιστορίες δεδομένων σε συλλογές εγγράφων και ομιλιών.

Λέξεις-κλειδιά: δημοσιογραφία δεδομένων, αδόμητα δεδομένα, ανάλυση κειμένου, εξόρυξη κειμένου, υπολογιστική δημοσιογραφία

Εξετάζοντας την παραγωγή της δημοσιογραφίας δεδομένων τα τελευταία χρόνια, ίσως παρατηρείτε ότι οι ιστορίες που βασίζονται σε αδόμητα δεδομένα (π.χ. κείμενο) είναι πολύ πιο ασυνήθιστες από ιστορίες με δομημένα δεδομένα. Για παράδειγμα, μια ανάλυση περισσότερων από 200 υποψηφιοτήτων για τα Βραβεία Δημοσιογραφίας Δεδομένων («Data Journalism Awards») από το 2012 έως το 2016 αποκάλυψε ότι τα έργα που διαγωνίστηκαν βασίζονταν κυρίως σε γεωγραφικά και οικονομικά δεδομένα, ακολουθούμενα από άλλους συχνούς τύπους πηγών, όπως δεδομένα αισθητήρων, κοινωνικο-δημογραφικά και προσωπικά δεδομένα, μεταδεδομένα και δημοσκοπήσεις (Loosen κ.α., 2020)· με άλλα λόγια, ως επί το πλείστον δομημένα δεδομένα[73]. Καθώς, όμως, οι αίθουσες σύνταξης καλούνται να διαχειριστούν ολοένα και περισσότερες αναρτήσεις στα μέσα κοινωνικής δικτύωσης, ομιλίες, μηνύματα ηλεκτρονικού ταχυδρομείου και μακροσκελείς επίσημες εκθέσεις, οι υπολογιστικές προσεγγίσεις για την επεξεργασία και την ανάλυση αυτών των πηγών αποκτούν συνεχώς μεγαλύτερη σημασία. Μπορεί να έχετε συναντήσει τέτοιες ιστορίες: σκεφτείτε τις στατιστικές περιλήψεις των tweet του προέδρου Τραμπ, ή τις οπτικοποιήσεις των κύριων θεμάτων που θίχτηκαν σε δημόσιες τοποθετήσεις ή κατά τη διάρκεια συζητήσεων από τους υποψηφίους για τις προεδρικές εκλογές στις ΗΠΑ.

Το να αντιμετωπίσεις το κείμενο ως δεδομένα δεν είναι μικρό επίτευγμα. Τα έγγραφα τείνουν να βρίσκονται στα πιο διαφορετικά φορμάτ, με τις πιο διαφορετικές διατάξεις και περιεχόμενα, γεγονός που καθιστά δύσχρηστες τις μονοδιάστατες λύσεις ή τις προσπάθειες αναπαραγωγής μιας έρευνας με ένα διαφορετικό σύνολο εγγράφων. Ο καθαρισμός, η προετοιμασία και η ανάλυση των δεδομένων μπορεί να διαφέρουν σημαντικά από τη μία συλλογή εγγράφων στην άλλη, και ορισμένα βήματα θα απαιτήσουν περαιτέρω ανθρώπινο έλεγχο, προτού μπορέσουμε να προβούμε σε ισχυρισμούς που χρήζουν ειδησεογραφικού ενδιαφέροντος ή να παρουσιάσουμε τα ευρήματα με τρόπο που να αποκαλύπτει κάτι ουσιαστικό όχι μόνο για τους ερευνητές αλλά και για το ευρύτερο κοινό[74].

73. Για περισσότερα σχετικά με τα Βραβεία Δημοσιογραφίας Δεδομένων, βλ. το κεφάλαιο της Λούζεν σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 9.2, «Χρυσά Πρότυπα Με Γνώμονα τα Δεδομένα: Τι Αποτιμά ο Κλάδος ως Δημοσιογραφία Δεδομένων που αξίζει Βραβείο και Πώς η Δημοσιογραφία Δεδομένων Εξελίσσεται μαζί με τη Δεδομενοποίηση της Κοινωνίας»).

74. Ο καθαρισμός και η προετοιμασία δεδομένων μπορεί να περιλαμβάνει ένα ή περισσότερα από τα ακόλουθα βήματα: κατάτμηση του κειμένου σε μονάδες ή «token» (μια διαδικασία γνωστή ως «tokenization»)· «ομαδοποίηση» λέξεων που μοιράζονται μια κοινή οικογένεια ή ρίζα (αποκοπή καταλήξεων και λημματοποίηση)· εξάλειψη των περιττών στοιχείων, όπως οι κοινές λέξεις και τα σημεία στίξης· αλλαγή της πτώσης του κειμένου· επιλογή της εστίασης στις λέξεις και αγνόηση της σειράς τους (ένα μοντέλο που ονομάζεται «bag of words»)· και μετατροπή του κειμένου σε διανυσματική αναπαράσταση.

Σε αυτό το κεφάλαιο, εξετάζω πέντε τρόπους, με τους οποίους οι δημοσιογράφοι μπορούν να χρησιμοποιήσουν την ανάλυση κειμένου, για να αφηγηθούν ιστορίες –που επεξηγούνται με αναφορά σε διάφορα παραδείγματα έργων δημοσιογραφίας δεδομένων.

Έκταση: Πόσο έγραψαν ή μίλησαν

Η καταμέτρηση προτάσεων ή λέξεων είναι η απλούστερη ποσοτική προσέγγιση όσον αφορά στα έγγραφα. Από υπολογιστική άποψη, αυτή είναι μια εργασία που υπάρχει εδώ και πολύ καιρό και μπορεί εύκολα να εκτελεστεί από τους περισσότερους επεξεργαστές κειμένου. Εάν είστε φοιτητής ή ρεπόρτερ που χρειάστηκε έστω και μία φορά να υποβάλετε κάποια εργασία ή να παραδώσετε κάποια ανάθεση με όριο λέξεων, δεν θα χρειαστείτε κάποια ειδική εκπαίδευση στα δεδομένα για να το καταλάβετε αυτό.

Το πρόβλημα με την καταμέτρηση των λέξεων έγκειται στην ερμηνεία των αποτελεσμάτων σε σχέση με ένα κατώτατο όριο. Τέτοιες μετρήσεις δεν είναι τόσο ευρέως γνωστές όσο η θερμοκρασία ή η ταχύτητα και, επομένως, μπορεί να μην είναι εξίσου απλό να βγει κάποιο νόημα από το γεγονός ότι μια ομιλία έχει έκταση 2.000 λέξεων. Στην πράξη, πολλές φορές η μόνη επιλογή είναι να δημιουργήσουμε εμείς οι ίδιοι αυτά τα κατώτατα όρια ή αναφορές για σύγκριση, κάτι που μπορεί να μεταφράζεται σε επιπλέον εργατοώρες.

Σε ορισμένες περιπτώσεις, είναι εφικτό να βρεθεί ένα πλαίσιο στην ιστορία του γεγονότος ή του ομιλητή που εξετάζετε. Για παράδειγμα, για την κάλυψη της ετήσιας ομιλίας του προέδρου των ΗΠΑ για την Κατάσταση της Ένωσης το 2016, το Vox υπολόγισε τη διάρκεια όλου του αρχείου των ιστορικών ομιλιών, για να διαπιστώσει ότι «ο πρόεδρος Ομπάμα ήταν από τους πιο φλύαρους ομιλητές που έχουν εκφωνήσει λόγο ποτέ για την Κατάσταση της Ένωσης» (Chang, 2016).

Σε εκδηλώσεις που περιλαμβάνουν περισσότερους από έναν ομιλητές, είναι δυνατόν να διερευνηθεί πόσο πολύ και πότε μιλάει ο καθένας σε σχέση με τον συνολικό αριθμό των λέξεων που εκφωνήθηκαν. Δείτε, για παράδειγμα, την Εικόνα 4.3.1.

Εικόνα 4.3.1: Οπτικοποίηση της αντιλογίας των Δημοκρατικών (2015). Πηγή: *The Washington Post*.

Αναφορές: Ποιος είπε τι, πότε και πόσες φορές

Η καταμέτρηση της συχνότητας χρήσης ενός όρου ή μιας έννοιας στην ομιλία ή στον γραπτό λόγο είναι μια άλλη απλή εργασία που παρέχει χρήσιμη στατιστική επισκόπηση των δεδομένων μας. Για να γίνει αυτό, είναι σημαντικό να βεβαιωθούμε ότι επιλέγουμε να μετρήσουμε τα πιο κατάλληλα στοιχεία.

Ανάλογα με τα ερωτήματα που θέλετε να θέσετε στα δεδομένα, μπορείτε να μετρήσετε τις επαναλήψεις κάθε λέξης ή μιας σειράς λέξεων που έχουν κοινή ρίζα, χρησιμοποιώντας λειτουργίες κανονικοποίησης όπως «αποκοπή καταλήξεων» («stemming») ή «λημματοποίηση» («lemmatization»)[75]. Μια άλλη προσέγγιση είναι η εστίαση στους πιο σχετικούς όρους σε κάθε έγγραφο με τη χρήση ενός σταθμισμένου μέτρου που καλείται «συχνότητα όρων/αντίστροφη συχνότητα εγγράφων» («Term Frequency-Inverse Document Frequency», TF-IDF)[76]. Ακολουθούν μερικά παραδείγματα.

Συχνοί όροι και θέματα. Για την κάλυψη των δημοτικών εκλογών του Λονδίνου το 2016, η εφημερίδα *Guardian* ανέλυσε πόσες φορές οι δύο υποψήφιοι είχαν ασχοληθεί με διάφορα θέματα της προεκλογικής εκστρατείας (π.χ. εγκληματικότητα, ρύπανση, στέγαση και μεταφορές) στο βρετανικό κοινοβούλιο κατά τα έξι χρόνια που προηγήθηκαν της αναμέτρησης (Barr, 2016). Τα προς ανάλυση θέματα μπορούν να αποφασιστούν εκ των προτέρων, όπως στην προκείμενη περίπτωση, και να διερευνηθούν μέσω ενός αριθμού σχετικών λέξεων-κλειδιών (ή ομάδων λέξεων-κλειδιών που συνδέονται με ένα θέμα) σε συγκρίσιμες ή ανάλογες συλλογές κειμένων. Οι όροι αναζήτησης μπορούν, επίσης, να είναι του ανάλογου ενδιαφέροντος, και όχι απαραίτητα οι ίδιοι. Πάρτε, για παράδειγμα, την ανάλυση του *FiveThirtyEight* για τον τρόπο με τον οποίο τα ίδια μέσα ενημέρωσης κάλυψαν τρεις διαφορετικούς τυφώνες το 2017, τους Χάρβεϊ, Ίρμα και Μαρία (Mehta, 2017). Μια άλλη προσέγγιση είναι απλά η εξέταση των πιο κοινών λέξεων σε ένα κείμενο ως στρατηγική ανίχνευσης θεμάτων.

Ομιλία σε βάθος χρόνου. Εξετάζοντας μια ομιλία σε βάθος χρόνου μπορούμε, επίσης, να επισημάνουμε θέματα που δεν έχουν αναφερθεί ποτέ πριν ή που δεν έχουν συζητηθεί για πολύ καιρό. Αυτή ήταν, για παράδειγμα, η προσέγγιση που επέλεξε η *Washington Post* για την κάλυψη της ομιλίας για την Κατάσταση της Ένωσης το 2018, σε ένα άρθρο που επεσήμαινε ποιος πρόεδρος χρησιμοποίησε ποιες λέξεις πρώτος στην ιστορία αυτής της εκδήλωσης (Fischer-Baum κ.α., 2018). Το γεγονός ότι οι αναγνώστες μπορούν πολύ γρήγορα να μάθουν ότι ο Τραμπ ήταν ο πρώτος πρόεδρος ιστορικά που ανέφερε τα καταστήματα Walmart (το 2017) ή τους «τζαμπατζήδες» (το 2019), χωρίς να χρειάζεται να διαβάσουν εκατοντάδες σελίδες ομιλιών, δείχνει πόσο αποτελεσματικές μπορούν να είναι οι περιλήψεις και οι οπτικοποιήσεις δεδομένων κειμένου.

Παραλείψεις. Ένας χαμηλός αριθμός ή η απουσία αναφορών μπορεί, επίσης, να είναι άξια λόγου. Αυτές οι παραλείψεις μπορούν να αναλυθούν σε βάθος χρόνου, όπως και με βάση την προσδοκία ότι ένα πρόσωπο ή ένας οργανισμός οφείλει να αναφέρει κάτι σε ένα δεδομένο πλαίσιο. Κατά τη διάρκεια της προεδρικής εκστρατείας του 2016 στις Ηνωμένες Πολιτείες, το *FiveThirtyEight* δημοσίευσε ότι ο υποψήφιος Ντόναλντ Τραμπ σταμάτησε να τουιτάρει για τις δημοσκοπήσεις, όταν διαπίστωσε ότι λέξεις-κλειδιά σχετικές με τις δημοσκοπήσεις αναφέρο-

75. Η αποκοπή καταλήξεων και η λημματοποίηση είναι λειτουργίες για την αναγωγή των παράγωγων λέξεων στη μορφή της ρίζας τους, έτσι ώστε οι εμφανίσεις των λέξεων «reporter», «reporting» και «reported» να μπορούν όλες να υπολογιστούν κάτω από την ομπρέλα της λέξης «report». Διαφέρουν ως προς τον τρόπο που ο αλγόριθμος καθορίζει τη ρίζα της λέξης. Σε αντίθεση με την αποκοπή καταλήξεων, η λημματοποίηση απογυμνώνει τις λέξεις από τα επιθήματά τους, χωρίς να λαμβάνει υπόψη της το μέρος του λόγου που αποτελούν.
76. Το TF-IDF είναι ένα μέτρο που χρησιμοποιείται από αλγορίθμους για την κατανόηση του βάρους μιας λέξης σε μια συλλογή. TF-IDF Βάρος (w, d) = TermFreq (w, d) · log (N / DocFreq (w)), όπου TermFreq (w, d) είναι η συχνότητα της λέξης στο έγγραφο (d), N είναι ο αριθμός όλων των εγγράφων και το DocFreq (w) είναι ο αριθμός των εγγράφων που περιέχουν τη λέξη w (Feldman & Sanger, 2007).

νται συγκριτικά λίγες φορές στις αναρτήσεις του (Mehta & Enten, 2016). Τέτοιες παραλείψεις μπορούν να εντοπιστούν με την παρακολούθηση του ίδιου ομιλητή σε βάθος χρόνου, όπως σε αυτή την περίπτωση, στην οποία, μήνες πριν, το *FiveThirtyEight* είχε ανακαλύψει ότι ο Τραμπ τουίταρε συχνά για δημοσκοπήσεις που τον εμφάνιζαν ως νικητή (Bialik & Enten, 2015). Αυτό αποτελεί επίσης ένα καλό παράδειγμα για το πώς τα δημοσιεύματα που βασίζονται στην ανάλυση κειμένου μπορούν αργότερα να αποτελέσουν το πλαίσιο για ένα συμπληρωματικό άρθρο, ως ένας τρόπος αντιμετώπισης του προαναφερθέντος προβλήματος στην πλαισίωση των στατιστικών κειμένου. Η απουσία αναφοράς σε ένα θέμα μπορεί, επίσης, να μετρηθεί με βάση την προσδοκία ότι ένα άτομο ή ένας οργανισμός θα το αναφέρει σε ένα δεδομένο πλαίσιο.

Άνθρωποι, μέρη, ουσιαστικά, ρήματα. Τα εργαλεία επεξεργασίας φυσικής γλώσσας («Natural Language Processing», NLP) επιτρέπουν την εξαγωγή κύριων ονομάτων, ονομάτων τοποθεσιών, εταιρειών και άλλων στοιχείων, μέσω μιας εργασίας που ονομάζεται αναγνώριση ονομαστικών οντοτήτων («Named Entity Recognition», NER)· καθώς και την αναγνώριση ουσιαστικών, επιθέτων και άλλων τύπων λέξεων, μέσω μιας εργασίας που ονομάζεται επισημείωση για μέρη του λόγου («Part-of-Speech tagging», POS). Στο άρθρο της *Washington Post* που αναφέρθηκε προηγουμένως, η απεικόνιση περιλαμβάνει φίλτρα που εστιάζουν σε εταιρείες, θρησκευτικούς όρους και ρήματα.

Συγκρίσεις

Ο προσδιορισμός της ομοιότητας δύο ή περισσότερων εγγράφων μπορεί να αποτελέσει το σημείο εκκίνησης για διάφορα είδη ιστοριών. Μπορούμε να χρησιμοποιήσουμε την κατά προσέγγιση αντιστοίχιση προτάσεων (γνωστή και ως «ασαφής αντιστοίχιση» –«fuzzy matching» στα αγγλικά), για να ξεσκεπάσουμε τη λογοκλοπή, να αποκαλύψουμε την κοινή νοοτροπία δημόσιων προσώπων ή να περιγράψουμε πώς έχει αλλάξει ένα νομοθετικό κείμενο. Το 2012, η *ProPublica* το έκανε αυτό για να παρακολουθήσει τις αλλαγές σε μηνύματα ηλεκτρονικού ταχυδρομείου που στάλθηκαν σε ψηφοφόρους από εκστρατείες, παρουσιάζοντας διαδοχικές εκδοχές των ίδιων μηνυμάτων πλάι πλάι και απεικονίζοντας τις διαγραφές, τις προσθήκες και το αμετάβλητο κείμενο (Larson & Shaw, 2012).

Ταξινόμηση

Το κείμενο μπορεί να ταξινομηθεί σε κατηγορίες σύμφωνα με ορισμένα προκαθορισμένα χαρακτηριστικά, χρησιμοποιώντας αλγορίθμους μηχανικής μάθησης («machine learning»). Σε γενικές γραμμές, η διαδικασία συνίσταται στην εκπαίδευση ενός μοντέλου για την ταξινόμηση εγγραφών με βάση ένα δεδομένο χαρακτηριστικό και, στη συνέχεια, στη χρήση του για την κατηγοριοποίηση νέων δεδομένων.

Για παράδειγμα, το 2015, η εφημερίδα *Los Angeles Times* ανέλυσε περισσότερες από 400.000 αστυνομικές αναφορές που ελήφθησαν μέσω αιτήματος για πρόσβαση σε δημόσια αρχεία και αποκάλυψε ότι περίπου 14.000 σοβαρές επιθέσεις είχαν κατηγοριοποιηθεί λανθασμένα από το αστυνομικό τμήμα του Λος Άντζελες ως ασήμαντα παραπτώματα (Poston κ.α., 2015). Αντί να χρησιμοποιήσουν τη MySQL, για να αναζητήσουν λέξεις-κλειδιά (π.χ. μαχαίρωμα, μαχαίρι) οι οποίες θα παρέπεμπαν σε βίαια αδικήματα –όπως είχαν κάνει σε προηγούμενη έρευνα που κάλυπτε μικρότερο όγκο δεδομένων–, οι δημοσιογράφοι χρησιμοποίησαν ταξινομητές μηχανικής μάθησης (SVM και MaxEnt), για να κατηγοριοποιήσουν εκ νέου και να επανεξετάσουν δεδομένα οκτώ ετών στον μισό χρόνο σε σχέση με την πρώτη έρευνα, η οποία κάλυπτε μόνο ένα έτος (Poston & Rubin, 2014). Αυτό το παράδειγμα δείχνει πώς οι προσεγγίσεις μηχανικής μάθησης μπορούν, επίσης, να εξοικονομήσουν χρόνο και να πολλαπλασιάσουν την ερευνητική μας δύναμη.

Συναίσθημα

Πολλοί δημοσιογράφοι θα αναγνώριζαν την αξία της ταξινόμησης προτάσεων ή εγγράφων ως θετικών, αρνητικών ή ουδέτερων (υπάρχουν και άλλες κλίμακες διαβάθμισης), ανάλογα με τη στάση του ομιλητή απέναντι στο εκάστοτε θέμα. Οι διάφορες εφαρμογές μπορεί να περιλαμβάνουν την ανάλυση ενός θέματος, ενός hashtag ή των αναρτήσεων ενός χρήστη του Twitter για την αξιολόγηση του συναισθήματος γύρω από ένα ζήτημα, καθώς και την εκτέλεση παρόμοιων αξιολογήσεων σε δελτία Τύπου ή σχόλια χρηστών σε έναν ιστότοπο. Πάρτε, για παράδειγμα, τη σύγκριση που έκανε ο *Economist* για τον τόνο των ομιλιών στα συνέδρια των κομμάτων της Χίλαρι Κλίντον και του Ντόναλντ Τραμπ (Economist, 2016). Αναλύοντας την πόλωση στις λέξεις που χρησιμοποίησαν αυτοί και προηγούμενοι υποψήφιοι, μπόρεσε να καταδείξει ότι ο Τραμπ «εκφώνησε την πιο αρνητική ομιλία στην πρόσφατη ιστορία» και η Κλίντον «μία από τις πιο ισορροπημένες ομιλίες των τελευταίων τεσσάρων δεκαετιών».

Πώς γίνεσαι δημοσιογράφος που κάνει εξόρυξη κειμένου

Η χρήση έτοιμου λογισμικού εξόρυξης κειμένου είναι ένα καλό σημείο εκκίνησης, για να εξοικειωθείτε με τις βασικές λειτουργίες ανάλυσης κειμένου και τα αποτελέσματά τους (μέτρηση λέξεων, εξαγωγή οντοτήτων, συσχετισμούς μεταξύ εγγράφων κλπ). Οι πλατφόρμες που έχουν σχεδιαστεί για δημοσιογράφους –όπως το DocumentCloud και το Overview– περιλαμβάνουν ορισμένα από αυτά τα χαρακτηριστικά[77]. Το API[78] του Google Cloud για NLP μπορεί να διεκπεραιώσει διάφορες εργασίες, συμπεριλαμβανομένης της ανάλυσης συναισθήματος, της ανάλυσης οντοτήτων, της ταξινόμησης περιεχομένου και της συντακτικής ανάλυσης[79].

Για όσους ενδιαφέρονται να μάθουν περισσότερα για την εξόρυξη κειμένου, υπάρχουν δωρεάν εργαλεία ανοιχτού κώδικα που επιτρέπουν πιο εξατομικευμένες αναλύσεις, συμπεριλαμβανομένων βιβλιοθηκών σε Python (NLTK, spaCy, gensim, textblob, scikit-learn) και R (tm, tidytext και πολλές άλλες), οι οποίες μπορεί να είναι πιο βολικές για δημοσιογράφους που είναι ήδη εξοικειωμένοι με αυτές τις γλώσσες. Η καλή γνώση των «regular expressions»[80] και των εργαλείων και τεχνικών που απαιτούνται για τη συλλογή κειμένων (scraping, αναζήτηση δεδομένων μέσω κλήσης API, αιτήματα FOIA[81]) και την επεξεργασία τους (οπτική αναγνώριση χαρακτήρων[82], μετατροπή μορφότυπων αρχείων κλπ) είναι επίσης απαραίτητη[83]. Και, φυσικά, μπορεί να είναι χρήσιμο να αποκτήσετε μια αντίληψη της θεωρίας και των αρχών που διέπουν την εργασία με δεδομένα κειμένου, συμπεριλαμβανομένης της ανάκτησης πληροφοριών, των σχετικών μοντέλων και αλγορίθμων και της οπτικοποίησης δεδομένων κειμένου[84].

Συμπεράσματα

Η δυνατότητα αποκάλυψης νέων πληροφοριών στο κοινό με και για τα έγγραφα και ο πολλαπλασιασμός των ικανοτήτων μας να αναλύουμε μεγάλα κείμενα που κανονικά θα χρειαζόταν

77. https://www.documentcloud.org/, https://www.overviewdocs.com/
78. «Application Programming Interface» (διεπαφή προγραμματισμού εφαρμογών)
79. https://cloud.google.com/natural-language/
80. ΣτΕ: Για έναν πρώτο οδηγό στην εκμάθηση και στη χρήση «regex», όπως συχνά λένε οι δημοσιογράφοι δεδομένων, μπορείτε να δείτε το regular-expressions.info.
81. ΣτΕ: «Freedom of Information Act» (νόμος για την ελευθερία της πληροφόρησης)
82. ΣτΕ: «Optical Character Recognition» (OCR)
83. http://regex.bastardsbook.com/
84. Για περαιτέρω μελέτη, βλ. το *Speech and Language Processing* των Ντάνιελ Ζουράφσκι και Τζέιμς Μάρτιν και το *The Text Mining Handbook* των Ρόνεν Φέλντμαν και Τζέιμς Σάνγκερ. Υπάρχουν, επίσης, πολυάριθμα δωρεάν διαδικτυακά μαθήματα για αυτά και συναφή θέματα.

μήνες ή χρόνια για να τα διαβάσουμε, είναι σοβαροί λόγοι για να εξετάσουμε σοβαρά την ανάπτυξη της ανάλυσης κειμένου ως χρήσιμου εργαλείου για τη δημοσιογραφία. Υπάρχουν ακόμη πολλές προκλήσεις, από ζητήματα ασάφειας –οι υπολογιστές μπορεί να δυσκολεύονται περισσότερο να «κατανοήσουν» το πλαίσιο της γλώσσας από ό,τι εμείς οι άνθρωποι– μέχρι ειδικά γλωσσικά προβλήματα που μπορεί να είναι ευκολότερο να επιλυθούν στα αγγλικά από ό,τι στα γερμανικά, ή που απλώς έχουν αντιμετωπιστεί περισσότερο σε ορισμένες γλώσσες από ό,τι σε άλλες. Η δουλειά των δημοσιογράφων μπορεί να συμβάλλει στην προώθηση αυτού του τομέα. Πολλά δημοσιογραφικά έργα θα μπορούσαν να θεωρηθούν τρόποι επέκτασης του αριθμού των διαθέσιμων επισημειωμένων συνόλων δεδομένων («annotated datasets») και εντοπισμού προκλήσεων, καθώς και νέες ιδέες για δημιουργία εφαρμογών. Κρίνοντας από τον αυξανόμενο αριθμό πρόσφατων ιστοριών που παράγονται με αυτή την προσέγγιση, η εξόρυξη κειμένου φαίνεται να είναι ένας πολλά υποσχόμενος και συναρπαστικός τομέας ανάπτυξης της δημοσιογραφίας δεδομένων.

Παραπομπές

Barr, C. (2016) «London mayor: Commons speeches reveal candidates' differing issue focus», *The Guardian*, 3 Μαΐου. Διαθέσιμο σε: https://www.theguardian.com/politics/datablog/2016/may/03/london-mayor-data-indicates-candidates-differing-focus-on-issues (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Bialik, C., & Enten, H. (2015) «Shocker: Trump tweets the polls that make him look most like a winner», *FiveThirtyEight*, 15 Δεκεμβρίου. Διαθέσιμο σε: https://fivethirtyeight.com/features/shocker-trump-tweets-the-polls-that-make-him-look-most-like-a-winner/ (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Chang, A. (2016, January 11) «President Obama is among the wordiest State of the Union speakers ever», *Vox*, 11 Ιανουαρίου. Διαθέσιμο σε: https://www.vox.com/2016/1/11/10736570/obama-wordy-state-of-the-union (Προσπελάστηκε 26 Νοεμβρίου 2021).

Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Κέιμπριτζ: Cambridge University Press.

Fischer-Baum, R., Mellnik, T., & Schaul, K. (2018) «The words Trump used in his State of the Union that had never been used before», *The Washington Post,* 30 Ιανουαρίου. Διαθέσιμο σε: https://www.washingtonpost.com/graphics/politics/presidential-lexicon-state-of-the-union/ (Προσπελάστηκε: 26 Νοεμβρίου 2021).

The Economist (2016) «How Clinton's and Trump's convention speeches compared to those of their predecessors», 29 Ιουλίου. Διαθέσιμο σε: https://www.economist.com/graphic-detail/2016/07/29/how-clintons-and-trumps-convention-speeches-compared-to-those-of-their-predecessors (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Jurafsky, D., & Martin, J. H. (2008). Speech and Language Processing. Νέα Υόρκη: Pearson.

Larson, J., & Shaw, A. (2012) «Message machine: Reverse engineering the 2012 campaign», *ProPublica*, 17 Ιουλίου. Διαθέσιμο σε: http://projects.propublica.org/emails/ (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Loosen, W., Reimer, J., & De Silva-Schmidt, F. (2020) «Data-driven reporting: An on-going (r) evolution? An analysis of projects nominated for the Data Journalism Awards 2013–2016», *Journalism*, 21(9), σσ. 1246–1263. DOI: https://doi.org/10.1177/1464884917735691

Mehta, D. (2017) «The media really has neglected Puerto Rico», *FiveThirtyEight*, 28 Σεπτεμβρίου. Διαθέσιμο σε: https://fivethirtyeight.com/features/the-media-really-has-neglected-puerto-rico/ (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Mehta, D., & Enten, Ha. (2016) «Trump isn't tweeting about the polls anymore», *FiveThirtyEight*, 19 Αυγούστου. Διαθέσιμο σε: https://fivethirtyeight.com/features/trump-isnt-tweeting-about-the-polls-anymore/ (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Poston, B., & Rubin, J. (2014) «Times Investigation: LAPD misclassified nearly 1,200 violent crimes as minor offenses», *Los Angeles Times*, 10 Αυγούστου. Διαθέσιμο σε: https://www.latimes.com/local/la-me-crimestats-lapd-20140810-story.html (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Poston, B., Rubin, J., & Pesce, A. (2015). LAPD underreported serious assaults, skewing crime stats for 8 years, *Los Angeles Times*, 15 Οκτωβρίου. Διαθέσιμο σε: https://www.latimes.com/local/cityhall/la-me-crime-stats-20151015-story.html (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Η συγγραφέας

Η Μπάρμπαρα Μασέντα είναι η ιδρύτρια και συντάκτρια δεδομένων στο *Proyecto Inventario* στην Κούβα.

4.4
Προγραμματισμός με Δεδομένα στην Αίθουσα Σύνταξης

Basile Simon

Περίληψη: Οι αίθουσες σύνταξης παρουσιάζουν μοναδικές προκλήσεις για τους προγραμματιστές και τους τεχνικά καταρτισμένους δημοσιογράφους.

Λέξεις-κλειδιά: υπολογιστική δημοσιογραφία, προγραμματισμός, καθαρισμός δεδομένων, βάσεις δεδομένων, οπτικοποίηση δεδομένων

Αναπόφευκτα, υπάρχει ένα σημείο όπου τα δεδομένα και ο κώδικας φτάνουν να συμπορεύονται. Ίσως όταν το Google Sheets αρχίζει να «κολλάει» εξαιτίας του μεγέθους ενός συνόλου δεδομένων, όταν οι φόρμουλες στο Excel γίνονται πολύ αινιγματικές ή όταν καθίσταται αδύνατο να βγει νόημα από δεδομένα που εκτείνονται σε εκατοντάδες σειρές. Ο προγραμματισμός μπορεί να κάνει την εργασία με δεδομένα απλούστερη, πιο κομψή, λιγότερο επαναλαμβανόμενη και πιο αναπαραγώγιμη. Αυτό δεν σημαίνει ότι θα εγκαταλείψουμε τα υπολογιστικά φύλλα, αλλά μάλλον ότι αυτά θα καταλήξουν να αποτελούν μόνο ένα από τα διάφορα διαθέσιμα εργαλεία. Οι δημοσιογράφοι δεδομένων συχνά περνούν από τη μία τεχνική στην άλλη ανάλογα με τις ανάγκες τους: κάνουν scraping σε Python Notebook, ρίχνουν το αποτέλεσμα σε ένα υπολογιστικό φύλλο και το αντιγράφουν στο Refine για να το καθαρίσουν, πριν να το επικολλήσουν ξανά.

Διαφορετικοί άνθρωποι μαθαίνουν διαφορετικές γλώσσες και τεχνικές προγραμματισμού· διαφορετικές αίθουσες σύνταξης παράγουν το έργο τους σε διαφορετικές γλώσσες, επίσης. Αυτό προέρχεται εν μέρει από την επιλογή της εργαλειοθήκης που κάνει κάθε οργανισμός –δηλαδή, το σύνολο των τεχνολογιών που χρησιμοποιούνται εσωτερικά: για παράδειγμα, το μεγαλύτερο μέρος των δεδομένων, των οπτικοποιήσεων και της ανάπτυξης λογισμικού στους *Times* του Λονδίνου γίνεται σε R, JavaScript και React. Στην άλλη πλευρά του Ατλαντικού, η *ProPublica* χρησιμοποιεί Ruby για πολλές από τις διαδικτυακές εφαρμογές της. Ενώ τα εργαλεία είναι ατομική επιλογή, οι πρακτικές και οι κουλτούρες των ειδησεογραφικών οργανισμών μπορούν να επηρεάσουν σημαντικά τις επιλογές αυτές. Για παράδειγμα, το BBC μεταφέρει σταδιακά τη ροή εργασιών οπτικοποίησης δεδομένων στην R (BBC Data Journalism team, άνευ ημερομηνίας). Ο *Economist* έχει αλλάξει το παγκοσμίως γνωστό Big Mac Index από υπολογισμούς βασισμένους στο Excel σε R και σε dashboard που δημιουργήθηκε με React/d3.js (González κ.α., 2018). Υπάρχουν πολλές επιλογές και όχι μία και μοναδική, σωστή απάντηση. Τα καλά νέα για εκείνους που ξεκινούν τώρα είναι ότι πολλές βασικές έννοιες ισχύουν για όλες τις γλώσσες προγραμματισμού. Μόλις καταλάβετε πώς να αποθηκεύετε δεδομένα σε μια λίστα (όπως θα κάνατε σε μια σειρά ή στήλη ενός υπολογιστικού φύλλου) και πώς να εκτελείτε διάφορες λειτουργίες στην Python, το να κάνετε το ίδιο πράγμα σε JavaScript, R ή Ruby είναι απλώς ζήτημα εκμάθησης της σύνταξης.

Για τους σκοπούς αυτού του κεφαλαίου, μπορούμε να θεωρήσουμε ότι ο προγραμματισμός της δημοσιογραφίας δεδομένων υποδιαιρείται σε τρεις βασικούς τομείς: εργασία με δεδομένα –συμπεριλαμβανομένων του scraping, του καθαρισμού, των στατιστικών αναλύσεων (εργασίες που θα μπορούσατε να εκτελέσετε σε ένα υπολογιστικό φύλλο)· εργασία backend –ο εσωτερι-

κός κόσμος των βάσεων δεδομένων, των διακομιστών και των API[85]· και εργασία frontend –τα περισσότερα από όσα συμβαίνουν σε ένα πρόγραμμα περιήγησης στο Διαδίκτυο, συμπεριλαμβανομένων των διαδραστικών οπτικοποιήσεων δεδομένων. Αυτό το κεφάλαιο διερευνά πώς αυτοί οι διαφορετικοί τομείς εργασίας διαμορφώνονται από διάφορους περιορισμούς που αντιμετωπίζουν συνήθως οι δημοσιογράφοι δεδομένων κατά την εργασία με κώδικα στις αίθουσες σύνταξης, όπως (α) ο χρόνος εκμάθησης, (β) το να δουλεύεις με προθεσμίες και (γ) η αναθεώρηση του κώδικα.

Ώρα για μάθημα

Ένα από τα υπέροχα χαρακτηριστικά που ενώνει την κοινότητα της δημοσιογραφίας δεδομένων είναι η όρεξη για μάθηση. Είτε είστε δημοσιογράφος που επιθυμεί να μάθει τα κατατόπια, είτε φοιτητής που θέλει να βρει δουλειά σε αυτόν τον τομέα, είτε επαγγελματίας, υπάρχουν πολλά να μάθετε. Καθώς η τεχνολογία εξελίσσεται πολύ γρήγορα και καθώς ορισμένα εργαλεία βγαίνουν εκτός μόδας, ενώ άλλα δημιουργούνται από ταλαντούχους και γενναιόδωρους ανθρώπους, υπάρχουν πάντα νέα πράγματα που μπορούν να γίνουν και νέα πράγματα που μπορούμε να μάθουμε. Υπάρχουν συχνά διαδοχικές επαναλήψεις και εκδόσεις εργαλείων για μια συγκεκριμένη εργασία (π.χ. βιβλιοθήκες για τη λήψη δεδομένων από το API του Twitter). Εργαλεία συχνά δημιουργούνται και επεκτείνονται με βάση προηγούμενά τους (π.χ. επεκτάσεις και πρόσθετα για τη βιβλιοθήκη οπτικοποίησης δεδομένων d3.js). Ο προγραμματισμός στη δημοσιογραφία δεδομένων είναι, επομένως, μια συνεχής διαδικασία εκμάθησης που απαιτεί χρόνο και ενέργεια, πέρα από επένδυση χρόνου για την αρχική εκπαίδευση.

Ένα πρόβλημα που προκύπτει με την εκμάθηση του προγραμματισμού είναι η αρχική μείωση της ταχύτητας και της αποδοτικότητας που συνεπάγεται η ενασχόληση με άγνωστες έννοιες. Ένα ταχύρρυθμο και εντατικό πρόγραμμα εκπαίδευσης στον προγραμματισμό μπορεί να σας προσφέρει άρτια κατάρτιση μέσα σε λίγες εβδομάδες, αν και ενδέχεται να κοστίζει αρκετά. Τα εργαστήρια σε συνέδρια είναι μικρότερης διάρκειας και φθηνότερα και απευθύνονται τόσο σε αρχάριους όσο και σε προχωρημένους χρήστες. Το να έχετε χρόνο να μαθαίνετε εν ώρα εργασίας, ως μέρος της δουλειάς σας, είναι μια αναγκαιότητα. Εκεί θα αντιμετωπίσετε πραγματικά, πρακτικά προβλήματα και, αν είστε τυχεροί, θα έχετε συναδέλφους να σας βοηθήσουν. Υπάρχει ένα κόλπο για να βρίσκετε λύσεις στα προβλήματά σας: ρωτώντας ξανά και ξανά και αναπτύσσοντας τη διαίσθησή σας για το τι δημιουργεί ένα συγκεκριμένο πρόβλημα.

Αυτή η επένδυση σε χρόνο και πόρους μπορεί να αποδώσει: ο προγραμματισμός ανοίγει πολλές νέες δυνατότητες και προσφέρει πολλές ανταμοιβές. Ένα ζήτημα που διατηρείται σε όλα τα στάδια της εμπειρίας είναι το γεγονός ότι είναι δύσκολο να εκτιμηθεί πόσο χρόνο θα πάρει μια εργασία. Αυτό αποτελεί πρόκληση, διότι η εργασία στην αίθουσα σύνταξης βασίζεται σε προθεσμίες.

Δουλεύοντας με προθεσμίες

Η έγκαιρη παράδοση αποτελεί ουσιαστικό μέρος της δημοσιογραφικής δουλειάς. Ο προγραμματισμός, όπως και το ρεπορτάζ, μπορεί να είναι απρόβλεπτος. Ανεξάρτητα από το επίπεδο της εμπειρίας σας, οι καθυστερήσεις μπορούν να συμβούν –και πάντοτε συμβαίνουν.

Μία πρόκληση για τους αρχάριους είναι η επιβράδυνση που προκαλείται από την εκμάθηση ενός νέου τρόπου εργασίας. Όταν ξεκινάτε να κάνετε κάτι καινούριο, ιδιαίτερα στην αρχή της εκπαίδευσής σας, βεβαιωθείτε ότι διαθέτετε αρκετό χρόνο για να μπορέσετε να ολοκληρώσετε την εργασία σας με ένα εργαλείο που γνωρίζετε (π.χ. υπολογιστικό φύλλο). Εάν μόλις τώρα αρχίζετε να μαθαίνε-

85. ΣτΕ: «Application Programming Interface» (διεπαφή προγραμματισμού εφαρμογών)

τε και ο χρόνος σας είναι περιορισμένος, ίσως να θέλετε να χρησιμοποιήσετε ένα οικείο εργαλείο και να περιμένετε μέχρι να έχετε περισσότερο χρόνο στη διάθεσή σας για να πειραματιστείτε.

Όταν αναλαμβάνουν μεγαλύτερα έργα, οι εταιρείες τεχνολογίας χρησιμοποιούν διάφορες μεθόδους για να «σπάνε» τα έργα σε επιμέρους εργασίες και υποεργασίες (έως ότου αυτές να καταστούν αρκετά μικρές και αυτοτελείς, προκειμένου να μπορεί να εκτιμηθεί ο χρόνος που θα απαιτηθεί για να υλοποιηθούν), καθώς και για να απαριθμούν και να ιεραρχούν τις εργασίες ανάλογα με τη σημασία τους.

Οι δημοσιογράφοι δεδομένων μπορούν να αξιοποιήσουν τέτοιες μεθόδους. Για παράδειγμα, σε ένα έργο των *Sunday Times* σχετικά με το ποσοστό των καταγγελλόμενων εγκλημάτων που μπορούν να εξιχνιάσουν οι αστυνομικές δυνάμεις στο Ηνωμένο Βασίλειο, δώσαμε προτεραιότητα στην προβολή των αριθμών που αφορούσαν την περιοχή του κάθε αναγνώστη. Μόλις τελείωσε αυτό και αφού υπήρχε κάποιος επιπλέον χρόνος, ασχοληθήκαμε με το επόμενο αντικείμενο της λίστας: μια απεικόνιση που συγκρίνει την περιοχή του εκάστοτε αναγνώστη με άλλες περιοχές και με τον εθνικό μέσο όρο. Το έργο θα μπορούσε να δημοσιευθεί ανά πάσα στιγμή χάρη στον τρόπο με τον οποίο δουλέψαμε. Αυτή η επαναλαμβανόμενη ροή εργασιών σας βοηθά να εστιάσετε και ταυτόχρονα να διαχειριστείτε τις προσδοκίες.

Αναθεώρηση κώδικα

Οι αίθουσες σύνταξης συχνά διαθέτουν συστήματα ελέγχου ποιότητας για πολλά από τα προϊόντα τους. Ένας ρεπόρτερ δεν καταθέτει απλώς το άρθρο του και αυτό τυπώνεται: το κείμενο ελέγχεται εξονυχιστικά τόσο από τους αρχισυντάκτες όσο και από τους επιμελητές.

Οι προγραμματιστές έχουν τα δικά τους συστήματα, για να διασφαλίζουν την ποιότητα και να αποφεύγουν την εμφάνιση σφαλμάτων σε συλλογικά έργα. Αυτό περιλαμβάνει τις «αναθεωρήσεις κώδικα», όπου ένας προγραμματιστής παρουσιάζει τη δουλειά του και άλλοι την ελέγχουν και την αξιολογούν, καθώς και τις αυτοματοποιημένες δοκιμές του κώδικα.

Σύμφωνα με την έρευνα Global Data Journalism Survey του 2017, το 40% των ομάδων δεδομένων που απάντησαν ήταν τριμελείς έως πενταμελείς και το 30% αυτών αριθμούσε μόνο ένα ή δύο μέλη (Heravi, 2017). Αυτοί οι μικροί αριθμοί αποτελούν πρόκληση για τις πρακτικές εσωτερικού ελέγχου του κώδικα. Έτσι, οι δημοσιογράφοι δεδομένων συχνά εργάζονται μόνοι τους, είτε επειδή δεν έχουν συναδέλφους, είτε επειδή δεν υπάρχουν συστήματα αξιολόγησης από ομότιμους, είτε επειδή δεν υπάρχει κάποιος με τις κατάλληλες δεξιότητες για να αξιολογήσει τον κώδικά τους.

Οι εσωτερικοί μηχανισμοί ποιοτικού ελέγχου, επομένως, μπορεί να αποτελούν μια πολυτέλεια που μόνο λίγες ομάδες δημοσιογραφίας δεδομένων μπορούν να αντέξουν οικονομικά (δεν υπάρχουν επιμελητές για τον προγραμματισμό!). Το κόστος της έλλειψης τέτοιου ελέγχου είναι η ύπαρξη πιθανών σφαλμάτων προγραμματισμού που περνούν απαρατήρητα, η μη βέλτιστη απόδοση ή, το χειρότερο από όλα, άλλα σφάλματα που παραμένουν κρυμμένα. Αυτοί οι περιορισμοί όσον αφορά στους πόρους ίσως είναι σε έναν βαθμό ο λόγος για τον οποίο είναι σημαντικό να αναζητούν οι δημοσιογράφοι τη συμβολή και συνεργασία άλλων οργανισμών, πέρα από τον δικό τους, για το έργο τους –για παράδειγμα, διαδικτυακών κοινοτήτων προγραμματισμού[86].

86. Περισσότερα σχετικά με τη διαφάνεια του κώδικα στη δημοσιογραφία δεδομένων και τις πρακτικές αναθεώρησης, μπορείτε να βρείτε στα κεφάλαια των Λίον και Μαζότε στον παρόντα τόμο (ΣτΕ: Κεφάλαιο 4.5, «Εξηγώντας τις Μεθόδους: Υπολογιστικά Φύλλα, Script και Notebook Προγραμματισμού», και Κεφάλαιο 4.6, «Δουλεύοντας Ανοιχτά στη Δημοσιογραφία Δεδομένων» αντίστοιχα).

Παραπομπές

BBC Data Journalism team (άνευ ημερομηνίας)
«What software do the BBC use?» [Interview], *Warwick*. Διαθέσιμο σε: https://warwick.ac.uk/fac/cross_fac/cim/news/bbc-r-interview/ (Προσπελάστηκε: 26 Νοεμβρίου 2021).

González, M., Hensleigh, E., McLean, M., Segger, M., & Selby-Boothroyd, A. (2018) «How we made the new Big Mac Index interactive», *Source*, 6 Αυγούστου. Διαθέσιμο σε: https://source.opennews.org/articles/how-we-made-new-big-mac-index-interactive/ (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Heravi, B. (2017) «State of data journalism globally: First insights into the global data journalism survey», *Medium*, 1 Αυγούστου. Διαθέσιμο σε: https://medium.com/ucd-ischool/state-of-data-journalism-globally-cb2f4696ad3d (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Ο συγγραφέας

Ο Μπαζίλ Σιμόν είναι συντάκτης γραφικών στο *Reuters Graphics* και Επισκέπτης Λέκτορας Διαδραστικής Δημοσιογραφίας στο Πανεπιστήμιο City στο Λονδίνο.

4.5
Εξηγώντας τις Μεθόδους: Υπολογιστικά Φύλλα, Script και Notebook Προγραμματισμού

Sam Leon

Περίληψη: Αυτό το κεφάλαιο διερευνά τους τρόπους με τους οποίους προγραμματιστικά περιβάλλοντα, όπως τα Jupyter Notebook, μπορούν να βοηθήσουν να γίνει η δημοσιογραφία δεδομένων αναπαραγώγιμη, λιγότερο επιρρεπής σε λάθη και πιο συνεργατική.

Λέξεις-κλειδιά: Jupyter Notebook, αναπαραγωγιμότητα, προγραμματισμός, Python, περιβάλλοντα εγγράμματου προγραμματισμού, δημοσιογραφία δεδομένων

Με την άνθηση της δημοσιογραφίας δεδομένων, οι ιδέες γύρω από το τι μπορεί να θεωρηθεί δημοσιογραφική πηγή αλλάζουν. Οι πηγές μπορούν να λάβουν πλέον πολλές μορφές: δημόσια σύνολα δεδομένων, ηλεκτρονικά μηνύματα που διέρρευσαν, σκαναρισμένα έγγραφα, δορυφορικές εικόνες και δεδομένα αισθητήρων. Παράλληλα, αναδύονται νέες μέθοδοι εύρεσης ιστοριών από αυτές τις πηγές. Η μηχανική μάθηση («machine learning»), η ανάλυση κειμένου και ορισμένες από τις άλλες τεχνικές που εξετάζονται αλλού σε αυτό το βιβλίο χρησιμοποιούνται όλο και περισσότερο στην υπηρεσία της δημοσιογραφικής επιτυχίας.

Ωστόσο, τα δεδομένα, μολονότι φαίνεται να αποπνέουν μια αύρα σκληρής αντικειμενικής αλήθειας, μπορούν να διαστρεβλωθούν και να παραποιηθούν. Υπάρχουν πολλοί τρόποι με τους οποίους οι δημοσιογράφοι δεδομένων μπορούν να σφάλλουν στην ερμηνεία ενός συνόλου δεδομένων και να δημοσιεύσουν μια παραπλανητική ιστορία. Μπορεί να υπάρχουν προβλήματα στο στάδιο της συλλογής των δεδομένων που δεν επιτρέπουν την εξαγωγή γενικών συμπερασμάτων για έναν ευρύτερο πληθυσμό. Αυτό θα μπορούσε, για παράδειγμα, να είναι αποτέλεσμα μεροληπτικής αυτοεπιλογής[87] στον τρόπο με τον οποίο συλλέχθηκε το δείγμα, ένα κοινό πρόβλημα στην εποχή των διαδικτυακών δημοσκοπήσεων και ερευνών. Σφάλματα μπορούν, επίσης, να υπάρξουν στο στάδιο της επεξεργασίας δεδομένων. Η επεξεργασία ή ο καθαρισμός των δεδομένων μπορεί να περιλαμβάνει γεωκωδικοποίηση («geocoding»), διόρθωση ορθογραφικών λαθών σε ονόματα, εναρμόνιση κατηγοριών ή τον αποκλεισμό ορισμένων δεδομένων, εάν, για παράδειγμα, αυτά θεωρούνται στατιστικά ακραία. Ένα καλό παράδειγμα αυτού του είδους σφάλματος είναι η ανακριβής γεωκωδικοποίηση διευθύνσεων IP σε μια ευρέως δημοσιευμένη μελέτη που υποτίθεται ότι κατέδειξε συσχέτιση μεταξύ της πολιτικής πεποίθησης και της κατανάλωσης πορνό (Harris, 2014). Στη συνέχεια, φυσικά, έχουμε την ουσία της εργασίας του δημοσιογράφου δεδομένων, την ανάλυση. Οποιαδήποτε στατιστική αστοχία μπορεί να επηρεάσει αυτό το τμήμα της εργασίας, όπως η λανθασμένη σύγχυση της συσχέτισης

87. ΣτΕ: «Self-selection», πρόκειται για πρακτική εθελοντικής δειγματοληψίας –για παράδειγμα, όταν κοινοποιείται μια έρευνα και οι ανάγκες της, και η συλλογή των στοιχείων βασίζεται στην αυτόκλητη συμμετοχή, δηλαδή σε όσους ανταποκριθούν. Τότε, η αντιπροσωπευτικότητα του δείγματος δεν είναι εγγυημένη.

(«correlation») με την αιτιώδη συνάφεια («causation») ή η επιλογή μιας ακατάλληλης στατιστικής μέτρησης για τη συνοπτική παρουσίαση του εξεταζόμενου συνόλου δεδομένων.

Λαμβάνοντας υπόψη τους τρόπους με τους οποίους η συλλογή, η επεξεργασία και η ανάλυση των δεδομένων μπορούν να μεταβάλλουν μια αφήγηση, πώς μπορεί ο δημοσιογράφος δεδομένων να διαβεβαιώσει τον αναγνώστη ότι οι πηγές που έχει χρησιμοποιήσει είναι αξιόπιστες και ότι η εργασία που έχει εκπονηθεί για την εξαγωγή των συμπερασμάτων του είναι ορθή;

Στην περίπτωση που ο δημοσιογράφος δεδομένων απλώς μεταδίδει τα δεδομένα ή τα ερευνητικά ευρήματα ενός τρίτου μέρους, δεν χρειάζεται να παρεκκλίνει από τα παραδοσιακά δημοσιογραφικά πρότυπα που υιοθετούνται από πολλά μεγάλα μέσα ενημέρωσης. Μια αναφορά στον φορέα που συνέλεξε και ανέλυσε τα δεδομένα σε γενικές γραμμές αρκεί. Για παράδειγμα, ένα πρόσφατο διάγραμμα των *Financial Times* σχετικά με το προσδόκιμο ζωής στο Ηνωμένο Βασίλειο συνοδεύεται από μια σημείωση που αναφέρει: «Πηγή: Υπολογισμοί του Club Vita, βάσει στοιχείων της Eurostat». Επί της αρχής, ο αναγνώστης μπορεί, στη συνέχεια, να αξιολογήσει την αξιοπιστία του αναφερόμενου φορέα. Ενώ ένας υπεύθυνος δημοσιογράφος δημοσιεύει μονάχα μελέτες που θεωρεί αξιόπιστες, ο τρίτος φορέας είναι σε μεγάλο βαθμό υπεύθυνος να λογοδοτήσει σχετικά με τις μεθόδους μέσω των οποίων κατέληξε στα συμπεράσματά του. Σε ένα ακαδημαϊκό πλαίσιο, αυτό πιθανότατα συμπεριλαμβάνει διαδικασίες αξιολόγησης από ομότιμους και, στην περίπτωση των επιστημονικών δημοσιεύσεων, περιλαμβάνει πάντα κάποιο επίπεδο μεθοδολογικής διαφάνειας.

Στην ολοένα και συχνότερη περίπτωση που ο δημοσιογραφικός οργανισμός παράγει την έρευνα δεδομένων, τότε είναι ο ίδιος υπόλογος απέναντι στον αναγνώστη για την αξιοπιστία των αποτελεσμάτων που δημοσιεύει. Οι δημοσιογράφοι έχουν ανταποκριθεί στην πρόκληση της λογοδοσίας για τις μεθόδους τους με διαφορετικούς τρόπους. Μια συνήθης προσέγγιση είναι η περιγραφή της γενικής μεθοδολογίας που χρησιμοποιήθηκε για την εξαγωγή των συμπερασμάτων στο πλαίσιο μιας ιστορίας. Οι περιγραφές αυτές θα πρέπει να διατυπώνονται σε, όσο το δυνατόν περισσότερο, απλή, μη τεχνική γλώσσα, ώστε να είναι κατανοητές από το ευρύτερο δυνατό κοινό. Ένα καλό παράδειγμα αυτής της προσέγγισης χρησιμοποιήθηκε από την εφημερίδα *Guardian* και το Global Witness, για να εξηγήσουν πώς μέτρησαν τους θανάτους περιβαλλοντικών ακτιβιστών για τη σειρά *Environmental Defenders* («Περιβαλλοντικοί υπερασπιστές»), όπως είναι ο πρωτότυπος τίτλος της (Leather, 2017· Leather & Kyte, 2017).

Όμως, όπως συμβαίνει με όλους τους τρόπους απολογισμού της κοινωνικής ζωής, οι γραπτές καταγραφές έχουν τα όριά τους. Το σημαντικότερο πρόβλημα που παρουσιάζουν είναι ότι γενικά δεν προσδιορίζουν τις ακριβείς διαδικασίες που χρησιμοποιούνται για την παραγωγή της ανάλυσης ή την προετοιμασία των δεδομένων. Αυτό καθιστά δύσκολη, ή σε ορισμένες περιπτώσεις αδύνατη, την ακριβή αναπαραγωγή των βημάτων που ακολούθησαν οι δημοσιογράφοι για να καταλήξουν στα συμπεράσματά τους. Με άλλα λόγια, ένας γραπτός απολογισμός, σε γενικές γραμμές, δεν είναι αναπαραγώγιμος. Στο παραπάνω παράδειγμα, όπου τα βήματα συλλογής, επεξεργασίας και ανάλυσης δεδομένων είναι σχετικά απλά, μπορεί να μην αξίζει τον κόπο να προχωρήσουμε πέρα από μια γενική γραπτή περιγραφή. Ωστόσο, όταν χρησιμοποιούνται πιο περίπλοκες τεχνικές, ενδέχεται να υπάρχουν ισχυρά επιχειρήματα για τη χρήση αναπαραγώγιμων προσεγγίσεων.

Αναπαραγώγιμη δημοσιογραφία δεδομένων

Η αναπαραγωγιμότητα θεωρείται ευρέως ως πυλώνας της σύγχρονης επιστημονικής μεθόδου. Βοηθά στη διαδικασία επιβεβαίωσης αποτελεσμάτων αλλά και εντοπισμού και αντιμετώπισης προβληματικών ευρημάτων ή αμφισβητήσιμων θεωριών. Καταρχήν, οι ίδιοι μηχανισμοί μπορούν να βοηθήσουν στην εξάλειψη των εσφαλμένων ή παραπλανητικών χρήσεων των δεδομένων σε δημοσιογραφικό πλαίσιο.

Μια ματιά σε ένα από τα πιο προβεβλημένα μεθοδολογικά λάθη στην πρόσφατη ακαδημαϊκή

ιστορία μπορεί να είναι χρήσιμη. Σε μια δημοσίευση του 2010, οι Κάρμεν Ρέινχαρτ και Κένεθ Ρόγκοφ από το Χάρβαρντ επιχείρησαν να αποδείξουν ότι ο μέσος ρυθμός πραγματικής οικονομικής ανάπτυξης είναι -0,1%, όταν το δημόσιο χρέος μιας χώρας αυξάνεται σε ποσοστό άνω του 90% του ακαθάριστου εγχώριου προϊόντος (Reinhart & Rogoff, 2010). Ο αριθμός αυτός χρησιμοποιήθηκε, στη συνέχεια, ως πυρομαχικό από τους πολιτικούς που υποστήριζαν τα μέτρα λιτότητας.

Όπως αποδείχθηκε, η ανάλυση ήταν βασισμένη σε ένα λάθος στο Excel: αντί να πάρουν τον μέσο όρο μιας ολόκληρης σειράς χωρών, οι Ρέινχαρτ και Ρόγκοφ είχαν κάνει ένα λάθος στη φόρμουλά τους, πράγμα που σήμαινε ότι είχαν ενσωματωθεί μόνο οι 15 από τις 20 χώρες που εξέτασαν. Λαμβάνοντας υπόψη όλες τις χώρες, η «πτώση» του 0,1% μετατράπηκε σε μέση αύξηση της οικονομικής ανάπτυξης κατά 2,2%.

Το λάθος έγινε αντιληπτό μόνο όταν ο υποψήφιος διδάκτορας Τόμας Χέρντον και οι καθηγητές Μάικλ Ας και Ρόμπερτ Πόλιν εξέτασαν το αρχικό υπολογιστικό φύλλο με το οποίο είχαν δουλέψει οι Ρέινχαρτ και Ρόγκοφ. Αυτό δείχνει πόσο σημαντικό είναι να μην είναι μόνο η μέθοδος γραμμένη σε απλή γλώσσα –αλλά το ίδιο και τα δεδομένα και η τεχνολογία που χρησιμοποιήθηκαν για την ανάλυση. Ωστόσο, το λάθος των Ρέινχαρτ και Ρόγκοφ ίσως καταδεικνύει και κάτι άλλο: το Microsoft Excel –και γενικά τα λογισμικά υπολογιστικών φύλλων– μπορεί να μην είναι η καλύτερη τεχνολογία για τη δημιουργία αναπαραγώγιμων αναλύσεων.

Από τον σχεδιασμό του το Excel αποκρύπτει μεγάλο μέρος της διαδικασίας της εργασίας με δεδομένα. Οι φόρμουλες –οι οποίες εκτελούν το μεγαλύτερο μέρος της δουλειάς της ανάλυσης σε ένα υπολογιστικό φύλλο– είναι ορατές μόνο όταν κανείς κάνει κλικ σε ένα κελί. Αυτό σημαίνει ότι είναι πιο δύσκολο να επανεξεταστούν τα πραγματικά βήματα που ακολουθήθηκαν για την εξαγωγή ενός συγκεκριμένου συμπεράσματος. Αν και δεν θα το μάθουμε ποτέ με βεβαιότητα, μπορούμε να φανταστούμε ότι, εάν η ανάλυση των Ρέινχαρτ και Ρόγκοφ είχε γίνει σε μια γλώσσα στην οποία τα βήματα έπρεπε να δηλωθούν ρητά (π.χ. σε γλώσσα προγραμματισμού), το σφάλμα θα μπορούσε να είχε εντοπιστεί πριν από τη δημοσίευση.

Η ροή εργασιών με βάση το Excel ενθαρρύνει γενικά την κατάργηση σταδίων για την εξαγωγή συμπερασμάτων. Είναι οι τιμές –και όχι οι φόρμουλες– που συχνά αντιγράφονται σε άλλα φύλλα ή στήλες, με αποτέλεσμα το πλήκτρο «αναίρεση» να είναι η μόνη οδός επιστροφής στον τρόπο με τον οποίο δημιουργήθηκε ένας συγκεκριμένος αριθμός. Το ιστορικό των «αναιρέσεων», βέβαια, γενικά διαγράφεται, όταν κλείνει μια εφαρμογή και, επομένως, δεν είναι το κατάλληλο μέρος για την αποθήκευση σημαντικών μεθοδολογικών πληροφοριών.

Η άνθηση του περιβάλλοντος εγγράμματου προγραμματισμού: τα Jupyter Notebook στην αίθουσα σύνταξης

Μια αναδυόμενη προσέγγιση της μεθοδολογικής διαφάνειας είναι η χρήση των λεγόμενων περιβαλλόντων «εγγράμματου προγραμματισμού». Οργανισμοί όπως το *Buzzfeed*, οι *New York Times* και το *Correctiv* τα χρησιμοποιούν, για να παράγουν έγγραφα αναγνώσιμα από τον άνθρωπο, τα οποία μπορούν, επίσης, να εκτελεστούν από μια μηχανή, προκειμένου να αναπαραχθούν ακριβώς τα βήματα που ακολουθήθηκαν σε μια συγκεκριμένη ανάλυση[88].

Ο εγγράμματος προγραμματισμός («literate programming»), στον οποίο αναφέρθηκε για πρώτη φορά ο Ντόναλντ Κνουτ τη δεκαετία του 1980, είναι μια προσέγγιση για τη συγγραφή κώδικα, όπου ο συντάκτης διανθίζει τον κώδικα με απλή ανθρώπινη γλώσσα με την οποία εξηγεί τα βήματα που ακολουθούνται (Knuth, 1992). Τα δύο βασικά περιβάλλοντα εγγράμματου προ-

88. https://github.com/TheUpshot

γραμματισμού που χρησιμοποιούνται σήμερα είναι τα Jupyter Notebook και R Markdown[89]. Και τα δύο παράγουν έγγραφα αναγνώσιμα από τον άνθρωπο που συνδυάζουν απλή γλώσσα, οπτικοποιήσεις και κώδικα σε ένα ενιαίο έγγραφο το οποίο μπορεί να αποθηκευτεί σε .html και να δημοσιευθεί στον Ιστό. Τα πρωτογενή δεδομένα μπορούν να συνδεθούν με σαφήνεια, ενώ τυχόν άλλες τεχνικές εξαρτήσεις, όπως βιβλιοθήκες τρίτων, προσδιορίζονται ξεκάθαρα.

Σε αυτά τα περιβάλλοντα, όχι μόνο δίνεται έμφαση σε επεξηγήσεις που διαβάζονται από οποιονδήποτε άνθρωπο, αλλά και ο κώδικας είναι διατεταγμένος έτσι ώστε να αντικατοπτρίζει την ανθρώπινη λογική. Επομένως, τα έγγραφα που συντάσσονται με αυτό το πρότυπο μπορούν να διαβαστούν σαν ένα σύνολο βημάτων ενός επιχειρήματος ή μια σειρά απαντήσεων σε ένα σύνολο ερευνητικών ερωτημάτων.

> Αυτός που χρησιμοποιεί τον εγγράμματο προγραμματισμό μπορεί να θεωρηθεί ένας δοκιμιογράφος, του οποίου το κύριο μέλημα είναι η παρουσίαση και η τελειότητα του ύφους. Ένας τέτοιος συντάκτης, με το λεξικό συνωνύμων στο χέρι, επιλέγει προσεκτικά τα ονόματα των μεταβλητών και εξηγεί τι σημαίνει η κάθε μεταβλητή. Πασχίζει να φτιάξει ένα πρόγραμμα που να είναι κατανοητό, επειδή οι έννοιές του έχουν εισαχθεί με τη σειρά που είναι η καλύτερη δυνατή για την ανθρώπινη κατανόηση, χρησιμοποιώντας ένα μείγμα τυπικών και άτυπων μεθόδων που αλληλοενισχύονται (Knuth, 1984).

Ένα καλό παράδειγμα αυτής της μορφής είναι το Jupyter Notebook του *Buzzfeed News* που περιγράφει λεπτομερώς πώς ανέλυσαν τις τάσεις στις πυρκαγιές της Καλιφόρνιας[90]. Ενώ το notebook περιέχει όλο τον κώδικα και τους συνδέσμους προς τα πρωτογενή δεδομένα που απαιτούνται για την αναπαραγωγή της ανάλυσης, η ουσία του εγγράφου είναι μια αφήγηση ή μια συνομιλία με τα αρχικά δεδομένα. Οι εξηγήσεις παρατίθενται σε τίτλους που ακολουθούν μια λογική ερευνητική γραμμή. Οι οπτικοποιήσεις και τα γραφήματα χρησιμοποιούνται για να αναδείξουν τα κύρια θέματα.

Μια πτυχή της «εγγράμματης» προσέγγισης του προγραμματισμού είναι ότι τα παραγόμενα έγγραφα (ως αρχεία Jupyter Notebook ή R Markdown) μπορεί να είναι ικανά να καθησυχάσουν ακόμη και εκείνους τους αναγνώστες που δεν μπορούν να διαβάσουν τον ίδιο τον κώδικα ότι τα βήματα που ακολουθήθηκαν για την παραγωγή των συμπερασμάτων είναι ορθά. Η ιδέα είναι παρόμοια με την περιγραφή του Στίβεν Σέιπιν και του Σάιμον Σέιφερ για την «εικονική μαρτυρία» («virtual witnessing») ως μέσο τεκμηρίωσης γεγονότων στην πρώιμη σύγχρονη επιστήμη. Χρησιμοποιώντας ως παράδειγμα το πειραματικό πρόγραμμα του Ρόμπερτ Μπόιλ, οι Σέιπιν και Σέιφερ παρουσίασαν τον ρόλο που έπαιζε η «εικονική μαρτυρία»:

> Η τεχνολογία της εικονικής μαρτυρίας περιλαμβάνει την παραγωγή στο μυαλό του αναγνώστη μιας τέτοιας εικόνας μιας πειραματικής σκηνής, η οποία παρακάμπτει την ανάγκη είτε άμεσης μαρτυρίας είτε αναπαραγωγής. Μέσω της εικονικής μαρτυρίας, ο πολλαπλασιασμός των μαρτύρων θα μπορούσε, επί της αρχής, να είναι απεριόριστος. Ως εκ τούτου, ήταν η πιο ισχυρή τεχνολογία για τη σύνθεση γεγονότων. Η επικύρωση των πειραμάτων και η αναγνώριση των αποτελεσμάτων τους ως γεγονότων, συνεπαγόταν αναγκαστικά την πραγματοποίησή τους στο εργαστήριο του νου και του οφθαλμού του νου. Αυτό που απαιτούνταν ήταν μια τεχνολογία εμπιστοσύνης και διαβεβαίωσης ότι τα πράγματα είχαν γίνει με τον τρόπο που ισχυρίζονταν πως έγιναν (Shapin & Schaffer, 1985).

Τα έγγραφα που παράγονται από περιβάλλοντα εγγράμματου προγραμματισμού, όπως τα Jupyter Notebook –όταν δημοσιεύονται παράλληλα με τα άρθρα– μπορεί να έχουν ένα πα-

89. http://jupyter.org/, https://rmarkdown.rstudio.com/
90. https://buzzfeednews.github.io/2018-07-wildfire-trends/

ρόμοιο αποτέλεσμα, καθώς επιτρέπουν στον αναγνώστη που δεν είναι προγραμματιστής να κάνει εικόνα τα βήματα που ακολουθήθηκαν για την παραγωγή των ευρημάτων μιας συγκεκριμένης ιστορίας. Ενώ ο αναγνώστης που δεν είναι προγραμματιστής μπορεί να μην είναι σε θέση να κατανοήσει ή να εκτελέσει τον ίδιο τον κώδικα, τα σχόλια και οι επεξηγήσεις στο έγγραφο μπορεί να είναι ικανά να τον καθησυχάσουν ότι ελήφθησαν τα κατάλληλα μέτρα για τον μετριασμό των σφαλμάτων.

Πάρτε, για παράδειγμα, ένα πρόσφατο ρεπορτάζ του *Buzzfeed News* σχετικά με τις επιθεωρήσεις σε δομές φροντίδας παιδιών στο Ηνωμένο Βασίλειο[91]. Το Jupyter Notebook περιλαμβάνει συγκεκριμένα βήματα, προκειμένου να ελεγχθεί ότι τα δεδομένα έχουν φιλτραριστεί σωστά (Εικόνα 4.5.1), παρέχοντας ένα εφεδρικό σύστημα κατά των απλών αλλά σοβαρών λαθών που εξέθεσαν τους Ρέινχαρτ και Ρόγκοφ[92]. Μολονότι το ακριβές περιεχόμενο του κώδικα μπορεί να μην είναι κατανοητό από τον αναγνώστη που δεν έχει τεχνικές γνώσεις, η παρουσία αυτών των δοκιμών και των δικλείδων ασφαλείας για την αποφυγή σφαλμάτων με εξηγήσεις σε επαρκώς κατανοητά αγγλικά μπορεί να δείξει ότι τα βήματα που ακολουθήθηκαν για την παραγωγή των ευρημάτων του δημοσιογράφου ήταν ορθά.


```
In [11]: # Make sure that we've identified all private-sector owners
         assert (
             as_at_data_filtered
             .loc[lambda df: df["Sector"] == "Private"]
             ["Owner"].isnull()
             .sum()
         ) == 0
```

Εικόνα 4.5.1: Κελί από το Jupyter Notebook του *Buzzfeed News*, με επεξηγήσεις ή σχόλια σε ανθρώπινη γλώσσα. Εδώ, εξηγείται ότι σκοπός αυτού του κώδικα είναι να ελέγξει πως το φιλτράρισμα των πρωτογενών δεδομένων έγινε σωστά. Πηγή: Jeremy Singer-Vine/*Buzzfeed News*.

Κάτι περισσότερο από απλή αναπαραγωγιμότητα

Η χρήση περιβαλλόντων εγγράμματου προγραμματισμού για ιστορίες δεδομένων δεν βοηθά απλώς να γίνουν αυτές πιο αναπαραγώγιμες.

Η δημοσίευση του κώδικα μπορεί να ενισχύσει τη συνεργασία μεταξύ οργανισμών. Το 2016, το Global Witness δημοσίευσε ένα εργαλείο scraping που αντλούσε στοιχεία για εταιρείες και τους μετόχους τους από το μητρώο εταιρειών της Παπούα-Νέας Γουινέας[93]. Η αρχική έρευνα αποσκοπούσε στον εντοπισμό των βασικών δικαιούχων του επιρρεπούς στη διαφθορά εμπορίου τροπικής ξυλείας, το οποίο έχει καταστροφικές επιπτώσεις στις τοπικές κοινότητες. Αν και το Global Witness δεν είχε άμεσα σχέδια για την επαναχρησιμοποίηση του εργαλείου που ανέπτυξε, ο κώδικας που είχε αναπτυχθεί δημοσιεύθηκε στο GitHub –τον δημοφιλή ιστότοπο διαμοιρασμού κώδικα.

Λίγο αργότερα, η οργάνωση κοινοτικής δράσης ACT NOW! κατέβασε τον κώδικα του scraper, τον βελτίωσε και τον ενσωμάτωσε στο δικό της πρότζεκτ iPNG, το οποίο επιτρέπει στο κοινό να διασταυρώνει τα ονόματα μετόχων και διευθυντών εταιρειών με άλλες πηγές δημοσίου συμφέ-

91. https://www.buzzfeed.com/richholmes/care-price
92. https://github.com/BuzzFeedNews/2018-07-ofsted-inspections/blob/master/notebooks/00-analyze-ofsted-data.ipynb
93. https://github.com/Global-Witness/papua-new-guinea-ipa

ροντος[94]. Το συγκεκριμένο εργαλείο αποτελεί πλέον μέρος της βασικής υποδομής δεδομένων του δικτυακού τόπου, αντλώντας δεδομένα από το μητρώο εταιρειών της Παπούα-Νέας Γουινέας δύο φορές τον χρόνο.

Η συγγραφή κώδικα σε ένα περιβάλλον εγγράμματου προγραμματισμού μπορεί, επίσης, να βοηθήσει στη βελτίωση ορισμένων εσωτερικών διαδικασιών, όταν άλλοι εντός ενός οργανισμού πρέπει να κατανοήσουν και να ελέγξουν μια ανάλυση πριν από τη δημοσίευση. Στο Global Witness, τα Jupyter Notebook έχουν χρησιμοποιηθεί για τη βελτιστοποίηση της διαδικασίας νομικού ελέγχου. Καθώς τα notebook παραθέτουν τα βήματα που ακολουθήθηκαν για την επίτευξη ενός συγκεκριμένου πορίσματος με λογική σειρά, οι δικηγόροι μπορούν στη συνέχεια να προβούν σε ακριβέστερη αξιολόγηση των νομικών κινδύνων που συνδέονται με έναν συγκεκριμένο ισχυρισμό.

Στο πλαίσιο της ερευνητικής δημοσιογραφίας, ένας τομέας όπου αυτό μπορεί να είναι ιδιαίτερα σημαντικό είναι όταν γίνονται υποθέσεις σχετικά με την ταυτότητα συγκεκριμένων ατόμων που αναφέρονται σε ένα σύνολο δεδομένων. Στο πλαίσιο του πρόσφατου έργου μας σχετικά με την κατάσταση της εταιρικής διαφάνειας στο Ηνωμένο Βασίλειο, θελήσαμε να προσδιορίσουμε ποια άτομα είχαν τον έλεγχο ενός πολύ μεγάλου αριθμού εταιρειών. Αυτό αποτελεί ένδειξη (αν και όχι απόδειξη) ότι τα άτομα αυτά λειτουργούν ως οι λεγόμενοι «εντολοδόχοι», γεγονός το οποίο σε ορισμένες περιπτώσεις –όπως όταν το άτομο αναφέρεται ως ο πραγματικός δικαιούχος– είναι παράνομο. Κατά τη δημοσίευση του καταλόγου με τα ονόματα των ατόμων που είχαν υπό τον έλεγχό τους τις περισσότερες εταιρείες, η νομική ομάδα ήθελε να μάθει πώς γνωρίζαμε ότι ένα συγκεκριμένο άτομο, ας πούμε ο Τζον Μπάρι Σμιθ, ήταν το ίδιο με ένα άλλο άτομο με το όνομα Τζον Μπ. Σμιθ[95]. Ένα Jupyter Notebook ήταν σε θέση να αποτυπώσει με σαφήνεια τον τρόπο με τον οποίο είχαμε εκτελέσει αυτό το είδος αποδελτίωσης, παρουσιάζοντας έναν πίνακα στο σχετικό βήμα που καθόριζε τα πεδία που χρησιμοποιήθηκαν, για να επιβεβαιωθεί η ταυτότητα των ατόμων[96]. Οι ίδιες διαδικασίες έχουν χρησιμοποιηθεί και στο Global Witness, προκειμένου να ελεγχθούν τα γεγονότα.

Τα Jupyter Notebook έχουν, επίσης, αποδειχθεί ιδιαίτερα χρήσιμα στο Global Witness, όταν υπάρχει η ανάγκη παρακολούθησης ενός συγκεκριμένου συνόλου δεδομένων με την πάροδο του χρόνου. Για παράδειγμα, το 2018 το Global Witness θέλησε να διαπιστώσει πώς μεταβλήθηκε ο κίνδυνος διαφθοράς στην αγορά ακινήτων του Λονδίνου κατά τη διάρκεια μιας διετίας[97]. Εξασφαλίσαμε από το κτηματολόγιο μια νέα αποτύπωση των ακινήτων που ανήκουν σε ξένες εταιρείες, και επαναχρησιμοποιήσαμε και δημοσιεύσαμε ένα notebook που είχαμε αναπτύξει για τον ίδιο σκοπό δύο χρόνια νωρίτερα[98]. Αυτό απέφερε συγκρίσιμα αποτελέσματα με ελάχιστο κόστος. Το notebook είχε ένα επιπλέον πλεονέκτημα και σε αυτό το πλαίσιο: επέτρεψε στο Global Witness να παρουσιάσει τη μεθοδολογία του, ελλείψει της δυνατότητας να αναδημοσιεύσει τα υποκείμενα πρωτογενή δεδομένα τα οποία, κατά τη στιγμή της ανάλυσης, είχαν ορισμένους περιορισμούς αδειοδότησης. Αυτό είναι κάτι που πολύ δύσκολα μπορεί να γίνει σε μια ροή εργασίας η οποία βασίζεται σε υπολογιστικά φύλλα. Φυσικά, ο πιο αποτελεσματικός τρόπος λογοδοσίας για τη μέθοδό σας θα είναι πάντα η δημοσίευση των ανεπεξέργαστων δεδομένων που χρησιμοποιήσατε. Ωστόσο, οι δημοσιογράφοι συχνά χρησιμοποιούν δεδομένα που δεν μπορούν να αναδημοσιευθούν για λόγους πνευματικών δικαιωμάτων, προστασίας της ιδιωτικής ζωής ή προστασίας των πηγών τους.

94. https://pngiportal.org/
95. https://www.globalwitness.org/en/campaigns/corruption-and-money-laundering/anonymous-company-owners/companies-we-keep/
96. https://github.com/Global-Witness/the-companies-we-keep/blob/master/companies_we_keep_analysis.ipynb
97. https://www.globalwitness.org/en/blog/two-years-still-dark-about-86000-anonymously-owned-uk-homes/
98. https://github.com/Global-Witness/overseas-companies-land-ownership-public/blob/master/overseas_companies_land_ownership_analysis.ipynb

Παρόλο που τα περιβάλλοντα εγγράμματου προγραμματισμού μπορούν σαφώς να ενισχύσουν τη δυνατότητα λογοδοσίας και την αναπαραγωγιμότητα της εργασίας ενός δημοσιογράφου δεδομένων, παράλληλα με άλλα οφέλη, υπάρχουν ορισμένοι σημαντικοί περιορισμοί.

Ένας τέτοιος περιορισμός είναι ότι, για να αναπαράξετε (και όχι απλώς να ακολουθήσετε ή να γίνετε «εικονικός μάρτυρας») μια προσέγγιση που περιγράφεται σε ένα Jupyter Notebook ή R Markdown, θα πρέπει να ξέρετε πώς να γράφετε, ή τουλάχιστον να εκτελείτε, κώδικα. Το γεγονός ότι η δημοσιογραφία δεδομένων βρίσκεται, σε γενικές γραμμές, σε ένα εκκολαπτόμενο στάδιο σημαίνει ότι εξακολουθεί να υπάρχει μια σχετικά μικρή ομάδα δημοσιογράφων (πόσο μάλλον οι γενικοί καταναλωτές της δημοσιογραφίας) που μπορούν να προγραμματίσουν. Αυτό σημαίνει ότι είναι απίθανο τα αποθετήρια των εφημερίδων στο GitHub να τύχουν του ίδιου επιπέδου ελέγχου όπως, ας πούμε, ο αναθεωρημένος από ομότιμους κώδικας που αναφέρεται σε ένα ακαδημαϊκό περιοδικό, όπου μεγαλύτερα τμήματα της κοινότητας μπορούν πραγματικά να εξετάσουν τον ίδιο τον κώδικα. Ως εκ τούτου, η δημοσιογραφία δεδομένων μπορεί να είναι πιο επιρρεπής σε κρυφά σφάλματα εντός του ίδιου του κώδικα σε σύγκριση με την έρευνα που απευθύνεται σε ένα πιο τεχνικά καταρτισμένο κοινό. Όπως επισημαίνει ο Τζεφ Χάρις (2013), μπορεί να μην αργήσουμε να δούμε διορθώσεις προγραμματισμού να δημοσιεύονται παράλληλα με τις παραδοσιακές δημοσιεύσεις διόρθωσης δημοσιογραφικών λαθών. Αξίζει να σημειωθεί σε αυτό το πλαίσιο ότι αρχίζουν να αναπτύσσονται για τους δημοσιογράφους εργαλεία όπως το Workbench (το οποίο αναφέρεται επίσης στο κεφάλαιο του Στρέι σε αυτό το βιβλίο)[99], τα οποία υπόσχονται να παρέχουν ορισμένες από τις λειτουργίες των περιβαλλόντων εγγράμματου προγραμματισμού χωρίς την ανάγκη συγγραφής ή κατανόησης κώδικα[100].

Στο σημείο αυτό, αξίζει, επίσης, να εξετάσουμε εάν οι νέοι μηχανισμοί λογοδοσίας στη δημοσιογραφία μπορεί να μην είναι απλώς νέα μέσα μέσω των οποίων ένα προϋπάρχον «κοινό» είναι σε θέση να ελέγξει τις μεθόδους που χρησιμοποιούνται, αλλά να παίζουν πράγματι ρόλο στη διαμόρφωση νέων τύπων «κοινού». Αυτό επισημαίνει ο Άντριου Μπάρι στο δοκίμιό του *Transparency as a Political Device* («Η Διαφάνεια ως Πολιτικό Μέσο»):

> Η διαφάνεια δεν συνεπάγεται μόνο τη δημοσίευση συγκεκριμένων πληροφοριών· συνεπάγεται, επίσης, τη διαμόρφωση μιας κοινωνίας που είναι σε θέση να αναγνωρίζει και να αξιολογεί την αξία των πληροφοριών που δημοσιοποιούνται –και, εάν είναι απαραίτητο, να τις τροποποιεί. Η λειτουργία της διαφάνειας απευθύνεται σε κατά τόπους μάρτυρες, αλλά οι μάρτυρες αυτοί αναμένεται να είναι κατάλληλα συγκροτημένοι και η παρουσία τους να επιβεβαιώνεται. Υπάρχει, επομένως, μια κυκλική σχέση μεταξύ της σύνθεσης των πολιτικών συγκεντρώσεων και των απολογισμών της πετρελαϊκής οικονομίας –η μία δημιουργεί την άλλη. Η διαφάνεια δεν αποσκοπεί απλώς στη δημοσιοποίηση των πληροφοριών αλλά στη διαμόρφωση ενός κοινού που ενδιαφέρεται να ενημερωθεί (Barry, 2010).

Οι μέθοδοι που αναπτύχθηκαν παραπάνω για την καταγραφή των έργων της δημοσιογραφίας δεδομένων από μόνες τους μπορεί να διαδραματίσουν ρόλο στην εμφάνιση νέων ομάδων κοινού με μεγαλύτερη τεχνική επίγνωση, που επιθυμούν να ελέγξουν και να θέσουν τους δημοσιογράφους προ των ευθυνών τους με τρόπους οι οποίοι δεν ήταν δυνατοί προηγουμένως –πριν από την έλευση και τη χρήση τεχνολογιών όπως τα περιβάλλοντα εγγράμματου προγραμματισμού.

Η ιδέα αυτή ανταποκρίνεται σε ορισμένες από τις εργασίες του Global Witness σχετικά με τις στοιχειώδεις γνώσεις πάνω στα δεδομένα, προκειμένου να ενισχυθεί η λογοδοσία του εξορυκτικού τομέα. Μια νομοθεσία-ορόσημο στην Ευρωπαϊκή Ένωση που υποχρεώνει τις εταιρείες

99. ΣτΕ: Κεφάλαιο 4.7, «Αξιοποιώντας τους Αλγορίθμους για το Ρεπορτάζ».
100. http://workbenchdata.com/

εξόρυξης να δημοσιεύουν τις πληρωμές σε επίπεδο έργου προς τις κυβερνήσεις για έργα πετρελαίου, φυσικού αερίου και εξόρυξης, έναν τομέα ιδιαίτερα επιρρεπή στη διαφθορά, έδωσε τη δυνατότητα για πολύ μεγαλύτερο έλεγχο του πού συσσωρεύονται στην πραγματικότητα αυτά τα έσοδα. Ωστόσο, το Global Witness και άλλες ομάδες πίεσης στο πλαίσιο της συμμαχίας Publish What You Pay έχουν παρατηρήσει εδώ και καιρό ότι δεν υφίσταται ένα προϋπάρχον «κοινό» που θα μπορούσε να παίξει άμεσα αυτόν τον ρόλο. Ως αποτέλεσμα, το Global Witness και άλλοι έχουν αναπτύξει πόρους και εκπαιδευτικά προγράμματα, για να συγκεντρώσουν δημοσιογράφους και ομάδες της κοινωνίας των πολιτών σε χώρες πλούσιες σε πόρους, με σκοπό να τους υποστηρίξουν ώστε να αναπτύξουν τις δεξιότητες για να χρησιμοποιούν αυτά τα δεδομένα, προκειμένου να θέτουν με μεγαλύτερη ευκολία τις εταιρείες προ των ευθυνών τους. Μια συνιστώσα αυτής της προσπάθειας ήταν η ανάπτυξη και δημοσίευση συγκεκριμένων μεθοδολογιών για τον εντοπισμό ύποπτων αναφορών πληρωμών που θα μπορούσαν να είναι προϊόν διαφθοράς[101].

Τα περιβάλλοντα εγγράμματου προγραμματισμού αποτελούν επί του παρόντος ένα πολλά υποσχόμενο μέσο, μέσω του οποίου οι δημοσιογράφοι δεδομένων εξασφαλίζουν μεγαλύτερη διαφάνεια και υπευθυνότητα για τις μεθοδολογίες τους. Ενώ τα δεδομένα θα παραμένουν πάντα ανοιχτά σε πολλαπλές ερμηνείες, οι τεχνολογίες που καθιστούν σαφείς τις υποθέσεις των δημοσιογράφων και τις μεθόδους τους αναπαραγώγιμες είναι πολύτιμες. Ενισχύουν τη συνεργασία και επιτρέπουν την εξέταση ενός ολοένα και περισσότερο τεχνικού κλάδου από διάφορες ομάδες κοινού. Δεδομένης της τρέχουσας κρίσης εμπιστοσύνης στη δημοσιογραφία, η ευρύτερη υιοθέτηση αναπαραγώγιμων προσεγγίσεων μπορεί να είναι ένας σημαντικός τρόπος με τον οποίο οι ομάδες δεδομένων μπορούν να διατηρήσουν την αξιοπιστία τους.

Παραπομπές

Barry, A. (2010) «Transparency as a political device», στο M. Akrich, Y. Barthe, F. Muniesa, & P. Mustar (επιμ.), *Débordements: Mélanges offerts à Michel Callon, σσ.* 21–39. Παρίσι: Presses des Mines. Διαθέσιμο σε: http://books.openedition.org/pressesmines/721 (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Harris, J. (2013) «The Times regrets the programmer error», *Source*, 19 Σεπτεμβρίου. Διαθέσιμο σε: https://source.opennews.org/articles/times-regrets-programmer-error/ (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Harris, J. (2014) «Distrust your data», *Source*, 22 Μαΐου. Διαθέσιμο σε: https://source.opennews.org/articles/distrust-your-data/ (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Knuth, D. E. (1984) «Literate programming», *The Computer Journal*, 27(2), *σσ.* 97-111. DOI: https://doi.org/10.1093/comjnl/27.2.97

Knuth, D. E. (1992). *Literate programming*. Στάνφορντ, Καλιφόρνια: Center for the Study of Language and Information.

Leather, B. (2017) «Environmental defenders: Who are they and how do we decide if they have died in defense of their environment?», *The Guardian*, 13 Ιουλίου. Διαθέσιμο σε: https://www.theguardian.com/environment/2017/jul/13/environmental-defenders-who-are-they-and-how-do-we-decide-if-they-have-died-in-defence-of-their-environment (Προσπελάστηκε: 26 Νοεμβρίου 2021).

Leather, B., & Kyte, B. (2017) «Defenders: Methodology», *Global Witness*, 13 Ιουλίου. Διαθέσιμο σε: https://www.globalwitness.org/en/campaigns/environmental-activists/defendersmethodology/ (Προσπελάστηκε: 26 Νοεμβρίου 2021).

101. https://www.globalwitness.org/en/campaigns/oil-gas-and-mining/finding-missing-millions/

Reinhart, C. M., & Rogoff, K. S. (2010) Growth in a time of debt (Working Paper No. 15639), *National Bureau of Economic Research*. DOI: https://doi.org/10.3386/w15639

Shapin, S., & Schaffer, S. (1985). *Leviathan and the air-pump: Hobbes, Boyle, and the experimental life.* Πρίνστον, Νιου Τζέρσεϊ: Princeton University Press.

Ο συγγραφέας

Ο Σαμ Λίον είναι επικεφαλής ερευνών δεδομένων στο Global Witness, όπου διεξάγει έρευνες και εκστρατείες κατά της διαφθοράς και της περιβαλλοντικής καταστροφής.

4.6
Δουλεύοντας Ανοιχτά στη Δημοσιογραφία Δεδομένων

Natalia Mazotte

Περίληψη: Αυτό το κεφάλαιο εξετάζει ορισμένα παραδείγματα και οφέλη από την περισσότερο ανοιχτή εργασία των δημοσιογράφων δεδομένων, καθώς και ορισμένους τρόπους για να ξεκινήσετε.

Λέξεις-κλειδιά: δημοσιογραφία δεδομένων, ανοιχτός κώδικας, ελεύθερο λογισμικό, διαφάνεια, εμπιστοσύνη, προγραμματισμός

Πολλά εξέχοντα έργα λογισμικού και Διαδικτύου –όπως το Linux, το Android, η Wikipedia, το WordPress και το TensorFlow– έχουν αναπτυχθεί συνεργατικά με βάση την ελεύθερη ροή της γνώσης[102]. Ο Στάλμαν, ένας διακεκριμένος χάκερ που ίδρυσε το GNU Project και το Ίδρυμα Ελεύθερου Λογισμικού («Free Software Foundation»), λέει ότι όταν άρχισε να εργάζεται στο MIT το 1971, ο διαμοιρασμός του πηγαίου κώδικα λογισμικού ήταν τόσο συνηθισμένος όσο η ανταλλαγή συνταγών (Stallman, 2002).

Για χρόνια μια τέτοια ανοιχτή προσέγγιση ήταν αδιανόητη για τη δημοσιογραφία. Στις αρχές της καριέρας μου ως δημοσιογράφος, συνεργάστηκα με κοινότητες ανοιχτού κώδικα στη Βραζιλία και άρχισα να βλέπω την ανοιχτότητα ως το μοναδικό βιώσιμο μονοπάτι για τη δημοσιογραφία. Όμως, η διαφάνεια δεν αποτελεί προτεραιότητα ή βασική αξία για τους δημοσιογράφους και τους οργανισμούς μέσων ενημέρωσης. Για μεγάλο μέρος της σύγχρονης ιστορίας της, η δημοσιογραφία ασκείται με βάση ένα παράδειγμα ανταγωνισμού για τις λιγοστές πληροφορίες.

Όταν η πρόσβαση στις πληροφορίες είναι προνόμιο των λίγων και όταν ένα σημαντικό πόρισμα είναι διαθέσιμο μόνο σε αυτόπτες μάρτυρες ή σε κλειστούς κύκλους, οι τρόποι διασφάλισης της λογοδοσίας είναι περιορισμένοι. Η παράθεση ενός εγγράφου ή η αναφορά μιας πηγής για μια συνέντευξη μπορεί να μην απαιτεί τόσο περίπλοκους μηχανισμούς διαφάνειας. Σε ορισμένες περιπτώσεις, η διατήρηση του απορρήτου σημαίνει την εξασφάλιση της ασφάλειας της πηγής, και αυτό είναι μέχρι και επιθυμητό. Όταν οι πληροφορίες είναι, όμως, άφθονες, το να μη διαμοιραζόμαστε πώς φτάσαμε ως εκεί μπορεί να στερήσει από τον αναγνώστη τα μέσα για να κατανοήσει την ιστορία και να βγάλει νόημα από αυτήν.

Καθώς οι δημοσιογράφοι και παρουσιάζουν τα δεδομένα και τους αλγορίθμους και βασίζονται σε αυτά, θα μπορούσαν να υιοθετήσουν ένα ήθος παρόμοιο με αυτό των κοινοτήτων ανοιχτού κώδικα; Ποια είναι τα πλεονεκτήματα των δημοσιογράφων που υιοθετούν τις αναδυόμενες ψηφιακές πρακτικές και αξίες που συνδέονται με αυτές τις κοινότητες; Αυτό το κεφάλαιο εξετάζει ορισμένα παραδείγματα και οφέλη από την περισσότερο ανοιχτή εργασία των δημοσιογράφων δεδομένων, καθώς και ορισμένους τρόπους για να ξεκινήσετε[103].

102. Αυτό το κεφάλαιο γράφτηκε από τη Νατάλια Μαζότε με τη συμβολή του Μάρκο Τούλιο Πίρες.
103. Για περισσότερα σχετικά με τη δημοσιογραφία δεδομένων και τον ανοιχτό κώδικα, βλ. επίσης τα κεφάλαια των Λίον, Μπαάκ, Πιτς και Μουσκάτο σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 4.5, «Εξηγώντας τις Μεθόδους: Υπολογιστικά Φύλλα, Script και Notebook Προγραμματισμού», Κεφάλαιο 7.4, «Οι Δεσμοί της Δημοσιογραφίας Δεδομένων με την Τεχνολογία των Πολιτών» και Κεφάλαιο 7.5, «Πρακτικές Ανοιχτού Κώδικα στη Δημοσιογραφία Δεδομένων»).

Παραδείγματα και οφέλη της ανοιχτότητας

Η *Washington Post* παρουσίασε μια άνευ προηγουμένου επισκόπηση της επιδημίας των συνταγογραφούμενων οπιοειδών στις Ηνωμένες Πολιτείες, ψάχνοντας σε μια βάση δεδομένων σχετικά με τις πωλήσεις εκατομμυρίων παυσίπονων[104]. Κατέστησαν, επίσης, το σύνολο δεδομένων και τη μεθοδολογία δημόσια προσβάσιμα. Αυτό επέτρεψε σε τοπικούς δημοσιογράφους από περισσότερες από 30 πολιτείες να δημοσιεύσουν περισσότερα από 90 άρθρα σχετικά με τον αντίκτυπο αυτής της κρίσης στις κοινότητές τους (Sánchez Díez, 2019)[105].

Δύο δημοσιογράφοι που ασκούν υπολογιστική δημοσιογραφία ανέλυσαν τον αλγόριθμο τιμολόγησης της Uber και αποκάλυψαν ότι η εταιρεία φαίνεται να προσφέρει καλύτερες υπηρεσίες σε περιοχές με περισσότερους λευκούς (Stark & Diakopoulos, 2016). Το θέμα δημοσιεύθηκε από την *Washington Post* και ο κώδικας συλλογής και ανάλυσης δεδομένων που χρησιμοποιήθηκε διατέθηκε ελεύθερα στο GitHub, μια διαδικτυακή πλατφόρμα που βοηθά τους προγραμματιστές να αποθηκεύουν και να διαχειρίζονται τον κώδικά τους[106]. Αυτό σήμαινε ότι, αν ένας αναγνώστης κοιτούσε τη βάση δεδομένων και εντόπιζε κάποιο σφάλμα, μπορούσε να το αναφέρει στους συντάκτες του άρθρου, οι οποίοι με τη σειρά τους μπορούσαν να αποκαταστήσουν το σφάλμα και να διορθώσουν την ιστορία.

Το *Gênero e Número* («Φύλο και Αριθμός»), ένα βραζιλιάνικο ψηφιακό περιοδικό του οποίου είμαι συνιδρύτρια, έτρεξε ένα πρόγραμμα ταξινόμησης περισσότερων από 800.000 ονομάτων οδών, προκειμένου να γίνει κατανοητή η έλλειψη γυναικείας εκπροσώπησης στους δημόσιους χώρους της Βραζιλίας. Αυτό το κάναμε εκτελώντας ένα πρόγραμμα γραμμένο σε Python, για τη διασταύρωση ονομάτων οδών με μια βάση δεδομένων με ονόματα από την Εθνική Στατιστική Υπηρεσία της Βραζιλίας (Mazotte & Justen, 2017). Το ίδιο πρόγραμμα χρησιμοποιήθηκε στη συνέχεια από άλλες πρωτοβουλίες για την ταξινόμηση συνόλων δεδομένων που δεν περιείχαν πληροφορίες σχετικά με το φύλο –όπως λίστες υποψηφίων για εκλογές και δικαστών (Justen, 2019).

Το να δουλεύουμε ανοιχτά και καθιστώντας τα διάφορα σύνολα δεδομένων, εργαλεία, κώδικες, μεθόδους και διαδικασίες διαφανή και διαθέσιμα είναι πιθανό να βοηθήσει τους δημοσιογράφους δεδομένων με διάφορους τρόπους. Πρώτον, μπορεί να τους βοηθήσει να βελτιώσουν την ποιότητα της εργασίας τους. Η καταγραφή των διαδικασιών μπορεί να ενθαρρύνει τους δημοσιογράφους να είναι πιο οργανωμένοι, πιο ακριβείς και λιγότερο επιρρεπείς στο να τους διαφεύγουν λάθη. Μπορεί, επίσης, να ελαφρύνει το βάρος της επεξεργασίας και της αναθεώρησης σύνθετων θεμάτων, επιτρέποντας στους αναγνώστες να επισημαίνουν τα προβλήματα. Δεύτερον, μπορεί να διευρύνει την απήχηση και τον αντίκτυπο. Ένα θέμα που χτίζεται πάνω σε αυτό μπορεί να αποκτήσει διαφορετικές προοπτικές και να εξυπηρετήσει διαφορετικές κοινότητες. Τα έργα μπορούν να αποκτήσουν τη δική τους ζωή, χωρίς να περιορίζονται πλέον από το αρχικό αντικείμενο και τους δημιουργούς τους. Και τρίτον, μπορεί να ενισχύσει τις στοιχειώδεις γνώσεις σχετικά με τα δεδομένα στους δημοσιογράφους και στο ευρύτερο κοινό. Οι βήμα προς βήμα περιγραφές της εργασίας σας σημαίνουν ότι άλλοι μπορούν να ακολουθήσουν και να μάθουν, γεγονός που μπορεί να εμπλουτίσει και να διαφοροποιήσει τα οικοσυστήματα, τις πρακτικές και τις κοινότητες δεδομένων.

Στη λεγόμενη εποχή της «μετα-αλήθειας» υπάρχει, επίσης, η δυνατότητα να αυξηθεί η εμπιστοσύνη του κοινού στα μέσα ενημέρωσης, η οποία έχει φθάσει σε νέο ιστορικό χαμηλό, σύμφωνα με το Βαρόμετρο Εμπιστοσύνης της Edelman για το 2018.[107] Η ανοιχτή εργασία θα μπορούσε να συμβάλλει στην επιβράδυνση ή ακόμη και στην αντιστροφή της συγκεκριμένης τάσης. Η εν λόγω εργασία μπορεί να περιλαμβάνει δημοσιογράφους που μιλούν πιο ανοιχτά για τον τρόπο με τον

104. https://www.washingtonpost.com/national/2019/07/20/opioid-files/
105. https://www.washingtonpost.com/national/2019/08/12/post-released-deas-data-pain-pills-heres-what-local-journalists-are-using-it/
106. https://github.com/comp-journalism/2016-03-wapo-uber
107. https://www.edelman.com/research/2018-edelman-trust-barometer, https://www.edelman.com/sites/g/files/aatuss191/files/2018-10/2018_Edelman_Trust_Barometer_Global_Report_FEB.pdf

οποίο καταλήγουν στα συμπεράσματά τους και παρέχουν πιο λεπτομερείς πρακτικές συμβουλές, προκειμένου να είναι ειλικρινείς σχετικά με τις προκαταλήψεις και τις αβεβαιότητές τους, καθώς και για να επιτρέψουν τη συζήτηση με το κοινό τους[108].

Ως προϋπόθεση, οι πρακτικές και οι κουλτούρες ανοιχτής και διαφανούς εργασίας στη δημοσιογραφία δεδομένων αποτελούν μια συνεχή διαδικασία εξερεύνησης και πειραματισμού. Ακόμα και καθώς προχωράμε στην κατανόηση των πιθανών ωφελειών, πρέπει να σκεφτούμε πότε η διαφάνεια είναι πολύτιμη, ή πότε μπορεί να αποτελέσει λιγότερο σημαντική προτεραιότητα, ή πότε μπορεί να είναι ακόμα και επιζήμια. Για παράδειγμα, μερικές φορές είναι σημαντικό να παραμείνουν εμπιστευτικά τα δεδομένα και οι τεχνικές, προκειμένου να προστατευθεί η ακεραιότητα της ίδιας της έρευνας, όπως συνέβη στην περίπτωση των *Panama Papers*.

Τρόποι ανοιχτής εργασίας

Εάν δεν υπάρχουν κωλύματα (και αυτό θα πρέπει να αναλύεται κατά περίπτωση), τότε μια συνήθης προσέγγιση της διαφάνειας είναι μέσω της μεθοδολογίας, γνωστής και ως το «μποξάκι του σπασίκλα»[109]. Αυτό μπορεί να πάρει διάφορες μορφές και να έχει διαφορετική έκταση, ανάλογα με την πολυπλοκότητα της διαδικασίας και το κοινό στο οποίο απευθύνεται.

Εάν η πρόθεσή σας είναι να απευθυνθείτε σε ένα ευρύτερο κοινό, ένα «box» μέσα στο άρθρο ή ακόμη και μια υποσημείωση με μια συνοπτική εξήγηση των μεθόδων σας μπορεί να είναι αρκετή. Ορισμένες εκδόσεις επιλέγουν να δημοσιεύουν αυτοτελή άρθρα που εξηγούν τον τρόπο με τον οποίο κάλυψαν το θέμα. Σε κάθε περίπτωση, είναι σημαντικό να αποφεύγετε την ορολογία, να εξηγείτε πώς αποκτήθηκαν και χρησιμοποιήθηκαν τα δεδομένα, να διασφαλίζετε ότι οι αναγνώστες δεν θα χάσουν σημαντικούς όρους και να εξηγείτε με τον πιο σαφή και άμεσο τρόπο πώς καταλήξατε στα συμπεράσματά σας. Πολλά μέσα ενημέρωσης που φημίζονται για τη δουλειά τους στη δημοσιογραφία δεδομένων –όπως το *FiveThirtyEight*, η *ProPublica*, οι *New York Times* και οι *Los Angeles Times*– διαθέτουν αποθετήρια σε πλατφόρμες διαμοιρασμού κώδικα, όπως το GitHub. Η ομάδα του *Buzzfeed News* διαθέτει, ακόμη, και ένα ευρετήριο όλων των δεδομένων, αναλύσεων, βιβλιοθηκών, εργαλείων και οδηγών ανοιχτού κώδικα[110]. Δημοσιεύουν όχι μόνο τη μεθοδολογία πίσω από τα ρεπορτάζ τους, αλλά και τα προγράμματα που χρησιμοποιούνται για την εξαγωγή, τον καθαρισμό, την ανάλυση και την παρουσίαση των δεδομένων. Αυτή η πρακτική καθιστά τη δουλειά τους αναπαραγώγιμη (όπως αναλύεται περαιτέρω στο κεφάλαιο του Λίον σε αυτόν τον τόμο)[111] και δίνει τη δυνατότητα στους ενδιαφερόμενους αναγνώστες να εξερευνήσουν οι ίδιοι τα δεδομένα. Όπως κάνουν οι επιστήμονες εδώ και αιώνες, αυτοί οι δημοσιογράφοι καλούν τους συναδέλφους τους να ελέγξουν τη δουλειά τους και να δουν εάν μπορούν να καταλήξουν στα ίδια συμπεράσματα ακολουθώντας τα καταγεγραμμένα βήματα.

Για πολλές αίθουσες σύνταξης δεν είναι απλό να ενσωματώσουν αυτά τα επίπεδα τεκμηρίωσης και συνεργασίας στη δουλειά τους. Εν μέσω περικοπών και συρρίκνωσης των ομάδων, οι δημοσιογράφοι που επιθυμούν να τεκμηριώσουν τις έρευνές τους μπορεί να αποθαρρύνονται από τους οργανισμούς τους. Αυτό μας φέρνει στους περιορισμούς που αντιμετωπίζουν οι δη-

108. Για περισσότερα θέματα σχετικά με την αβεβαιότητα στη δημοσιογραφία δεδομένων, ανατρέξτε στο κεφάλαιο του Άντερσον σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 9.1, «Γενεαλογίες της Δημοσιογραφίας Δεδομένων»).
109. ΣτΕ: Επεξηγηματικό κείμενο (εν προκειμένω, σχετικά με τη μεθοδολογία εργασίας) το οποίο συνοδεύει το δημοσιογραφικό θέμα. Η λέξη «μποξάκι», προερχόμενη από το αγγλικό «box», χρησιμοποιείται στη δημοσιογραφική αργκό –ιδίως στα έντυπα μέσα, όπου γραφιστικά συνηθίζεται ο διαχωρισμός των επεξηγηματικών κειμένων, συνοδευτικών του κυρίως θέματος, να γίνεται με την τοποθέτησή τους σε ένα «κουτί» που καθιστά εμφανή τη διαφοροποίησή τους από τη δημοσιογραφική ιστορία, ακόμη και αν τα δύο κείμενα δημοσιεύονται στην ίδια σελίδα.
110. https://github.com/BuzzFeedNews/everything#guides
111. ΣτΕ: Κεφάλαιο 4.5, «Εξηγώντας τις Μεθόδους: Υπολογιστικά Φύλλα, Script και Notebook Προγραμματισμού»

μοσιογράφοι: πολλοί ειδησεογραφικοί οργανισμοί δίνουν μάχη για να επιβιώσουν, καθώς ο ρόλος τους στον κόσμο και τα επιχειρηματικά τους μοντέλα αλλάζουν. Παρά τις προκλήσεις αυτές, η υιοθέτηση ορισμένων από τις πρακτικές των κοινοτήτων ελεύθερου και ανοιχτού κώδικα μπορεί να αποτελέσει έναν τρόπο για να ξεχωρίσει κανείς –ως δείγμα καινοτομίας και ως τρόπος οικοδόμησης εμπιστοσύνης και σχέσεων με το κοινό, σε έναν ολοένα και πιο σύνθετο, και ταχέως μεταβαλλόμενο, κόσμο.

Παραπομπές

Justen, A. (2019) «Classif icando nomes por gênero usando dados públicos», *Brasil.IO—blog.brasil.io*, 31 Μαΐου. Διαθέσιμο σε: https://blog.brasil.io/2019/05/31/classificando-nomes-por-genero-usando-dados-publicos/ (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Mazotte, N., & Justen, A. (2017) «Como classificamos mais de 800 mil logradouros brasileiros por gênero», *Gênero e Número,* 5 Απριλίου. Διαθέσιμο σε: http://www.generonumero.media/como-classificamos-mais-de-800-mil-logradouros-brasileiros-por-genero/ (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Sánchez Díez, M. (2019) «*The Post* released the DEA's data on pain pills. Here's what local journalists are using it for», *The Washington Post*, 26 Νοεμβρίου. Διαθέσιμο σε: https://www.washingtonpost.com/national/2019/08/12/post-released-deas-data-pain-pills-heres-what-local-journalists-are-using-it/ (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Stallman, R. M. (2002). *Free software, free society: Selected essays of Richard M. Stallman* (J. Gay, επιμ.). Βοστώνη: GNU Press.

Stark, J., & Diakopoulos, N. (2016) «Uber seems to offer better service in areas with more White people. That raises some tough questions», *The Washington Post*, 10 Μαρτίου. Διαθέσιμο σε: https://www.washingtonpost.com/news/wonk/wp/2016/03/10/uber-seems-to-offer-better-service-in-areas-with-more-white-people-that-raises-some-tough-questions/ (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Η συγγραφέας

Η Ναταλία Μαζότε είναι Βραζιλιάνα δημοσιογράφος δεδομένων και επιχειρηματίας. Είναι συνιδρύτρια του νεοσύστατου ειδησεογραφικού οργανισμού *Gênero e Número* και του School of Data Brazil.

4.7
Αξιοποιώντας τους Αλγορίθμους για το Ρεπορτάζ

Jonathan Stray

Περίληψη: Οι εξελιγμένοι αλγόριθμοι ανάλυσης δεδομένων μπορούν να ωφελήσουν σημαντικά το ερευνητικό ρεπορτάζ, αλλά το μεγαλύτερο μέρος της εργασίας είναι η συλλογή και ο καθαρισμός των δεδομένων.

Λέξεις-κλειδιά: αλγόριθμοι, μηχανική μάθηση, υπολογιστική δημοσιογραφία, δημοσιογραφία δεδομένων, ερευνητική δημοσιογραφία, καθαρισμός δεδομένων

Το βρόμικο μυστικό της υπολογιστικής δημοσιογραφίας είναι ότι το «αλγοριθμικό» μέρος μιας ιστορίας δεν είναι αυτό που απαιτεί όλο τον χρόνο και την προσπάθεια.

Μην με παρεξηγήσετε: οι εξελιγμένοι αλγόριθμοι μπορούν να είναι εξαιρετικά χρήσιμοι στο ρεπορτάζ, ιδίως στο ερευνητικό ρεπορτάζ. Η μηχανική μάθηση (το «machine learning», όπως λέμε, δηλαδή η εκπαίδευση υπολογιστών για την εύρεση μοτίβων) έχει χρησιμοποιηθεί για τον εντοπισμό εγγράφων-κλειδιών σε τεράστιους όγκους δεδομένων. Η επεξεργασία φυσικής γλώσσας («natural language processing» ή, κοινώς, NLP –δηλαδή, η εκπαίδευση υπολογιστών για την κατανόηση της γλώσσας) μπορεί να εξάγει ονόματα ανθρώπων και εταιρειών από έγγραφα, δίνοντας στους δημοσιογράφους τη δυνατότητα να καταλάβουν ποιοι εμπλέκονται σε μια ιστορία. Και οι δημοσιογράφοι έχουν χρησιμοποιήσει ποικίλες στατιστικές αναλύσεις για να εντοπίσουν παραβάσεις ή μεροληψία.

Ωστόσο, στην πραγματικότητα, η εκτέλεση ενός αλγορίθμου είναι το εύκολο μέρος. Η εξασφάλιση των δεδομένων, ο καθαρισμός τους και η παρακολούθηση των αλγοριθμικών στοιχείων είναι το δύσκολο μέρος.

Για να γίνει αυτό κατανοητό, ας πάρουμε για παράδειγμα μια επιτυχία της μηχανικής μάθησης στην ερευνητική δημοσιογραφία, το αξιοσημείωτο θέμα της εφημερίδας *The Atlanta Journal-Constitution* για τη σεξουαλική κακοποίηση από γιατρούς, *License to Betray* (Teegardin κ.α., 2016). Οι δημοσιογράφοι ανέλυσαν περισσότερα από 100.000 αρχεία πειθαρχικού ελέγχου γιατρών από κάθε πολιτεία των ΗΠΑ και εντόπισαν 2.400 περιπτώσεις όπου επετράπη σε γιατρούς, που είχαν κακοποιήσει σεξουαλικά ασθενείς τους, να συνεχίσουν να ασκούν το επάγγελμα. Αντί να διαβάσουν το κάθε αρχείο, αρχικά μείωσαν δραστικά τον όγκο των εγγράφων, εφαρμόζοντας μηχανική μάθηση, για να βρουν τις αναφορές που ήταν πιθανό να αφορούν σε σεξουαλική κακοποίηση. Κατάφεραν να μειώσουν τον όγκο τους πάνω από 10 φορές, σε μόλις 6.000 έγγραφα, τα οποία στη συνέχεια διάβασαν και εξέτασαν με το χέρι.

Αυτό δεν θα μπορούσε να γίνει θέμα εθνικής εμβέλειας χωρίς τη μηχανική μάθηση, σύμφωνα με τον δημοσιογράφο Τζεφ Ερνστχάουζεν. «Ίσως θα μπορούσαμε να το είχαμε κάνει τοπικό θέμα», είπε αργότερα (Diakopoulos, 2019).

Αυτή είναι η μεγαλύτερη νίκη των αλγορίθμων στη δημοσιογραφία που έχουμε δει ποτέ, και αυτή η τεχνική θα μπορούσε να χρησιμοποιηθεί πολύ ευρύτερα. Ωστόσο, η ίδια η μηχανική μάθηση δεν είναι το δύσκολο κομμάτι. Η μέθοδος που χρησιμοποίησε ο Ερνστχάουζεν, η «λο-

γιστική παλινδρόμηση» («logistic regression»), είναι μια τυπική στατιστική προσέγγιση για την ταξινόμηση εγγράφων με βάση τις λέξεις που περιέχουν. Μπορεί να υλοποιηθεί με μόλις καμιά δεκαριά γραμμές κώδικα σε Python και υπάρχουν πολλά καλά tutorial στο Διαδίκτυο.

Για τα περισσότερα θέματα, το μεγαλύτερο μέρος της δουλειάς είναι το στήσιμο των πραγμάτων και, στη συνέχεια, η εκμετάλλευση των αποτελεσμάτων. Τα δεδομένα πρέπει να συλλεχθούν, να καθαριστούν, να μορφοποιηθούν, να φορτωθούν, να ελεγχθούν και να διορθωθούν –μια ατελείωτη προετοιμασία. Και τα αποτελέσματα της αλγοριθμικής ανάλυσης είναι συχνά μόνο ενδείξεις ή υποδείξεις, οι οποίες γίνονται θέμα μόνο μετά από πολύ και χειροκίνητο ρεπορτάζ, συχνά από ομάδες δημοσιογράφων που χρειάζονται εργαλεία συνεργασίας και όχι εργαλεία ανάλυσης. Αυτό είναι το πιο άχαρο μέρος της εργασίας με δεδομένα, και αυτός είναι ο λόγος που δεν το διδάσκουμε πολύ καλά ούτε μιλάμε πολύ για αυτό. Ωστόσο, αυτή η προετοιμασία και η παρακολούθηση είναι που καταλαμβάνουν το μεγαλύτερο μέρος του χρόνου και της προσπάθειας για ένα θέμα που βασίζεται σε δεδομένα.

Για το *License to Betray*, η συγκέντρωση των δεδομένων από μόνη της ήταν μια τεράστια πρόκληση. Δεν υπάρχει εθνική βάση δεδομένων για το πειθαρχικό μητρώο των γιατρών, παρά μόνο μια σειρά από βάσεις δεδομένων σε επίπεδο πολιτείας. Πολλές από αυτές τις βάσεις δεδομένων δεν περιέχουν ένα πεδίο που να υποδεικνύει τον λόγο για τον οποίο ένας γιατρός τιμωρήθηκε πειθαρχικά. Όπου υπάρχει πεδίο, συχνά δεν κωδικοποιεί αξιόπιστα τη σεξουαλική κακοποίηση. Αρχικά, η ομάδα προσπάθησε να λάβει τις πληροφορίες μέσω αιτημάτων στο πλαίσιο της ελευθερίας της πληροφόρησης («freedom of information», FOI). Αυτό αποδείχθηκε απαγορευτικά ακριβό, με ορισμένες πολιτείες να ζητούν χιλιάδες δολάρια για να διαθέσουν τα δεδομένα. Έτσι, η ομάδα στράφηκε στην ανεύρεση εγγράφων από τις ιστοσελίδες των πολιτειακών ιατρικών συμβουλίων (Ernsthausen, 2017). Αυτά τα έγγραφα έπρεπε να μετατραπούν σε κείμενο με τη χρήση οπτικής αναγνώρισης χαρακτήρων («optical character recognition», OCR) και να φορτωθούν σε μια προσαρμοσμένη διαδικτυακή εφαρμογή για συνεργατική επισημείωση («tagging») και αναθεώρηση.

Στη συνέχεια, οι δημοσιογράφοι έπρεπε να χαρακτηρίσουν χειροκίνητα αρκετές εκατοντάδες έγγραφα για την παραγωγή δεδομένων εκπαίδευσης («training data»)[112]. Αφού η μηχανική μάθηση κατέταξε τα υπόλοιπα 100.000, χρειάστηκαν αρκετοί μήνες για να διαβαστούν χειροκίνητα τα 6.000 έγγραφα που είχε υπολογιστεί ότι αφορούσαν σεξουαλική κακοποίηση, καθώς και χιλιάδες άλλα έγγραφα τα οποία περιείχαν λέξεις-κλειδιά που επιλέχθηκαν με το χέρι. Και στη συνέχεια, φυσικά, υπήρχε και το υπόλοιπο ρεπορτάζ, όπως η διερεύνηση εκατοντάδων συγκεκριμένων περιπτώσεων, προκειμένου να πάρει η ιστορία σάρκα και οστά. Αυτό βασίστηκε σε άλλες πηγές, όπως προηγούμενες ειδησεογραφικές δημοσιεύσεις και, φυσικά, προσωπικές συνεντεύξεις με τους εμπλεκόμενους.

Η χρήση ενός αλγορίθμου (μηχανική μάθηση) ήταν ένα βασικό, κρίσιμο κομμάτι της έρευνας, αλλά ήταν μόνο ένα ελάχιστο μέρος του χρόνου και της προσπάθειας που δαπανήθηκε. Οι έρευνες των επιστημόνων δεδομένων δείχνουν σταθερά ότι το μεγαλύτερο μέρος της εργασίας τους είναι η «επεξεργασία» και ο καθαρισμός των δεδομένων –συχνά σε ποσοστό έως και 80%–, και η δημοσιογραφία δεν διαφέρει (Lohr, 2014).

Οι αλγόριθμοι συχνά θεωρούνται ένα είδος μαγικού συστατικού. Μπορεί να φαίνονται πολύπλοκοι ή αδιαφανείς, αλλά είναι αναμφισβήτητα ισχυροί. Είναι πολύ πιο διασκεδαστικό να συζητάς για τη μαγεία αυτή παρά να κάνεις την τετριμμένη εργασία της προετοιμασίας δεδομένων ή της παρακολούθησης μιας μακράς λίστας με στοιχεία. Στους τεχνολόγους αρέσει να διαφημίζουν την τεχνολογία τους και όχι την εξίσου σημαντική εργασία που συμβαίνει γύρω

112. ΣτΕ: Δεδομένα με τα οποία τροφοδοτείται και εκπαιδεύεται το εκάστοτε αλγοριθμικό μοντέλο που, στη συνέχεια, «αποφαίνεται», αντί του ανθρώπου, για τα υπόλοιπα δεδομένα, τα λεγόμενα «δεδομένα δοκιμής» («testing data»).

από αυτήν –και αυτή η προτίμηση στα νέα και εξελιγμένα εργαλεία μεταφέρεται μερικές φορές και στη δημοσιογραφία. Οπωσδήποτε θα πρέπει να διδάσκουμε και να αξιοποιούμε τις τεχνολογικές εξελίξεις, αλλά η πρωταρχική μας ευθύνη είναι να ασκήσουμε τη δημοσιογραφία, και αυτό σημαίνει ότι πρέπει να ασχοληθούμε και με το υπόλοιπο κομμάτι της ροής δεδομένων.

Σε γενικές γραμμές, υποτιμούμε τα εργαλεία που χρησιμοποιούνται για την προετοιμασία των δεδομένων. Το OpenRefine είναι μια παλιά μας αγάπη για όλα τα είδη εργασιών που αφορούν καθαρισμό δεδομένων. Το Dedupe.io είναι μηχανική μάθηση που εφαρμόζεται στο πρόβλημα της συγχώνευσης σχεδόν διπλότυπων ονομάτων σε μια βάση δεδομένων. Οι κλασικές μέθοδοι καθαρισμού κειμένου, όπως τα «regular expressions»[113], θα πρέπει να αποτελούν μέρος της εκπαίδευσης κάθε δημοσιογράφου δεδομένων. Στο πνεύμα αυτό, το τρέχον έργο μου, το Workbench, επικεντρώνεται στη χρονοβόρα αλλά κυρίως αόρατη εργασία της προετοιμασίας των δεδομένων για δημοσίευση –σε όλα όσα συμβαίνουν πριν από τον «αλγόριθμο». Ως εκ τούτου, στοχεύει να καταστήσει την όλη διαδικασία πιο συνεργατική, προκειμένου οι δημοσιογράφοι να μπορούν να εργάζονται από κοινού σε μεγάλα έργα δεδομένων και να μαθαίνουν ο ένας από το έργο του άλλου, μαζί με τις μηχανές.

Οι αλγόριθμοι είναι σημαντικοί για το ρεπορτάζ, αλλά για να λειτουργήσουν, πρέπει να μιλήσουμε για όλα τα άλλα σκέλη της δημοσιογραφίας με βάση τα δεδομένα. Πρέπει να ενεργοποιήσουμε ολόκληρη τη ροή εργασιών, όχι μόνο τα ιδιαίτερα γοητευτικά, υψηλής τεχνολογίας μέρη.

Παραπομπές

Diakopoulos, N. (2019) *Automating the news: How algorithms are rewriting the media*. Κέιμπριτζ, Μασαχουσέτη: Harvard University Press.

Ernsthausen, J. (2017) «Doctors and sex abuse», *NICAR 2017*. Τζάκσονβιλ, Φλόριντα, 2-5 Μαρτίου. Διαθέσιμο σε: https://docs.google.com/presentation/d/1keGeDk_wpBPQgUOOhbRarPPFbyCculTObGLeAhOMmEM/edit#slide=id.p (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Lohr, S. (2014) «For big-data scientists, "janitor work" is key hurdle to insights», *The New York Times*, 17 Αυγούστου. Διαθέσιμο σε: https://www.nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-is-janitor-work.html (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Teegardin, C., Robbins, D., Ernsthausen, J., & Hart, A (2016) «License to betray», *The Atlanta Journal-Constitution*, Doctors & Sex Abuse, 5 Ιουλίου. Διαθέσιμο σε: http://doctors.ajc.com/doctors_sex_abuse/?ecmp=doctorssexabuse_microsite_nav (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Ο συγγραφέας

Ο Τζόναθαν Στρέι, Επιστημονικός Συνεργάτης του *Partnership on AI*, δίδασκε στο παρελθόν στο διπλό μεταπτυχιακό πρόγραμμα σπουδών στην Επιστήμη των Υπολογιστών και στη Δημοσιογραφία στο Πανεπιστήμιο Columbia και έχει δημιουργήσει λογισμικό για την ερευνητική δημοσιογραφία.

113. ΣτΕ: Για έναν πρώτο οδηγό στην εκμάθηση και στη χρήση «regex», όπως συχνά λένε οι δημοσιογράφοι δεδομένων, μπορείτε να δείτε το regular-expressions.info.

4.8

Δημοσιογραφία με Μηχανές; Από την Υπολογιστική Σκέψη στην Κατανεμημένη Νόηση

Eddy Borges-Rey

Περίληψη: Αυτό το κεφάλαιο εξετάζει τους τρόπους με τους οποίους η αυτοματοποίηση των ειδήσεων διαδίδει ολοένα και περισσότερο τη δημοσιογραφική γνώση και συλλογιστική.

Λέξεις-κλειδιά: αυτοματοποιημένη δημοσιογραφία, κατανεμημένη νόηση, υπολογιστική σκέψη, εκτεταμένος νους, βάσεις δεδομένων, μηχανική μάθηση

Φανταστείτε ότι είστε δημοσιογράφος στο όχι και τόσο μακρινό μέλλον. Δουλεύετε πάνω σε ένα θέμα και, προκειμένου να αποκτήσετε τις πληροφορίες που ψάχνετε, ζητάτε από τον διαλογικό σας πράκτορα (τον οποίο αποκαλείτε χαϊδευτικά Τουίκι) να συνδυάσει πάνω από 15 ανωνυμοποιημένες βάσεις δεδομένων. Δεδομένου του μεγέθους και της πολυπλοκότητας των συγχωνευμένων συνόλων δεδομένων, το λογισμικό οπτικοποίησης είναι πολύ υποτυπώδες, για να απομονώσει τις ανωμαλίες που αναζητάτε. Έτσι, χρησιμοποιώντας το εγκεφαλικό σας εμφύτευμα, συνδέεστε στο σύστημα και πλοηγήστε εύκολα στην εξαγωγή των συνόλων δεδομένων. Αν και, μεμονωμένα, το κάθε επεξεργασμένο σύνολο δεδομένων προστατεύει αποτελεσματικά την ταυτότητα και τα προσωπικά δεδομένα των ατόμων που αναφέρονται, όταν αυτά συνδυαστούν, είστε σε θέση να συμπεράνετε την ταυτότητα ορισμένων ατόμων υψηλού προφίλ και να τοποθετήσετε τα προσωπικά τους δεδομένα σε ένα πλαίσιο. Αντιλαμβανόμενοι τις πιθανές νομικές συνέπειες της αποκάλυψης των ονομάτων και των δεδομένων που συνδέονται με αυτά, ζητάτε από τον Τουίκι να «τρέξει» ένα νευρωνικό δίκτυο, για να καθορίσει κατά πόσο η αποκάλυψη αυτών των πληροφοριών έχει ηθικές ή νομικές συνέπειες. Το δίκτυο εκτελεί έναν v+ αριθμό προσομοιώσεων όπου εικονικοί δημοσιογράφοι λαμβάνουν αποφάσεις με βάση διάφορους κώδικες δεοντολογίας και ρυθμιστικά πλαίσια. Ενώ αυτό εκτελείται στο παρασκήνιο, καταφέρνετε να απομονώσετε μερικές ακραίες τιμές και να εντοπίσετε ορισμένες ενδιαφέρουσες τάσεις. Δεδομένου ότι θέλετε να βεβαιωθείτε πως οι ανωμαλίες έχουν κάτι να προσθέσουν στην ιστορία και δεν είναι απλά σφάλματα, ζητάτε από τον Τουίκι να ελέγξει τα ιστορικά αρχεία, για να δει αν οι ακραίες τιμές συμπίπτουν με κάποιο σημαντικό ιστορικό γεγονός και κατόπιν να εκτελέσει ένα προγνωστικό μοντέλο για να υπολογίσει την πιθανότητα οι τάσεις που εντοπίστηκαν να εξακολουθήσουν να υπάρχουν στο άμεσο μέλλον, πυροδοτώντας έτσι ανησυχητικές συνέπειες.

Αυτή η σύντομη, φανταστική εισαγωγή βασίζεται σε μια συναρπαστική συζήτηση που είχα με την πρώην δημοσιογράφο δεδομένων των *Times*, Νίκολα Χιουτζ πριν από μερικά χρόνια. Αν και η σκηνή που περιγράφεται θα μπορούσε κάλλιστα να έχει βγει από το *The Minority Report* του Φίλιπ Κ. Ντικ, στην πραγματικότητα αναφέρεται σε μια σειρά εργαλείων και τεχνικών που είτε είναι ήδη διαθέσιμες και χρησιμοποιούνται ευρέως, είτε βρίσκονται σε ταχεία ανάπτυξη. Το πιο σημαντικό είναι ότι αναφέρεται, επίσης, σε ένα είδος ροής δημοσιογραφικής εργασίας και επαγγελματικής νοοτροπίας που αναπτύσσεται στις αίθουσες σύνταξης, σε έναν κόσμο όπου οι δημοσιογράφοι ασχολούνται ολοένα και περισσότερο με τα δεδομένα και οι υπολογισμοί είναι πλέον απαραίτητοι.

Αυτές οι πρόσφατες αλλαγές αντικατοπτρίζουν τον τρόπο με τον οποίο ιστορικά, κάθε φορά που μια σημαντική τεχνολογική καινοτομία εισάγεται στη ροή εργασίας της παραγωγής ειδήσεων, η ίδια η ειδησεογραφία όχι μόνο διαταράσσεται και, κατά συνέπεια, μετασχηματίζεται, αλλά και οι διαδικασίες σκέψης των δημοσιογράφων και τα επαγγελματικά ιδεώδη εργασίας τροποποιούνται αναπόφευκτα.

Σήμερα, καθώς περνάμε από την εποχή των μεγάλων δεδομένων («big data») στην εποχή της τεχνητής νοημοσύνης («artificial intelligence», AI) και της αυτοματοποίησης, οι αρχές και οι πρακτικές εργασίας που προέρχονται από την πληροφορική και την επιστήμη των δεδομένων γίνονται όλο και πιο διαδεδομένες στη δημοσιογραφία. Όπως το θέτει η Έμιλι Μπελ, ιδρυτική διευθύντρια του Tow Center for Digital Journalism στο Πανεπιστήμιο Columbia:

> Κάθε εταιρεία, σε κάθε τομέα, και κάθε οργανισμός, είτε πρόκειται για επιχειρήσεις είτε για το δημόσιο, θα πρέπει να σκεφτεί πώς θα προσανατολιστεί εκ νέου γύρω από την τεχνητή νοημοσύνη με τον ίδιο ακριβώς τρόπο που πριν από 20 χρόνια έπρεπε να σκεφτεί πώς θα προσανατολιστεί εκ νέου γύρω από τις τεχνολογίες του Διαδικτύου (Bell, προσωπική επικοινωνία, 7 Σεπτεμβρίου 2017).

Σε αυτό το πλαίσιο, το παρόν κεφάλαιο εξετάζει τους τρόπους με τους οποίους οι δημοσιογράφοι που εργάζονται στενά με δεδομένα και αυτοματοποιημένες διαδικασίες εσωτερικεύουν μια σειρά από αρχές της πληροφορικής, οι οποίες αφενός ενισχύουν τις δημοσιογραφικές τους ικανότητες και αφετέρου έχουν αρχίσει να τροποποιούν τον ίδιο τον ακρογωνιαίο λίθο των δημοσιογραφικών τους προσεγγίσεων και ιδανικών. Ως εκ τούτου, το κεφάλαιο διερευνά μια σειρά από θεωρητικές έννοιες που θα μπορούσαν να χρησιμεύσουν ως πλαίσιο για να οραματιστεί κανείς τη δημοσιογραφική νόηση σε ένα περιβάλλον διεισδυτικής υπολογιστικής. Υιοθετώ την έννοια της εκτεταμένης νόησης («extended cognition»), για να ενθαρρύνω την περαιτέρω συζήτηση σχετικά με τους τρόπους με τους οποίους η δημοσιογραφική νόηση εξαρτάται σήμερα από τις μηχανές που χρησιμοποιούνται για την κάλυψη των ειδήσεων (και, συνεπώς, κατανέμεται σε αυτές). Μέσα από αυτή τη συζήτηση, ελπίζω να ενθαρρύνω μελλοντικές εργασίες για τη διερεύνηση του ρόλου των υπολογισμών σε δημοσιογραφικά συμφραζόμενα, συμπεριλαμβανομένης της εμπειρικής εργασίας για τη δοκιμή και τον περαιτέρω προσδιορισμό της έννοιας της κατανεμημένης δημοσιογραφικής νόησης. Αυτό το σκεπτικό θα μπορούσε να είναι ιδιαίτερα χρήσιμο για τους επαγγελματίες δημοσιογράφους που θέλουν να γνωρίζουν και να ασχοληθούν με τις αλλαγές που είναι πιθανό να υποστεί η δημοσιογραφία, αν η πληροφορική και η αυτοματοποίηση επικρατήσουν παντού στην παραγωγή ειδήσεων.

Υπολογιστική σκέψη

Σε μια απόπειρα να ανιχνεύσουν την ιστορική σημασία της έννοιας του υπολογισμού, οι Ντένινγκ και Μαρτέλ προτείνουν ότι «[ως] υπολογισμός θεωρήθηκαν τα μηχανικά βήματα που ακολουθούνται για την αξιολόγηση μαθηματικών συναρτήσεων [και] υπολογιστές ήταν οι άνθρωποι που έκαναν [τους] υπολογισμούς». Ωστόσο, τη δεκαετία του 1980, η έννοια συνδέθηκε συχνότερα με έναν νέο τρόπο άσκησης της επιστήμης, μετατοπίζοντας έτσι την έμφαση από τις μηχανές στην επεξεργασία πληροφοριών (Denning & Martell, 2015).

Αυτή η μετατόπιση της έμφασης είναι κρίσιμη για το επιχείρημά μου, καθώς ευθυγραμμίζει τους απώτερους στόχους του ειδησεογραφικού ρεπορτάζ και του υπολογισμού: Σε πολύ γενικές γραμμές, η δουλειά του δημοσιογράφου συνίσταται στον εξορθολογισμό της ροής των πληροφοριών, στην επιμέλειά τους και στο «πακετάρισμά» τους σε μορφή που να είναι εύπεπτη για το κοινό. Εδώ, θα έλεγα ότι η διείσδυση της υπολογιστικής νοοτροπίας στην ειδησεογραφία οφείλεται εν μέρει στις ομοιότητες που υπάρχουν μεταξύ των δύο επαγγελματικών πρακτικών.

Τόσο η πληροφορική όσο και η δημοσιογραφία είναι τυποποιημένες εργασίες, αφορούν στην

επίλυση προβλημάτων και απαιτούν συντακτική γνώση. Η Γουίνγκ παρατηρεί ότι «[λ]ειτουργικά, η πληροφορική προσπαθεί να απαντήσει στο ερώτημα "πώς θα έκανα έναν υπολογιστή να λύσει αυτό το πρόβλημα;"» (Wing, 2008, σελ. 3719), και αυτό απαιτεί σχετικά υψηλό επίπεδο υπολογιστικής σκέψης. Καθώς οι υπολογισμοί γίνονται πλέον ο κανόνας στις αίθουσες σύνταξης, όλο και περισσότεροι δημοσιογράφοι χρησιμοποιούν την υπολογιστική σκέψη, για να προσεγγίσουν ιστορίες με δεδομένα. Ο Μπράντσο, για παράδειγμα, υποστηρίζει ότι η υπολογιστική σκέψη «βρίσκεται στο επίκεντρο της δουλειάς ενός δημοσιογράφου δεδομένων», επιτρέποντάς του «να επιλύει τα προβλήματα που αποτελούν τόσο μεγάλο μέρος της σύγχρονης δημοσιογραφίας, και να είναι σε θέση να το κάνει με την ταχύτητα και την ακρίβεια που απαιτεί η ειδησεογραφική διαδικασία» (Bradshaw, 2017).

Η υπολογιστική σκέψη είναι η αναστοχαστική διαδικασία μέσω της οποίας πραγματοποιείται ένα σύνολο προγραμματιστικών βημάτων για την επίλυση ενός προβλήματος (Bradshaw, 2017· Wing, 2006, 2008). Η Γουίνγκ υποστηρίζει ότι «η ουσία της υπολογιστικής σκέψης είναι η αφαίρεση» και ότι «στην πληροφορική, αφαιρούμε έννοιες πέρα από τις φυσικές διαστάσεις του χρόνου και του χώρου» (Wing, 2008, σ. 3717) για να επιλύσουμε προβλήματα, να σχεδιάσουμε συστήματα και να κατανοήσουμε την ανθρώπινη συμπεριφορά (Wing, 2006). Η συγγραφέας διατυπώνει αυτήν την τοποθέτηση σε απάντηση της ερώτησης «πώς να βάλω έναν υπολογιστή να λύσει αυτό το πρόβλημα;». Οι επαγγελματίες της πληροφορικής πρέπει να βρουν τις σωστές αφαιρέσεις (Wing, 2008, σ. 3717), οι οποίες είναι κατάλληλες για τον σχεδιασμό και την εφαρμογή ενός προγραμματικού σχεδίου για την επίλυση του εκάστοτε προβλήματος.

Από την εισαγωγή των τεχνολογιών αυτοματισμού στις αίθουσες σύνταξης, οι δημοσιογράφοι που συνεργάζονται με επαγγελματίες της πληροφορικής αντιμετωπίζουν ένα παρόμοιο ερώτημα: «πώς θα κάνω έναν υπολογιστή να ερευνήσει ή να γράψει μια είδηση, σύμφωνα με τα ανθρώπινα πρότυπα;». Η Τζίνιλντ προτείνει ότι η ενσωμάτωση της υπολογιστικής σκέψης στην επαγγελματική δημοσιογραφία αμφισβητεί το «θεμελιώδες σύστημα σκέψης της δημοσιογραφίας από την περιγραφική εξιστόρηση σε αφηρημένους συλλογισμούς, αυτόνομη έρευνα και οπτικοποίηση ποσοτικών γεγονότων», το οποίο εξοπλίζει τους δημοσιογράφους με «συμπληρωματικές, λογικές και αλγοριθμικές δεξιότητες, νοοτροπίες και αξίες» (Gynnild, 2014).

Φυσικά, αυτό δεν σημαίνει ότι η ιδέα της «υπολογιστικής» αφαίρεσης είναι καινούρια για τους δημοσιογράφους. Στην πραγματικότητα, οι δημοσιογράφοι που ασχολούνται με θέματα όπως τα οικονομικά, οι επιχειρήσεις, η αγορά ακινήτων ή η εκπαίδευση ασκούν την αφαίρεση σε καθημερινή βάση, για να κατανοήσουν πολύπλοκες δυναμικές όπως οι επιδόσεις της αγοράς, οι αποδόσεις των μετοχών, η καθαρή αξία των νοικοκυριών κλπ. Και είναι ενδιαφέρον ότι, όπως παρατηρεί ο Μάιλς (2019), σε αντίθεση με τις προσδοκίες ότι η αυτοματοποίηση θα απελευθέρωνε τους δημοσιογράφους από επαχθή καθήκοντα, εισήγαγε μια σειρά από νέες επιμελητικές δραστηριότητες που δεν εκτελούνταν προηγουμένως από τους δημοσιογράφους. Για παράδειγμα, εξηγεί ότι η εισαγωγή της αναγνώρισης εικόνων στη ροή εργασίας του *Associated Press* οδήγησε τους δημοσιογράφους και τους φωτογράφους να ασχοληθούν με εργασίες που παραδοσιακά συνδέονται με τη μηχανική μάθηση, όπως η επισημείωση των δεδομένων εκπαίδευσης («training data»), η αξιολόγηση των αποτελεσμάτων των δοκιμών, η διόρθωση των μεταδεδομένων ή η δημιουργία ορισμών για έννοιες (Myles, 2019).

Γνωστική προβολή και εκτεταμένη δημιουργικότητα

Μέχρι στιγμής, υποστήριξα ότι οι δημοσιογράφοι που, ως μέρος της δουλειάς τους, πρέπει να ασχοληθούν με τα υπολογιστικά προβλήματα που φέρνει η αυτοματοποίηση των ειδήσεων, βλέπουν τη ροή εργασίας τους και τις επιμελητικές τους ευθύνες να μεταλλάσσονται. Η *Wall Street Journal*, για παράδειγμα, δημοσίευσε πρόσφατα αγγελίες για θέσεις όπως Δημοσιογράφος Μηχανικής Μάθησης («Machine Learning Journalist»), Συντάκτης Αυτοματισμού

(«Automation Editor») και Συντάκτης Αναδυόμενων Διαδικασιών («Emerging Processes Editor») –όλες τους σχετίζονται με την εξάπλωση της τεχνητής νοημοσύνης και της αυτοματοποίησης. Ως αποτέλεσμα αυτής της επέκτασης των υποδομών και της επακόλουθης διαφοροποίησης των συντακτικών αρμοδιοτήτων που επιφέρουν, οι δημοσιογράφοι συχνά βρίσκονται να θέτουν ερωτήματα που τους τοποθετούν στη θέση μιας μηχανής η οποία πρέπει να σκεφτεί και να λειτουργήσει ως δημοσιογράφος –ένα ενδιαφέρον παράδοξο, το οποίο συνεπάγεται εξίσου ενδιαφέρουσες προκλήσεις. Αυτή η ιδέα της προβολής, πιστεύω, γίνεται ολοένα και πιο διαδεδομένη στην αυτοματοποίηση των ειδήσεων. Πάρτε, για παράδειγμα, την πεμπτουσία της δημοσιογραφικής δραστηριότητας: τη σύνταξη μιας είδησης. Εάν αποδομήσουμε τη διαδικασία, σε γενικές γραμμές, οι δημοσιογράφοι πρέπει να χρησιμοποιήσουν τη δημιουργικότητά τους, για να συνθέσουν έναν απολογισμό των γεγονότων που θα προσελκύει ή/και θα ενημερώνει το κοινό. Το ερώτημα, λοιπόν, είναι: πώς μπορώ να κάνω μια μηχανή να γράφει ειδήσεις που να διαβάζονται σαν να τις έχει γράψει ένας άνθρωπος-ρεπόρτερ; Δημοσιογράφοι και τεχνολόγοι έχουν συνεργαστεί τα τελευταία πέντε χρόνια, για να προβάλλουν τους εαυτούς τους, σε μια προσπάθεια να επιλύσουν αυτό το ζήτημα. Ένα καλό παράδειγμα, σε αυτό το μέτωπο, είναι η εφαρμογή τεχνολογιών παραγωγής φυσικής γλώσσας («natural language generation», NLG) για την αυτοματοποίηση της παραγωγής ειδήσεων. Ωστόσο, σε αντίθεση με ό,τι θα μπορούσαμε να περιμένουμε, η διαδικασία εξακολουθεί να περιλαμβάνει ανθρώπους-ρεπόρτερ που γράφουν πρότυπα ειδησεογραφικών ιστοριών, τα οποία περιέχουν κενά που στη συνέχεια συμπληρώνονται από το λογισμικό αυτοματισμού με τη χρήση μιας βάσης δεδομένων. Η διαδικασία αυτή, η οποία λειτουργεί με αρκετή επιτυχία σε ειδησεογραφικούς οργανισμούς όπως το *Associated Press* και το RADAR, μια συνεργασία μεταξύ του Press Association και του Urbs Media, αποσκοπεί στην αύξηση της ταχύτητας και της κλίμακας της παραγωγής ειδήσεων σε τομείς όπως ο αθλητισμός, τα επιχειρηματικά έσοδα και οι τοπικές ειδήσεις.

Η δημιουργικότητα σε αυτόν τον χώρο παίρνει μια νέα μορφή, στην οποία οι προγραμματιστές-δημοσιογράφοι έπρεπε να επανεξετάσουν την εξιστόρηση ως μια μηχανή που αποκωδικοποιεί και επανακωδικοποιεί τη διαδικασία σύνταξης ειδήσεων. Αντί να εντοπίσουμε ποια συνέντευξη θα τεκμηρίωνε καλύτερα ένα επιχείρημα ή ποιες λέξεις θα αποτελούσαν έναν ισχυρότερο τίτλο, ο στόχος έχει μετατοπιστεί: να επιλέξουμε ποια διαμόρφωση των υπό όρους δηλώσεων θα ήταν πιο αποτελεσματική, ώστε να κάνει το αυτοματοποιημένο σύστημα να αποφασίσει ποιος τίτλος θα απευθυνόταν αποτελεσματικότερα στο κοινό του ειδησεογραφικού οργανισμού όπου λειτουργεί. Ακολουθώντας τις αρχές της αλληλεπίδρασης ανθρώπου-υπολογιστή («human-computer interaction», HCI) και του χρηστοκεντρικού σχεδιασμού («user experience design» ή «UX design»), οι προγραμματιστές-δημοσιογράφοι πρέπει να προβλέψουν τους τρόπους με τους οποίους οι χρήστες θέλουν να εμπλακούν με τις αυτοματοποιημένες πληροφοριακές εμπειρίες, τους πιθανούς τρόπους με τους οποίους θα περιηγηθούν στα διάφορα επίπεδα πληροφοριών και τα όρια της είδησης. Ο Γουίλερ, διατυπώνοντας την έννοια της εκτεταμένης δημιουργικότητας, εξηγεί ότι υπάρχουν περιπτώσεις διανοητικής δημιουργίας στις οποίες «τα υλικά οχήματα που υλοποιούν τη σκέψη και τις σχετικές σκέψεις είναι χωρικά κατανεμημένα στον εγκέφαλο, στο σώμα και στον κόσμο» (Wheeler, 2018). Η έννοια της εκτεταμένης δημιουργικότητας, λοιπόν, αποδίδει καλά ως πλαίσιο για να εξηγήσει την ιδέα ότι το μυαλό ενός δημοσιογράφου που εργάζεται με δεδομένα και αυτοματισμούς λειτουργεί πλέον σε στενή σύνδεση με μια σειρά από αυτοματισμούς, που εκτείνονται σε μια σειρά από βιβλιοθήκες της Python, Jupyter Notebook, σύνολα δεδομένων, εργαλεία ανάλυσης δεδομένων και διαδικτυακές πλατφόρμες. Αυτή η δυναμική φέρνει κατά συνέπεια μια σειρά πρόσθετων προκλήσεων που αξίζουν την προσοχή μας.

Για παράδειγμα, η Μεβάν Μπαμπακάρ, επικεφαλής του αυτοματοποιημένου ελέγχου γεγονότων στο Full Fact, εξηγεί ότι μία από τις προκλήσεις που αντιμετωπίζουν εκεί με τον αυτοματοποιημένο έλεγχο γεγονότων είναι το πλαίσιο. Χρησιμοποιεί ως παράδειγμα τον ισχυρισμό της πρώην πρωθυπουργού του Ηνωμένου Βασιλείου, Τερέζα Μέι, ότι η κυβέρνησή της διέθεσε περισσότερους πόρους στο Εθνικό Σύστημα Υγείας («National Health Service», NHS) από ό,τι

είχε υποσχεθεί το αντιπολιτευόμενο Εργατικό Κόμμα στο μανιφέστο του. Και παρόλο που ο ισχυρισμός ελέγχθηκε ως ακριβής, για να έχει νόημα και να είναι χρήσιμος στο κοινό, πρέπει να γίνει κατανοητός μέσα σε ένα ευρύτερο πλαίσιο: η χρηματοδότηση δεν ήταν αρκετή για την αποτελεσματική απόδοση του NHS (Babakar, προσωπική επικοινωνία, 16 Αυγούστου 2018). Ως εκ τούτου, καθώς τα αυτοματοποιημένα συστήματα δεν είναι ακόμη ικανά να κάνουν τέτοιες συνδέσεις μεταξύ των πηγών πληροφοριών, η Μπαμπακάρ και η ομάδα της πρέπει να καταφύγουν σε ερωτήσεις όπως «πώς μπορώ να κάνω έναν αυτοματοποιημένο ελεγκτή γεγονότων να κατανοήσει τις διαφοροποιήσεις του πλαισίου;».

Δημοσιογραφική κατανεμημένη νόηση

Εν κατακλείδι, θα ήθελα να διερευνήσω περαιτέρω την ιδέα της δημοσιογραφικής κατανεμημένης νόησης και τα ερωτήματα που εγείρει. Οι Άντερσον, Γουίλερ και Σπρέβακ υποστηρίζουν ότι, καθώς οι υπολογιστές καθίστανται παντοχού παρόντες στην ανθρώπινη δραστηριότητα, η νόηση «εξαπλώνεται στον εγκέφαλο, στο μη νευρικό σώμα και... σε ένα περιβάλλον που αποτελείται από αντικείμενα, εργαλεία, άλλα τεχνουργήματα, κείμενα, άτομα, ομάδες ή/και κοινωνικές/θεσμικές δομές» (Anderson κ.α., 2018). Στη δημοσιογραφία, αυτό σημαίνει ότι, επί του παρόντος, καθώς οι δημοσιογράφοι χρησιμοποιούν δικτυωμένο λογισμικό και υλικό για να αυξήσουν την ικανότητά τους να παράγουν ειδήσεις σε κλίμακα και με ταχύτητα, η νόησή τους κατανέμεται στο φάσμα των πλατφορμών και των εργαλείων που χρησιμοποιούν. Φυσικά, αυτό τους παρέχει απεριόριστη πρόσβαση στο μεγαλύτερο μέρος της ανθρώπινης γνώσης στο Διαδίκτυο.

Ωστόσο, αυτή η ιδέα της φορητής γνώσης και της κατανεμημένης νόησης εγείρει το ερώτημα ποιος κατέχει και διαχειρίζεται την πρόσβαση των δημοσιογράφων σε αυτόν τον πλούτο γνώσεων και στην «ελεύθερη» αναλυτική δύναμη. Ποιος καθιστά δυνατή τη δημοσιογραφική κατανεμημένη γνώση; Το ζήτημα αυτό, που αξίζει να συζητηθεί διεξοδικότερα, είναι ακανθώδες, όπως διαπιστώσαμε όταν η Google διέκοψε τη λειτουργία του διαδικτυακού εργαλείου απεικόνισης δεδομένων Google Fusion Tables. Μετά το κλείσιμο της πλατφόρμας, δεκάδες έργα δημοσιογραφίας δεδομένων που είχαν αναπτυχθεί με το εργαλείο δεν ήταν πια διαθέσιμα, καθώς δεν υποστηρίζονταν πλέον από την εταιρεία.

Στο πλαίσιο αυτό, καθώς οι δημοσιογράφοι ασχολούνται καθημερινά με την υπολογιστική δυναμική, η υπολογιστική τους σκέψη κανονικοποιείται και διευκολύνει την προβολή της νόησής τους στις μηχανές που χρησιμοποιούν για τις καθημερινές τους δημοσιογραφικές εργασίες. Καθώς η δημοσιογραφική γνώση κατανέμεται, συμβαίνει άραγε το ίδιο με τη δημοσιογραφική εξουσία και τον έλεγχο; Αναπόφευκτα, η κατανομή μετατοπίζει τα όρια που παρέχουν στους δημοσιογράφους τον έλεγχο της καθημερινής τους εργασίας και της επαγγελματικής τους νοοτροπίας, επηρεάζοντας έτσι το επιστημολογικό τους κύρος. Κοιτάζοντας μπροστά, όπως κάναμε στη φανταστική εισαγωγή αυτού του κεφαλαίου, η κατανομή θα μπορούσε, επίσης, να δημιουργήσει μια σειρά από συναφείς κινδύνους, όταν οι δημοσιογράφοι αρχίσουν να αναθέτουν σημαντικές ηθικές θεωρήσεις και αποφάσεις σε μηχανές. Είναι σημαντικό, λοιπόν, η υποδομή που χρησιμοποιούν για να κατανέμουν τη νόησή τους να είναι ανοιχτή και διαθέσιμη για δημόσιο έλεγχο, προκειμένου να διαφυλαχθούν τα ακρογωνιαία ιδανικά της δημοσιογραφίας στην εποχή των δεδομένων και της αυτοματοποίησης.

Παραπομπές

Anderson, M., Wheeler, M., & Sprevak, M. (2018) «Distributed cognition and the humanities», στο M. Anderson, D. Cairns, & M. Sprevak (επιμ.) *Distributed cognition in classical antiquity*, σσ. 1-17. Εδιμβούργο: Edinburgh University Press.

Bradshaw, P. (2017) «Computational thinking and the next wave of data journalism», *Online Journalism Blog*. Διαθέσιμο σε: https://onlinejournalismblog.com/2017/08/03/computational-thinking-data-journalism/ (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Denning, P. J., & Martell, C. H. (2015). *Great principles of computing.* Κέιμπριτζ, Μασαχουσέτη: The MIT Press.

Gynnild, A. (2014) «Journalism innovation leads to innovation journalism: The impact of computational exploration on changing mindsets», *Journalism*, 15(6), σσ. 713–730. DOI: https://doi.org/10.1177/1464884913486393

Myles, S. (2019, February 1). «Photomation or fauxtomation? Automation in the newsroom and the impact on editorial labour: A case study» [Technology], *Computation + Journalism Symposium.* Πανεπιστήμιο του Μαϊάμι, Φλόριντα, 1 Φεβρουαρίου.

Wheeler, M. (2018) «Talking about more than heads: The embodied, embedded and extended creative mind», στο B. Gaut & M. Kieran (επιμ.), *Creativity and philosophy*, Νέα Υόρκη: Routledge, σσ. 230-250. Διαθέσιμο σε: http://dspace.stir.ac.uk/handle/1893/26296 (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Wing, J. M. (2006) «Computational thinking», *Communications of the ACM*, 49, σσ. 33-35. DOI: https://doi.org/10.1145/1118178.1118215

Wing, J. M. (2008) «Computational thinking and thinking about computing», *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1881), σσ. 3717-3725. DOI: https://doi.org/10.1098/rsta.2008.0118

Ο συγγραφέας

Ο Έντι Μπόρχες-Ρέι είναι Αναπληρωτής Καθηγητής Ψηφιακής Δημοσιογραφίας και Αναδυόμενων Μέσων στο Πανεπιστήμιο Northwestern στο Κατάρ και πρώην δημοσιογράφος ραδιοτηλεοπτικών εκπομπών.

5.

Η Εμπειρία της Εργασίας με Δεδομένα

5.1
Τρόποι Άσκησης της Δημοσιογραφίας Δεδομένων

Sarah Cohen

Περίληψη: Αυτό το κεφάλαιο διερευνά τους ποικίλους τρόπους με τους οποίους έχει εξελιχθεί η δημοσιογραφία δεδομένων και τις διαφορετικές μορφές που μπορεί να πάρει, από το παραδοσιακό ερευνητικό ρεπορτάζ έως τις ενημερωτικές εφαρμογές και τις οπτικοποιήσεις.

Λέξεις-κλειδιά: ερευνητική δημοσιογραφία, ενημερωτικές εφαρμογές, οπτικοποίηση δεδομένων, επεξηγηματική δημοσιογραφία, δημοσιογραφία ακριβείας

δεδομένα: ένα σύνολο γεγονότων ή πληροφοριών· μεμονωμένα γεγονότα, στατιστικές ή στοιχεία («Data», άνευ ημερομηνίας)

δημοσιογραφία: το επάγγελμα του ρεπορτάζ, της συγγραφής, της επεξεργασίας, της φωτογράφισης ή της μετάδοσης ειδήσεων, ή της διεύθυνσης οποιουδήποτε ειδησεογραφικού οργανισμού ως επιχείρηση («Journalism», άνευ ημερομηνίας)

Εάν διαβάζετε αυτό το εγχειρίδιο, θα είναι επειδή έχετε αποφασίσει ότι θέλετε να μάθετε κάποια πράγματα για την πρακτική που έχει γίνει γνωστή ως δημοσιογραφία δεδομένων. Όμως, τι ακριβώς σημαίνει αυτό σε μια εποχή όπου οι πύλες δεδομένων είναι ανοιχτές, οι οπτικοποιήσεις εκθαμβωτικές και δίνονται μάχες για την ελευθερία της πληροφόρησης σε ολόκληρο τον κόσμο;

Ένας λεξικογραφικός ορισμός των δύο λέξεων δεν βοηθάει πολύ –εάν τις συνδυάσουμε, προκύπτει το συμπέρασμα ότι η δημοσιογραφία δεδομένων είναι ένα επάγγελμα παραγωγής ειδήσεων που απαρτίζονται από γεγονότα ή πληροφορίες. Η δημοσιογραφία δεδομένων έχει καταλήξει να σημαίνει σχεδόν κάθε δημοσιογραφική πράξη που αγγίζει ηλεκτρονικά αρχεία και στατιστικά στοιχεία –με άλλα λόγια, σχεδόν το σύνολο της δημοσιογραφίας.

Αυτός είναι ο λόγος για τον οποίο πολλοί από τους ανθρώπους στον τομέα δεν θεωρούν τους εαυτούς τους δημοσιογράφους δεδομένων –είναι πιο πιθανό να θεωρούν τους εαυτούς τους επεξηγηματικούς συντάκτες, γραφικούς ή οπτικούς δημοσιογράφους, ρεπόρτερ, αναλυτές κοινού ή προγραμματιστές ενημερωτικών εφαρμογών– όλα πιο ακριβείς τίτλοι για τις πολλές παραφυάδες αυτού του αναπτυσσόμενου τομέα. Και σαν να μη φτάνει αυτό, προσθέστε οτιδήποτε σε μια αίθουσα σύνταξης απαιτεί τη χρήση αριθμών ή οτιδήποτε απαιτεί προγραμματισμό. Αυτό που κάποτε ήταν μια μπάντα σε γκαράζ έχει πλέον εξελιχθεί σε μια ολόκληρη ορχήστρα.

Η δημοσιογραφία δεδομένων δεν είναι κάτι καινούριο. Στην πραγματικότητα, εάν θεωρείτε τα «δεδομένα» ως κάποιο είδος συστηματικής συλλογής, τότε κάποια από τα πρώτα δημοσιογραφικά δεδομένα στις Ηνωμένες Πολιτείες χρονολογούνται στα μέσα του 19ου αιώνα, όταν ο Φρανκ Λέσλι, εκδότης του περιοδικού *Frank Leslie's Illustrated Newspaper*, προσέλαβε ντετέκτιβ για να παρακολουθήσουν τα καροτσάκια με τα γαλακτοκομικά προϊόντα σε όλη τη Νέα Υόρκη, προκειμένου να καταγράψουν το χαλασμένο γάλα με την ακατάλληλη σήμανση. Ο Σκοτ

Κλάιν, βοηθός αρχισυντάκτη στον μη κερδοσκοπικό ερευνητικό ιστότοπο της *ProPublica*, κατέγραψε μια συναρπαστική ιστορία της δημοσιογραφίας δεδομένων που χρονολογείται, επίσης, από τις αρχές του 19ου αιώνα (Klein, 2016), στην οποία οι εφημερίδες δίδαξαν στους αναγνώστες πώς να κατανοούν ένα ραβδόγραμμα. Ο Κρις Άντερσον διερευνά, επίσης, διαφορετικές γενεαλογίες της δημοσιογραφίας δεδομένων στις δεκαετίες του 1910, του 1960 και του 2010 στο δικό του κεφάλαιο σε αυτόν τον τόμο[114].

Με αυτές τις ιστορίες, οι ταξινομήσεις των διαφορετικών κλάδων της δημοσιογραφίας δεδομένων μπορούν να βοηθήσουν τους φοιτητές και τους επαγγελματίες να ξεκαθαρίσουν ποια σταδιοδρομία προτιμούν και ποιες δεξιότητες απαιτούνται για να επιτύχουν σε αυτή. Οι προαναφερθέντες διαφορετικοί τρόποι άσκησης της δημοσιογραφίας δεδομένων παρατίθενται εδώ σε μια κατά προσέγγιση χρονολογική σειρά της ανάπτυξης του κλάδου.

Εμπειρική δημοσιογραφία ή τα δεδομένα στην υπηρεσία των ιστοριών

Ο Μόρις Τάμαν του *Reuters* επινόησε τον όρο «εμπειρική δημοσιογραφία» («empirical journalism»), προκειμένου να συνδυάσει δύο παραδόσεις της δημοσιογραφίας δεδομένων. Η δημοσιογραφία ακριβείας («precision journalism»), που αναπτύχθηκε τη δεκαετία του 1960 από τον Φίλιπ Μάγιερ, επεδίωκε τη χρήση μεθόδων των κοινωνικών επιστημών στα δημοσιογραφικά θέματα. Το εύρος της εργασίας του εκτεινόταν από τη διενέργεια μιας έρευνας για τους εξεγερμένους στο Ντιτρόιτ μέχρι την επίβλεψη της συλλογής και ανάλυσης δεδομένων σε έρευνα για τις φυλετικές προκαταλήψεις στα δικαστήρια της Φιλαδέλφειας. Έθεσε τις βάσεις για την ερευνητική εργασία για μια ολόκληρη γενιά. Η εμπειρική δημοσιογραφία μπορεί, επίσης, να περιλαμβάνει αυτό που έγινε γνωστό ως ρεπορτάζ με τη βοήθεια υπολογιστή («computer-assisted reporting») τη δεκαετία του 1990, ένα είδος που εισήγαγε ο Έλιοτ Τζάσπιν στο Πρόβιντενς του Ρόουντ Άιλαντ. Σε αυτόν τον κλάδο, οι δημοσιογράφοι αναζητούν ντοκουμέντα σε ηλεκτρονική μορφή –ή τα δημιουργούν, όταν πρέπει–, για να ερευνήσουν μια πληροφορία ή μια ιδέα για ένα θέμα.

Πιο πρόσφατα, αυτοί οι δημοσιογράφοι έχουν αρχίσει να χρησιμοποιούν την τεχνητή νοημοσύνη («artificial intelligence») και τη μηχανική μάθηση («machine learning»), για να βρουν ή να απλοποιήσουν την ανάπτυξη ιστοριών. Αυτά τα εργαλεία μπορούν να χρησιμοποιηθούν για την απάντηση απλών ερωτημάτων, όπως το φύλο ενός ασθενούς που υπέστη βλάβη από ιατρικές συσκευές, ενώ η κυβέρνηση προσπαθούσε να αποκρύψει αυτή τη λεπτομέρεια. Διαφορετικά, μπορούν να χρησιμοποιηθούν για τον εντοπισμό δύσκολων μοτίβων, όπως η ανάλυση των κατασκοπευτικών αεροσκαφών του Πίτερ Άλντους για το *Buzzfeed News* (Aldhous, 2017· Woodman, 2019).

Αυτοί οι δημοσιογράφοι είναι σχεδόν καθαροί συλλέκτες ειδήσεων –ο στόχος τους δεν είναι να δημιουργήσουν μια οπτικοποίηση ούτε να αφηγηθούν ιστορίες με δεδομένα. Αντ' αυτού, χρησιμοποιούν τα αρχεία για να εξερευνήσουν μια πιθανή ιστορία. Η δουλειά τους είναι αναπόσπαστο μέρος της διαδικασίας του ρεπορτάζ και συχνά καθορίζει την εξέλιξη μιας έρευνας. Συνήθως εμπλέκονται λιγότερο στην παρουσίαση της ιστορίας.

Αναμφισβήτητα, η νεότερη προσθήκη σε αυτόν τον κόσμο της «δημοσιογραφίας δεδομένων» θα μπορούσε να είναι η αυξανόμενη επίδραση των οπτικών ερευνών και των ερευνών ανοιχτού κώδικα σε παγκόσμιο επίπεδο. Αυτό το είδος, το οποίο προέρχεται από την έρευνα στον τομέα των πληροφοριών και των ανθρωπίνων δικαιωμάτων, επεκτείνει την έννοια των «δεδομένων» σε βίντεο, μέσα κοινωνικής δικτύωσης με τη συμμετοχή του πλήθους και άλλα ψηφιακά τεχνουργήματα. Μολονότι εξαρτάται λιγότερο από τον προγραμματισμό, εντάσσεται σταθερά στην παράδοση της δημοσιογραφίας δεδομένων, αποκαλύπτοντας –μέσω πρωτότυπης έρευνας– αυτά που άλλοι θα ήθελαν να κρατήσουν μυστικά.

114. ΣτΕ: Κεφάλαιο 9.1, «Γενεαλογίες της Δημοσιογραφίας Δεδομένων»

Ένα από τα πιο διάσημα παραδείγματα, το *Anatomy of a Killing* από την ενότητα ντοκιμαντέρ *Africa Eye* του BBC, αποκαλύπτει πού ακριβώς έλαβε χώρα η δολοφονία μιας οικογένειας στο Καμερούν και πότε συνέβη, και βοηθά στον εντοπισμό των εμπλεκομένων –ενώ η κυβέρνηση του Καμερούν είχε αρνηθεί το περιστατικό ως «ψευδή είδηση» (BBC News, 2018). Η ομάδα χρησιμοποίησε εργαλεία από το Google Earth, για τον εντοπισμό του περιγράμματος μιας κορυφογραμμής, μέχρι το Facebook, για την αναγνώριση των ρούχων που φορούσαν οι δολοφόνοι.

Οπτικοποίηση δεδομένων

Κοιτάζοντας τους νικητές των διεθνών Βραβείων Δημοσιογραφίας Δεδομένων, ο αναγνώστης θα μπορούσε να σκεφτεί ότι η οπτικοποίηση είναι το κλειδί για κάθε είδους δημοσιογραφία δεδομένων[115]. Εάν τα στατιστικά στοιχεία είναι το νόμισμα, η οπτικοποίηση είναι η τιμή της εισόδου στο κλαμπ.

Οι οπτικοποιήσεις μπορούν να αποτελέσουν σημαντικό τμήμα της εργαλειοθήκης ενός δημοσιογράφου δεδομένων. Ωστόσο, απαιτούν μια συλλογή εργαλείων που προέρχεται τόσο από τον κόσμο του σχεδιασμού και της τέχνης όσο και από τον κόσμο των δεδομένων, της στατιστικής και του ρεπορτάζ. Ο Αλμπέρτο Κάιρο, ένας από τους πιο διάσημους οπτικούς δημοσιογράφους («visual journalists»), που εργάζεται σήμερα στον ακαδημαϊκό χώρο, προέρχεται από τον κόσμο των ενημερωτικών γραφικών («infographics») σε περιοδικά και εφημερίδες. Η δουλειά του επικεντρώνεται στην αφήγηση ιστοριών μέσω της οπτικοποίησης –ένας ρόλος τόσο αφηγητή ιστοριών όσο και συλλέκτη ειδήσεων.

Ενημερωτικές εφαρμογές

Στην *ProPublica*, οι περισσότερες μεγάλες έρευνες ξεκινούν ή τελειώνουν με μια ενημερωτική εφαρμογή («news application» ή «news app») –έναν ιστότοπο ή μια λειτουργικότητα που παρέχει πρόσβαση σε τοπικά ή μεμονωμένα δεδομένα μέσω μιας ελκυστικής και κατατοπιστικής διεπαφής. Η *ProPublica* έχει γίνει γνωστή για τις ενημερωτικές εφαρμογές της και μηχανικοί που ξεκίνησαν την καριέρα τους στον προγραμματισμό εξελίχθηκαν σε δημοσιογράφους που χρησιμοποιούν κώδικα, αντί για λέξεις, προκειμένου να αφηγηθούν ιστορίες.

Ο Κεν Σβένκε της *ProPublica*, ένας αυτοδίδακτος προγραμματιστής, ο οποίος έχει εργαστεί σε αίθουσες σύνταξης εντύπων όπως οι *Los Angeles Times* και οι *New York Times*, έγινε ένας από τους κορυφαίους δημοσιογράφους στην κάλυψη εγκλημάτων μίσους στις Ηνωμένες Πολιτείες, στο πλαίσιο του έργου *Documenting Hate*, το οποίο περιστρέφεται γύρω από ιστορίες που συλλέγονται από το κοινό μέσω της ενημερωτικής εφαρμογής της *ProPublica*.

Ιστορίες δεδομένων

Ο όρος «δημοσιογραφία δεδομένων» ενηλικιώθηκε καθώς δημοσιογράφοι, στατιστικολόγοι και άλλοι ειδικοί άρχισαν να γράφουν για τα δεδομένα ως ένα αυτοτελές είδος δημοσιογραφίας. Ο Σάιμον Ρότζερς, ο δημιουργός του Datablog του *Guardian*, καθιέρωσε το είδος. Το *FiveThirtyEight*, το *Vox* και, αργότερα, το Upshot των *New York Times* έγιναν οι σημαιοφόροι αυτού του κλάδου. Ο καθένας αντιμετώπιζε τον ρόλο του με διαφορετικό τρόπο, αλλά συνέκλιναν στην ιδέα ότι οι στατιστικές και οι αναλύσεις είναι από μόνες τους ειδησεογραφικά ενδιαφέρουσες.

115. Βλ. το κεφάλαιο της Λούζεν σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 9.2, «Χρυσά Πρότυπα Με Γνώμονα τα Δεδομένα: Τι Αποτιμά ο Κλάδος ως Δημοσιογραφία Δεδομένων που αξίζει Βραβείο και Πώς η Δημοσιογραφία Δεδομένων Εξελίσσεται μαζί με τη Δεδομενοποίηση της Κοινωνίας»).

Κάποια από αυτά τα μέσα έγιναν περισσότερο γνωστά για τις πολιτικές τους προβλέψεις, ποντάροντας σε αποδόσεις για τις προεδρικές εκλογές των ΗΠΑ. Άλλα έγιναν γνωστά για την εύρεση ιδιόμορφων συνόλων δεδομένων που προσφέρουν μια γεύση από την ψυχοσύνθεση του κοινού. Ένα τέτοιο παράδειγμα αποτελεί ο χάρτης του 2014 για τις προτιμήσεις στο μπέιζμπολ που προέκυψε από τις προτιμήσεις του Facebook στις ΗΠΑ. Η αφετηρία για αυτό το είδος δημοσιογραφίας είναι ένα σύνολο δεδομένων και η εξειδίκευση σε ένα θέμα είναι ο τρόπος με τον οποίο οι συγκεκριμένοι επαγγελματίες ξεχωρίζουν από τους υπόλοιπους του τομέα. Στην πραγματικότητα, ο Νέιτ Σίλβερ και άλλοι που καθόρισαν τον κλάδο δεν προέρχονταν από τον χώρο της δημοσιογραφίας αλλά από τον κόσμο της στατιστικής και της πολιτικής επιστήμης.

Η Αμάντα Κοξ, αρχισυντάκτρια του Upshot των *New York Times*, έχει δηλώσει ότι θεωρεί πως ο ρόλος του ιστοτόπου είναι να καταλαμβάνει τον χώρο μεταξύ των γνωστών αδιάσειστων στοιχείων και του άγνωστου –ένα είδος δημοσιογραφίας που παρέχει πληροφορίες μέσω της ανάλυσης των διαθέσιμων δεδομένων από εμπειρογνώμονες και που κινείται στα όρια ανάμεσα στα σκέτα γεγονότα και τη σκέτη άποψη (Cox, προσωπική επικοινωνία, άνευ ημερομηνίας).

Ερευνώντας τους αλγορίθμους

Ένας αναδυόμενος τομέας της δημοσιογραφίας δεδομένων είναι, στην πραγματικότητα, δημοσιογραφία για την τεχνολογία –ο τομέας της «αλγοριθμικής λογοδοσίας» («algorithmic accountability»), όρος που επινοήθηκε από τον Νίκολας Ντιακόπουλος στο Πανεπιστήμιο Northwestern[116]. Οι δημοσιογράφοι Τζούλια Άνγκουιν και Τζεφ Λάρσον άφησαν την *ProPublica*, για να ακολουθήσουν αυτόν τον εξειδικευμένο κλάδο, ιδρύοντας τον ιστότοπο The Markup που, σύμφωνα με την Άνγκουιν, θα θέτει τις εταιρείες τεχνολογίας προ των ευθυνών τους για τα αποτελέσματα που παράγουν οι αλγόριθμοι μηχανικής μάθησης και τεχνητής νοημοσύνης για την κοινωνία μας, από τις αποφάσεις για τις ποινές φυλάκισης μέχρι τις τιμές χρέωσης με βάση τον ταχυδρομικό κώδικα ενός καταναλωτή.

Το εν λόγω ρεπορτάζ έχει ήδη ωθήσει το YouTube να αναθεωρήσει τις μηχανές συστάσεών του («recommendation engines»), για να μειώσει την τάση του να οδηγεί τους θεατές σε ολοένα και πιο βίαια βίντεο, ανάγκασε το Facebook να αναλάβει τις ευθύνες του για τις δυνητικά μεροληπτικές αγγελίες ακινήτων και έχει εντοπίσει διακρίσεις στις τιμές ηλεκτρονικών καταστημάτων, με βάση την περιοχή του χρήστη (Dwoskin, 2019).

Παραπομπές

Aldhous, P. (2017) «We trained a computer to search for hidden spy planes. This is what it found», *BuzzFeed* News, 8 Αυγούστου. Διαθέσιμο σε: https://www.buzzfeednews.com/article/peteraldhous/hidden-spy-planes (Προσπελάστηκε: 29 Νοεμβρίου 2021).

BBC News (2018) «Anatomy of a killing», *BBC Africa Eye* [You Tube] 23 Σεπτεμβρίου. Διαθέσιμο σε: https://www.youtube.com/watch?v=4G9S-eoLgX4 (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Dictionary.com (άνευ ημερομηνίας) *Data*. Διαθέσιμο σε: https://www.dictionary.com/browse/data (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Dictionary.com (άνευ ημερομηνίας) *Journalism*. Διαθέσιμο σε: https://www.dictionary.com/browse/journalism (Προσπελάστηκε: 29 Νοεμβρίου 2021).

116. Για περισσότερα σχετικά με αυτόν τον τομέα, βλ. τα κεφάλαια των Ντιακόπουλος και Έλμερ σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 6.1, «Το Ρεπορτάζ των Αλγορίθμων: Οπτικές και Μέθοδοι Έρευνας», και Κεφάλαιο 6.5, «Ρίχνοντας Φως στους Αλγορίθμους: Συνεργατικές Έρευνες στο *Der Spiegel*», αντίστοιχα).

Dwoskin, E. (2019) «YouTube is changing its algorithms to stop recom-mending conspiracies», *The Washington Post*, 25 Ιανουαρίου. Διαθέσιμο σε: https://www.washingtonpost.com/technology/2019/01/25/youtube-is-changing-its-algorithms-stop-recommending-conspiracies/ (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Klein, S. (2016) «Infographics in the time of cholera», *ProPublica*, 16 Μαρτίου. Διαθέσιμο σε: https://www.propublica.org/nerds/infographics-in-the-time-of-cholera (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Woodman, S. (2019) «Using the power of machines to complete impossible reporting tasks», *ICIJ*, 22 Οκτωβρίου, Διαθέσιμο σε: https://www.icij.org/blog/2019/10/using-the-power-of-machines-to-complete-impossible-reporting-tasks/ (Προσπελάστηκε: 29 Νοεμβρίου 2021).

Η συγγραφέας

Η Σάρα Κοέν κατέχει την έδρα Knight στη Δημοσιογραφία στη Σχολή Δημοσιογραφίας Walter Cronkite και προηγουμένως είχε εργαστεί ως αρχισυντάκτρια μιας ομάδας δημοσιογραφίας δεδομένων στους *New York Times*.

5.2

Οπτικοποιήσεις Δεδομένων: Τάσεις στις Αίθουσες Σύνταξης και Καθημερινές Αλληλεπιδράσεις

Helen Kennedy, William Allen, Martin Engebretsen, Rosemary Lucy Hill, Andy Kirk, Wibke Weber

Περίληψη: Αυτό το κεφάλαιο εξετάζει την παραγωγή οπτικοποιήσεων δεδομένων στις αίθουσες σύνταξης και τις καθημερινές αλληλεπιδράσεις του κοινού με αυτές.

Λέξεις-κλειδιά: οπτικοποίηση δεδομένων, εμπλοκή του κοινού, αίθουσες σύνταξης, αφήγηση, κοινό των δεδομένων

Αυτό το κεφάλαιο εξετάζει τόσο την παραγωγή οπτικοποιήσεων δεδομένων («data visualizations» ή, στην «αργκό» του τομέα, «dataviz») στις αίθουσες σύνταξης όσο και την καθημερινή αλληλεπίδραση του κοινού με αυτές, βασιζόμενο σε δύο ξεχωριστά ερευνητικά έργα. Το πρώτο είναι το *Seeing Data*, το οποίο διερεύνησε τον τρόπο με τον οποίο οι άνθρωποι δίνουν νόημα στις οπτικοποιήσεις δεδομένων, και το δεύτερο είναι το *Innovative Data Visualization and Visual-Numeric Literacy* (INDVIL), το οποίο διερεύνησε τις οπτικοποιήσεις δεδομένων ως σημειωτικούς, αισθητικούς και διαλεκτικούς πόρους στην κοινωνία[117]. Το κεφάλαιο ξεκινά συνοψίζοντας τα κύρια ευρήματα ενός υποέργου του INDVIL που επικεντρώθηκε στις οπτικοποιήσεις δεδομένων στις ειδήσεις, όπου διαπιστώσαμε ότι οι οπτικοποιήσεις γίνονται αντιληπτές με διάφορους τρόπους και χρησιμοποιούνται για διάφορους σκοπούς. Στη συνέχεια, συνοψίζει τα κύρια ευρήματά μας από το *Seeing Data*[118], όπου, επίσης, διαπιστώσαμε μεγάλη ποικιλομορφία, αυτή τη φορά στον τρόπο με τον οποίο το κοινό κατανοεί τις οπτικοποιήσεις. Αυτή η ποικιλομορφία είναι σημαντική για το μελλοντικό έργο τόσο των ερευνητών όσο και των επαγγελματιών που ασχολούνται με την οπτικοποίηση δεδομένων.

Οπτικοποίηση δεδομένων στις αίθουσες σύνταξης: τάσεις και προκλήσεις

Πώς ενσωματώνεται η οπτικοποίηση δεδομένων στην πρακτική της αίθουσας σύνταξης; Ποιες τάσεις εμφανίζονται και ποιες προκλήσεις προκύπτουν; Για να απαντήσουμε σε αυτά τα ερωτήματα, το 2016 και το 2017 πραγματοποιήσαμε 60 συνεντεύξεις σε 26 αίθουσες σύνταξης σε έξι ευρωπαϊκές χώρες: Νορβηγία, Σουηδία, Δανία, Γερμανία, Ελβετία και στο Ηνωμένο Βασίλειο (ΗΒ). Οι ερωτηθέντες από κυρίαρχα, διαδικτυακά μέσα ενημέρωσης περιλάμβαναν αρχισυντάκτες, επικεφαλής εξειδικευμένων ομάδων οπτικοποίησης δεδομένων, δημοσιογράφους δεδομένων, οπτικούς δημοσιογράφους («visual journalists»), σχεδιαστές γραφικών/οπτικοποίη-

117. http://seeingdata.org/, https://indvil.org/

118. Το δεύτερο μισό του κεφαλαίου συνοψίζει ένα μεγαλύτερο άρθρο που είναι διαθέσιμο στο Διαδίκτυο: Kennedy, H., Hill, R. L., Allen, W., & Kirk, A. (2016) «Engaging with (big) data visualizations: Factors that affect engagement and resulting new definitions of effectiveness», *First Monday*, 21(11). DOI: https://doi.org/10.5210/fm.v21i11.6389

σης δεδομένων και προγραμματιστές (αν και ορισμένοι δεν είχαν τίτλους εργασίας, σημάδι από μόνο του ότι πρόκειται για έναν ταχέως αναπτυσσόμενο τομέα). Παρουσιάζουμε μερικά από τα κυριότερα σημεία της έρευνάς μας εδώ.

Αλλάζοντας τη δημοσιογραφική αφήγηση

Η χρήση της οπτικοποίησης δεδομένων στη δημοσιογραφία σημαίνει ότι υπάρχει μια μετατόπιση από τη γραφή, ως κύριο σημειωτικό τρόπο, στα δεδομένα και στην οπτικοποίηση, ως κεντρικά στοιχεία της δημοσιογραφικής αφήγησης. Πολλοί ερωτηθέντες δήλωσαν ότι η οπτικοποίηση των δεδομένων είναι η κινητήρια δύναμη μιας ιστορίας, ακόμη και όταν πρόκειται για ένα απλό γραφικό ή διάγραμμα.

> Οι στατιστικές των αναγνωστών μάς λένε ότι, όταν εντάσσουμε μια απλή οπτικοποίηση δεδομένων σε μια ιστορία, οι αναγνώστες παραμένουν στη σελίδα λίγο περισσότερο (Σουηδία).

Οι οπτικοποιήσεις δεδομένων χρησιμοποιούνται με ένα ευρύ φάσμα επικοινωνιακών προθέσεων, όπως: «για να προσφέρουν πληροφορίες» (ΗΒ), «για ευκολότερη επεξήγηση» (Σουηδία), «για να επικοινωνήσουν με σαφήνεια, πιο ξεκάθαρα από ό,τι μπορούν οι λέξεις» (ΗΒ), «για να αφηγηθούν διάφορες πτυχές με λεπτομέρεια, κάτι που στο κείμενο είναι δυνατό μόνο σε συνοπτική μορφή» (Γερμανία), για να κάνουν τις ιστορίες «πιο προσιτές» (Δανία), «για να αποκαλύψουν θλιβερές καταστάσεις» (Ελβετία), «για να βοηθήσουν τους ανθρώπους να κατανοήσουν τον κόσμο» (ΗΒ). Η οπτικοποίηση δεδομένων χρησιμοποιείται για να τονίσει ένα σημείο, για να προσθέσει εμπειρικά στοιχεία, για να επιτρέψει στους χρήστες να εξερευνήσουν σύνολα δεδομένων, ως αισθητικός πόλος έλξης για την τόνωση του ενδιαφέροντος, και για να ανοίξει τον δρόμο σε αθέατες ιστορίες.

Οι αλλαγές αυτές συνοδεύονται από την εμφάνιση εξειδικευμένων ομάδων πολλαπλών ειδικοτήτων στις αίθουσες σύνταξης, με τις δεξιότητες που αφορούν στα δεδομένα και στις οπτικοποιήσεις τους να ιεραρχούνται ως προτεραιότητα στις νέες προσλήψεις. Ωστόσο, δεν υπάρχουν καθορισμένα μοτίβα στην οργάνωση της παραγωγής των οπτικοποιήσεων δεδομένων μέσα στις αίθουσες σύνταξης –σε ορισμένες, οι οπτικοποιήσεις παράγονται από τις ομάδες δεδομένων, σε άλλες, από τις ομάδες δημιουργίας οπτικού υλικού (ένας από τους συνεντευξιαζόμενους σχεδιαστές οπτικοποιήσεων δεδομένων δούλευε, επίσης, σε ένα έργο εικονικής πραγματικότητας κατά τη διάρκεια της συνέντευξης) και αλλού, από εντελώς διαφορετικές ομάδες. Και όπως ακριβώς αναδύονται νέες δομές, για να φιλοξενήσουν αυτή τη νέα αναπτυσσόμενη οπτική μορφή, έτσι και το προσωπικό της αίθουσας σύνταξης πρέπει να προσαρμοστεί, για να μάθει νέα εργαλεία (που είτε έχουν αναπτυχθεί εσωτερικά είτε είναι εμπορικά διαθέσιμα), να αναπτύξει νέες δεξιότητες και να κατανοήσει τον τρόπο επικοινωνίας μεταξύ ομάδων και τομέων εξειδίκευσης, προκειμένου να παράγει αποτελεσματικές ιστορίες δεδομένων.

Το μότο «mobile first» και οι συνέπειές του

Η ευρεία παραδοχή ότι το κοινό καταναλώνει ολοένα και περισσότερο τις ειδήσεις σε μικρές, κινητές οθόνες έχει οδηγήσει στην εξίσου ευρεία υιοθέτηση του μότο «mobile first» («πρώτα για κινητές συσκευές»), όταν πρόκειται για την παραγωγή οπτικοποιήσεων δεδομένων στις αίθουσες σύνταξης. Αυτό σηματοδοτεί την απομάκρυνση από τις περίτεχνες και διαδραστικές οπτικοποιήσεις που χαρακτήριζαν τις πρώτες ημέρες των οπτικοποιήσεων στις ειδήσεις, προς χάριν μεγαλύτερης απλότητας και γραμμικότητας ή απλών οπτικών μορφών με χαμηλά επίπεδα διαδραστικότητας. Είναι κάτι που έχει οδηγήσει στην επικράτηση ορισμένων τύπων διαγραμμάτων, όπως τα ραβδογράμματα και τα γραμμικά διαγράμματα, καθώς και στην εμφάνιση του scrollytelling, δηλαδή των ιστοριών που ξεδιπλώνονται καθώς οι χρήστες μετακινούνται προς τα κάτω στη σελίδα με κύλιση –με τις οπτικοποιήσεις που είναι ενσωματωμένες στο άρθρο να εμφανίζονται την κατάλληλη στιγμή. Η διαδικασία της κύλισης επιφέρει, επίσης, αλλαγές στις ίδιες τις απεικονίσεις, όπως η σμίκρυνση.

> Συχνά στις ιστορίες μας χρησιμοποιούμε την τεχνική κύλισης. Δεν είναι το κλικ απαραίτητο, αλλά η κύλιση· εάν κάνετε κύλιση προς τα κάτω, κάτι θα συμβεί στην ιστορία (Γερμανία).

Τα εργαλεία που αυτοματοποιούν την παραγωγή οπτικοποιήσεων δεδομένων και καθιστούν δυνατή την παραγωγή τους από δημοσιογράφους που δεν είναι ειδικοί στις οπτικοποιήσεις οδηγούν, επίσης, στη διάδοση απλουστευμένων μορφών διαγραμμάτων. Παρ' όλα αυτά, ορισμένοι ερωτηθέντες επιθυμούν να εκπαιδεύσουν τους αναγνώστες, παρουσιάζοντας λιγότερο συνηθισμένα διαγράμματα (ένα διάγραμμα διασποράς, για παράδειγμα), συνοδευόμενα από πληροφορίες σχετικά με το πώς να τα κατανοήσουν. Ορισμένοι πιστεύουν ότι οι εικόνες μπορούν, επίσης, να παρουσιάσουν τα δεδομένα αποτελεσματικά –μια σκανδιναβική ταμπλόιντ εθνικής κυκλοφορίας κατέδειξε το μέγεθος ενός φορτηγού αεροσκάφους γεμίζοντάς το με 427.000 πίτσες. Άλλοι αναγνωρίζουν την αξία της κινούμενης εικόνας, για παράδειγμα, προκειμένου να δείξουν την αλλαγή με την πάροδο του χρόνου, ή να πειραματιστούν με τη δυνατότητα μεγέθυνσης στις οπτικοποιήσεις.

Ο κοινωνικός ρόλος της δημοσιογραφίας

Η σύνδεση μιας οπτικοποίησης με μια πηγή δεδομένων, η παροχή πρόσβασης στα ανεπεξέργαστα δεδομένα και η επεξήγηση των μεθοδολογιών θεωρούνται από ορισμένους συμμετέχοντες ως ηθικές πρακτικές που εξασφαλίζουν τη διαφάνεια και αντισταθμίζουν την υποκειμενικότητα της επιλογής και της ερμηνείας, η οποία, για ορισμένους, αποτελεί μια αναπόφευκτη πτυχή της οπτικοποίησης δεδομένων. Ωστόσο, για άλλους, η σύνδεση με πηγές δεδομένων σημαίνει ότι δίνουν στο κοινό «όλα τα δεδομένα» και αυτό έρχεται σε σύγκρουση με τον δημοσιογραφικό κανόνα της εύρεσης και ακολούθως της αφήγησης μιας ιστορίας. Για ορισμένους, η σύγκρουση αυτή αντιμετωπίζεται με πολύπλοκες διαδικασίες διαμοιρασμού διαφορετικών στοιχείων των δεδομένων και της διαδικασίας εργασίας σε διαφορετικές πλατφόρμες (Twitter, Pinterest, GitHub).

Αυτό είναι κάτι που οδηγεί τους δημοσιογράφους δεδομένων και τους σχεδιαστές οπτικοποιήσεων σε προβληματισμό σχετικά με το πόσα δεδομένα πρέπει να κοινοποιηθούν, τον ρόλο που διαδραματίζουν οι ίδιοι ως πάροχοι στοιχείων και τον κοινωνικό τους ρόλο γενικότερα. Ο δημοσιογράφος δεδομένων Πολ Μπράντσο συνοψίζει τα παραπάνω στο ιστολόγιό του:

> Πόση ευθύνη έχουμε για τις ιστορίες που λένε οι άνθρωποι με τις δικές μας πληροφορίες; Και πόση ευθύνη έχουμε να παρέχουμε σε κάποιον όσες πληροφορίες χρειάζεται (Bradshaw, 2017).

Ο πρώην αρχισυντάκτης της εφημερίδας *Guardian*, Άλαν Ράσμπριτζερ (2010), έθεσε ένα παρόμοιο ερώτημα σχετικά με τον κοινωνικό ρόλο της δημοσιογραφίας, όταν επισήμανε το εύρος των φορέων που κάνουν αυτό που ιστορικά έκανε η δημοσιογραφία, δηλαδή ενεργούν ως φύλακες των δεδομένων και των επίσημων πληροφοριών (π.χ. FixMyStreet και TheyWorkForYou στο Ηνωμένο Βασίλειο). Καταλήγει: «Δεν γνωρίζω αν αυτό είναι ή όχι δημοσιογραφία. Δεν ξέρω και αν έχει σημασία» (Baack, 2018). Ορισμένοι από τους ερωτώμενους εργάζονται σε έργα μεγάλης κλίμακας, παρόμοια με εκείνα που συζητήθηκαν από τον Ράσμπριτζερ –για παράδειγμα, ένα έργο συγκέντρωσε όλα τα διαθέσιμα δεδομένα σχετικά με τα σχολεία στο Ηνωμένο Βασίλειο και έδωσε τη δυνατότητα εξερεύνησής τους ανά ταχυδρομικό κώδικα, για να παρέχει πληροφορίες που θα βοηθούσαν τη λήψη αποφάσεων για το εκάστοτε σχολείο επιλογής. Έτσι, το ερώτημα του τι μετράει ως δημοσιογραφία στο πλαίσιο της ευρείας διάδοσης των δεδομένων και των οπτικοποιήσεων δεν είναι εύκολο να απαντηθεί.

Επιπλέον, ο διαμοιρασμός συνόλων δεδομένων προϋποθέτει ότι το κοινό θα αλληλεπιδράσει με αυτά. Ωστόσο, μελέτες δείχνουν ότι η διαδικτυακή διαδραστικότητα είναι τόσο μύθος όσο

και πραγματικότητα –με την εξιδανικευμένη εικόνα ενός ενεργού και με κίνητρα εξερευνητή ενός οπτικοποιημένου συνόλου δεδομένων να έρχεται σε αντίθεση με τον πιο συνηθισμένο αναγνώστη ειδήσεων που διαβάζει «σκρολάροντας» στα γρήγορα (Burmester κ.α., 2010). =Παρομοίως, μια μελέτη έργων δημοσιογραφίας δεδομένων που υποβλήθηκαν στα Nordic Data Journalism Awards καταλήγει στο συμπέρασμα ότι τα διαδραστικά στοιχεία συχνά προσφέρουν απλώς μια ψευδαίσθηση διαδραστικότητας, καθώς οι περισσότερες επιλογές έχουν ήδη γίνει ή προκαθοριστεί από τους δημοσιογράφους (Appelgren, 2017). Αυτό θέτει και πάλι υπό αμφισβήτηση την πρακτική του διαμοιρασμού «όλων των δεδομένων» και εγείρει ερωτήματα σχετικά με τον μεταβαλλόμενο κοινωνικό ρόλο της δημοσιογραφίας.

Εμπιστοσύνη, αλήθεια και οπτικοποιήσεις «στην άγρια ζούγκλα»

Άλλα στοιχεία της διαδικασίας οπτικοποίησης δεδομένων εγείρουν ζητήματα εμπιστοσύνης και αλήθειας και σχετίζονται, επίσης, με τον τρόπο με τον οποίο στοχάζονται οι δημοσιογράφοι για τον κοινωνικό ρόλο της δημοσιογραφίας. Μια πτυχή της εργασίας οπτικοποίησης δεδομένων που παραπέμπει σε αυτά τα ζητήματα είναι ο τρόπος με τον οποίο σκέφτονται οι ίδιοι οι δημοσιογράφοι του τομέα για τα δεδομένα και την οπτική αναπαράστασή τους. Ορισμένοι το βλέπουν ως μια μορφή κατάθεσης της αλήθειας, άλλοι ως μια διαδικασία επιλογής και ερμηνείας. Άλλοι, πάλι, πιστεύουν ότι η διαμόρφωση οπτικοποιήσεων δεδομένων μέσω επιλογών είναι ένας τρόπος αποκάλυψης μιας ιστορίας και, άρα, είναι ακριβώς αυτό που πρέπει να κάνουν οι δημοσιογράφοι. Οι προβληματισμοί αυτοί αναδεικνύουν τη σχέση μεταξύ (ελλείμματος) εμπιστοσύνης και παρουσίασης και μεταξύ οπτικής και (αν)ειλικρίνειας.

Στο τρέχον πλαίσιο, το αποκαλούμενο πλαίσιο της «μετα-αλήθειας», στο οποίο λέγεται ότι το κοινό έχει βαρεθεί τα γεγονότα, τα δεδομένα και τους ειδήμονες και στο οποίο οι «ψευδείς ειδήσεις» κυκλοφορούν γρήγορα και ευρέως, οι συμμετέχοντές μας ήταν σε εγρήγορση για τους πιθανούς τρόπους με τους οποίους το κοινό θα μπορούσε να ανταποκριθεί στις απεικονίσεις των δεδομένων τους. Αυτοί οι τρόποι μπορεί να περιλαμβάνουν την αφελή αποδοχή, την αμφισβήτηση με σκεπτικισμό, την απομόνωση από το συγκείμενο μέσω του διαμοιρασμού στα μέσα κοινωνικής δικτύωσης, ή ακόμη και την αλλαγή και την παραποίηση. Οι ερωτώμενοι θεώρησαν ότι οι δημοσιογράφοι χρειάζονται όλο και περισσότερο «soft γνώσεις για την κουλτούρα του Διαδικτύου» (HB), όπως το έθεσε ένας συμμετέχων. Αυτό περιλαμβάνει την κατανόηση του τρόπου με τον οποίο το διαδικτυακό περιεχόμενο μπορεί να είναι πιο ανοιχτό στην αμφισβήτηση από το αντίστοιχο «offline» περιεχόμενο και του τρόπου με τον οποίο οι οπτικοποιήσεις δεδομένων είναι πιο πιθανό να κυκλοφορούν στο Διαδίκτυο απ' ό,τι τα κείμενα, αποκομμένες από το αρχικό τους πλαίσιο, ως συνδυασμοί αριθμών και εικόνων «στην άγρια ζούγκλα» (Espeland & Sauder, 2007). Αυτό με τη σειρά του απαιτεί την κατανόηση των στρατηγικών που θα μπορούσαν να αντιμετωπίσουν αυτούς τους κινδύνους, όπως η ενσωμάτωση επεξηγηματικού κειμένου σε ένα αρχείο οπτικοποίησης, έτσι ώστε η εικόνα να μην μπορεί να κυκλοφορήσει χωρίς την επεξήγηση. Αυτά τα ζητήματα, μαζί με την ανησυχία για την παιδεία του κοινού όσον αφορά στα δεδομένα και την οπτικοποίηση, επηρεάζουν και αναδιαμορφώνουν την αντίληψη των δημοσιογράφων για το κοινό τους.

Πώς αλληλεπιδρούν οι άνθρωποι με τις οπτικοποιήσεις δεδομένων;

Σε αυτήν την ενότητα, εξετάζουμε τις οπτικοποιήσεις στις ειδήσεις από την οπτική γωνία του κοινού. Πώς αλληλεπιδρά το κοινό με τις οπτικοποιήσεις που συναντά στα ενημερωτικά μέσα και πώς τις κατανοεί; Οι δημοσιογράφοι δεδομένων είναι συχνά πολύ απασχολημένοι για να ασχοληθούν με αυτό το ερώτημα. Οι ερευνητές οπτικοποίησης δεδομένων δεν έχουν αυτή τη δικαιολογία, αλλά παρ' όλα αυτά σπανίως εστιάζουν την προσοχή τους στο τι σκέφτονται οι τελικοί χρήστες για τις οπτικοποιήσεις που βλέπουν.

Εδώ έρχεται το *Seeing Data*, ένα ερευνητικό πρόγραμμα που εξέτασε τον τρόπο με τον οποίο οι άνθρωποι αλληλεπιδρούν με τις οπτικοποιήσεις δεδομένων που συναντούν στην καθημερινή τους ζωή, συχνά στα μέσα ενημέρωσης. Εξέτασε τους παράγοντες που επηρεάζουν την αλληλεπίδραση και τι σημαίνει αυτό για τον τρόπο με τον οποίο σκεφτόμαστε τι καθιστά μια οπτικοποίηση αποτελεσματική. Στο *Seeing Data* χρησιμοποιήσαμε ομάδες εστίασης («focus groups») και συνεντεύξεις, για να εξετάσουμε αυτά τα ερωτήματα, ώστε να μπορέσουμε να προσεγγίσουμε τις νοοτροπίες, τα συναισθήματα και τις πεποιθήσεις που διαμορφώνουν την ενασχόληση των ανθρώπων με την οπτικοποίηση δεδομένων. Στην έρευνα συμμετείχαν 46 άτομα. Σε αυτά συμπεριλαμβάνεται ένας συνδυασμός συμμετεχόντων: πήραν μέρος, αφενός, εκείνοι που θα μπορούσε να θεωρηθεί ότι ενδιαφέρονται για τα δεδομένα, την οπτικοποίηση ή τη μετανάστευση (η οποία ήταν το αντικείμενο αρκετών από τις οπτικοποιήσεις που τους δείξαμε) και, έτσι, ήδη «ασχολούνται» με ένα από τα ζητήματα που βρίσκονται στο επίκεντρο του έργου μας και, αφετέρου, εκείνοι για τους οποίους δεν μπορούσαμε να κάνουμε αυτές τις παραδοχές.

Στις ομάδες εστίασης, ζητήσαμε από τους συμμετέχοντες να αξιολογήσουν οκτώ οπτικοποιήσεις, τις οποίες επιλέξαμε (μετά από πολλή συζήτηση), επειδή αντιπροσώπευαν μια ποικιλία θεμάτων, τύπων διαγραμμάτων, πρωτότυπων πηγών και φορμάτ, και είχαν ως στόχο είτε να εξηγήσουν είτε να προσκαλέσουν σε διερεύνηση. Οι μισές από τις οπτικοποιήσεις προέρχονται από τον δημοσιογραφικό χώρο (BBC, *New York Times*, *The Metro*, μια εφημερίδα που διανέμεται ελεύθερα στο Ηνωμένο Βασίλειο, και το περιοδικό *Scientific American*). Άλλες προέρχονταν από οργανισμούς που οπτικοποιούν και διαμοιράζονται δεδομένα στο πλαίσιο της εργασίας τους: το Παρατηρητήριο Μετανάστευσης του Πανεπιστημίου της Οξφόρδης, την Εθνική Στατιστική Υπηρεσία του Ηνωμένου Βασιλείου («UK Office for National Statistics», ONS) και τον Οργανισμό Οικονομικής Συνεργασίας και Ανάπτυξης (ΟΟΣΑ).

Μετά τις ομάδες εστίασης, επτά συμμετέχοντες κράτησαν ημερολόγιο για έναν μήνα, προκειμένου να μας προσφέρουν περισσότερες πληροφορίες σχετικά με την έκθεσή τους σε οπτικοποιήσεις «στην άγρια ζούγκλα», οι οποίες δεν επιλέχθηκαν από εμάς.

Παράγοντες που επηρεάζουν την αλληλεπίδραση με τις οπτικοποιήσεις

Θέμα. Οι οπτικοποιήσεις δεν υφίστανται απομονωμένες από το θέμα που αναπαριστούν. Στις περιπτώσεις που η θεματολογία ανταποκρινόταν στα ενδιαφέροντα των συμμετεχόντων στην έρευνα, το ενδιαφέρον αυτό τους ωθούσε στην αλληλεπίδραση –για παράδειγμα, επαγγελματίες της κοινωνίας των πολιτών που ενδιαφέρονταν για θέματα σχετικά με τη μετανάστευση και, ως εκ τούτου, για τις οπτικοποιήσεις της μετανάστευσης. Αντίθετα, ένας συμμετέχων (άνδρας, 38 ετών, λευκός, Βρετανός, γεωργικός εργάτης) δεν ενδιαφέρθηκε για καμία από τις οπτικοποιήσεις που του δείξαμε στη διάρκεια των συζητήσεων ούτε έδειξε ότι είχε τη διάθεση να τις κοιτάξει. Ωστόσο, η έλλειψη ενδιαφέροντος και διάθεσης και η δυσπιστία του απέναντι στα μέσα ενημέρωσης (είπε ότι ένιωθε πως προσπαθούν «να σε μπερδέψουν») δεν τον εμπόδισαν εντελώς από το να κοιτάζει οπτικοποιήσεις γενικότερα: μας είπε ότι, όταν συναντούσε οπτικοποιήσεις στο *The Farmer's Guide*, μια έκδοση που διάβαζε τακτικά επειδή ανταποκρίνεται στα ενδιαφέροντά του, αφιέρωνε χρόνο για να τις κοιτάξει.

Τοποθέτηση της πηγής ή των μέσων. Η πηγή των οπτικοποιήσεων είναι σημαντική: έχει επιπτώσεις στον βαθμό εμπιστοσύνης των χρηστών. Πολλοί συμμετέχοντες μοιράστηκαν τις ανησυχίες τους σχετικά με τα μέσα ενημέρωσης που επιδιώκουν να προκαλέσουν σύγχυση και έχουν οδηγήσει ορισμένους να θεωρούν ύποπτες τις οπτικοποιήσεις που συναντούν σε συγκεκριμένα μέσα. Αντίθετα, ορισμένοι συμμετέχοντες εμπιστεύτηκαν τις οπτικοποιήσεις για τη μετανάστευση που έφεραν τον λογότυπο του Πανεπιστημίου της Οξφόρδης, επειδή θεώρησαν ότι το «εμπορικό σήμα» του εν λόγω πανεπιστημίου παραπέμπει σε ποιότητα και κύρος. Όμως, κατά την περίοδο που κρατούσαν το ημερολόγιο, προέκυψε μια διαφορετική εικόνα. Οι συμμετέχοντες είχαν την τάση να βλέπουν

οπτικοποιήσεις στα αγαπημένα τους μέσα ενημέρωσης, τα οποία εμπιστεύονταν, οπότε ήταν πιθανό να εμπιστεύονται και τις οπτικοποιήσεις που έβλεπαν εκεί. Ένας συμμετέχων (άνδρας, 24 ετών, λευκός, Βρετανός, γεωργικός εργάτης), ο οποίος διαβάζει την *Daily Mail*, επιβεβαίωσε αυτό το γεγονός, όταν κατά τη συνέντευξή του παρατήρησε: «νομίζω πως βλέπεις περισσότερα λάθη ή τυπογραφικά λάθη στη *Sun*». Λαμβάνοντας υπόψη τις ιδεολογικές ομοιότητες μεταξύ αυτών των δύο εντύπων, το σχόλιο αυτό υποδεικνύει τη σημασία της τοποθέτησης των μέσων ενημέρωσης στην αλληλεπίδραση με τις οπτικοποιήσεις.

Πεποιθήσεις και απόψεις. Οι συμμετέχοντες εμπιστεύονταν τις εφημερίδες που διάβαζαν τακτικά και, ως εκ τούτου, εμπιστεύονταν τις οπτικοποιήσεις τους, επειδή τόσο οι εφημερίδες όσο και οι οπτικοποιήσεις συχνά ταίριαζαν με τις απόψεις τους για τον κόσμο. Αυτό υποδεικνύει τη σημασία των πεποιθήσεων και των απόψεων που επηρεάζουν τον τρόπο και την πιθανότητα οι άνθρωποι να αφιερώνουν χρόνο για να ασχοληθούν με συγκεκριμένες οπτικοποιήσεις. Ορισμένοι συμμετέχοντες δήλωσαν ότι τους άρεσαν οι οπτικοποιήσεις που επιβεβαίωναν τις πεποιθήσεις και τις απόψεις τους. Οι πεποιθήσεις δεν έχουν, όμως, σημασία μόνο όταν τις επιβεβαιώνουν οι οπτικοποιήσεις. Ένας συμμετέχων (άνδρας, 34 ετών, λευκός, Βρετανός, τεχνικός υπολογιστών) εξεπλάγη από τα μεταναστευτικά δεδομένα σε μια απεικόνιση του ONS (Εικόνα 5.2.1). Είπε ότι δεν είχε συνειδητοποιήσει πόσοι άνθρωποι που ζουν στο Ηνωμένο Βασίλειο είχαν γεννηθεί στην Ιρλανδία. Τα δεδομένα αυτά έθεταν υπό αμφισβήτηση τα πιστεύω του και ο ίδιος απόλαυσε τη συγκεκριμένη εμπειρία. Σε ορισμένους ανθρώπους αρέσουν τα δεδομένα σε οπτικοποιήσεις που θέτουν υπό αμφισβήτηση τις υπάρχουσες πεποιθήσεις τους, ή τα βρίσκουν ενδιαφέροντα, επειδή προκαλούν και δοκιμάζουν τους ορίζοντές τους. Επομένως, οι πεποιθήσεις και οι απόψεις έχουν σημασία και κατ' αυτόν τον τρόπο.

Χρόνος. Η αλληλεπίδραση με τις οπτικοποιήσεις αντιμετωπίζεται ως κόπος από ανθρώπους που δεν τα καταφέρνουν να αλληλεπιδράσουν με ευκολία. Η ύπαρξη διαθέσιμου χρόνου είναι ζωτικής σημασίας για να αποφασίσει κανείς εάν είναι πρόθυμος να κάνει αυτό τον «κόπο». Οι

Εικόνα 5.2.1: Απογραφή πληθυσμού που δεν γεννήθηκε στο Ηνωμένο Βασίλειο, 1951-2011. Πηγή: Εθνική Στατιστική Υπηρεσία του Ηνωμένου Βασιλείου (ONS).

περισσότεροι συμμετέχοντες που δήλωσαν ότι δεν είχαν χρόνο να εξετάσουν οπτικοποιήσεις ήταν γυναίκες και απέδωσαν την έλλειψη χρόνου σε επαγγελματικές, οικογενειακές και οικιακές υποχρεώσεις. Μια εργαζόμενη μητέρα μίλησε για το πώς ο συνδυασμός αμειβόμενης και οικιακής εργασίας τής ήταν τόσο κουραστικός που, όταν τελείωνε την ημέρα της, δεν ήθελε να δει ειδήσεις, και αυτό περιλάμβανε και τις οπτικοποιήσεις. Τέτοιες δραστηριότητες της φαίνονταν «κόπος» και ήταν πολύ κουρασμένη για να τις εκτελέσει στο τέλος της πολυάσχολης ημέρας της. Ένας γεωργικός εργάτης μάς είπε σε ένα μήνυμα ηλεκτρονικού ταχυδρομείου ότι οι ώρες εργασίας του ήταν πάρα πολλές και αυτό επηρέασε την ικανότητά του να κρατήσει το μηνιαίο ημερολόγιο ενασχόλησής του με τις οπτικοποιήσεις μετά την ομάδα εστίασης.

Εμπιστοσύνη και ικανότητες. Το κοινό πρέπει να αισθάνεται ότι διαθέτει τις απαραίτητες δεξιότητες για την αποκωδικοποίηση των οπτικοποιήσεων και πολλοί συμμετέχοντες ανέφεραν έλλειψη εμπιστοσύνης στο θέμα αυτό. Ένας σύμβουλος σταδιοδρομίας που εργάζεται με μερική απασχόληση δήλωσε για μια οπτικοποίηση: «Είχε όλους αυτούς τους κύκλους και τα χρώματα και σκέφτηκα ότι έδειχνε πολύ δύσκολο· δεν ξέρω αν καταλαβαίνω». Πολλοί από τους συμμετέχοντες εξέφρασαν την ανησυχία τους για έλλειψη δεξιοτήτων ή έδειξαν ότι δεν διέθεταν τις απαιτούμενες δεξιότητες, είτε επρόκειτο για δεξιότητες οπτικού αλφαβητισμού («visual literacy»), είτε για γλωσσικές δεξιότητες, είτε για μαθηματικές και στατιστικές δεξιότητες (όπως το να γνωρίζουν πώς διαβάζονται συγκεκριμένοι τύποι διαγραμμάτων), είτε για δεξιότητες κριτικής σκέψης.

Συναισθήματα. Αν και τελευταίο στη λίστα μας, ένα σημαντικό εύρημα της έρευνάς μας ήταν ο σημαντικός ρόλος που παίζουν τα συναισθήματα στην αλληλεπίδραση των ανθρώπων με τις οπτικοποιήσεις δεδομένων[119]. Προέκυψε ένα ευρύ φάσμα συναισθημάτων σε σχέση με την αλληλεπίδραση με τις οπτικοποιήσεις, όπως ευχαρίστηση, θυμός, θλίψη, ενοχή, ντροπή, ανακούφιση, ανησυχία, αγάπη, ενσυναίσθηση, ενθουσιασμός, δυσφορία. Οι συμμετέχοντες ανέφεραν συναισθηματικές αντιδράσεις για τις οπτικοποιήσεις εν γένει, για τα δεδομένα που απεικονίζονταν, για το οπτικό ύφος, για το αντικείμενο των οπτικοποιήσεων δεδομένων, για την πηγή ή την αρχική τοποθέτηση των οπτικοποιήσεων, για το επίπεδο των δικών τους δεξιοτήτων κατανόησης των οπτικοποιήσεων.

Για παράδειγμα, δύο επαγγελματίες της κοινωνίας των πολιτών χρησιμοποίησαν έντονη γλώσσα, για να περιγράψουν τα συναισθήματά τους, όταν είδαν τις οπτικοποιήσεις για τη μετανάστευση στο Ηνωμένο Βασίλειο που παρουσιάζονται στην Εικόνα 5.2.2. Τα δεδομένα τούς έβαλαν σε σκέψεις για το πώς μπορεί να αισθάνεται ένας μετανάστης που έρχεται στο Ηνωμένο Βασίλειο και βρίσκεται αντιμέτωπος με τους αντιμεταναστευτικούς πηχυαίους τίτλους των μέσων ενημέρωσης. Περιέγραψαν ότι αισθάνονται «ένοχοι» και ότι «ντρέπονται» που είναι Βρετανοί.

Άλλοι συμμετέχοντες είχαν έντονες συναισθηματικές αντιδράσεις απέναντι στο οπτικό ύφος ορισμένων οπτικοποιήσεων. Μια οπτικοποίηση των εισπράξεων από το κινηματογραφικό box office στους *New York Times* δίχασε τους συμμετέχοντες, με άλλους να έλκονται από την αισθητική της και άλλους να τους απωθεί (Bloch κ.α., 2008):

> Ήταν ευχάριστο να βλέπεις αυτήν την οπτική παρουσίαση, λόγω της συνοχής ανάμεσα στην εικόνα και στο μήνυμα που μεταφέρει.
>
> Εκνευρίστηκα. Καταρχάς, ήταν μια άσχημη απεικόνιση, δυσκολευόσουν να δεις καθαρά, δεν περιείχε πληροφορίες, ένα σκέτο χάλι.

119. Για περισσότερες πληροφορίες σχετικά με τον ρόλο των συναισθημάτων στις αλληλεπιδράσεις με την οπτικοποίηση δεδομένων, βλ. Kennedy, H., & Hill, R. L. (2018) «The feeling of numbers: Emotions in everyday engagements with data and their visualization», Sociology, 52(4), σσ. 830–848. https://doi.org/10.1177/0038038516674675

Εικόνα 5.2.2: Η μετανάστευση στο Ηνωμένο Βασίλειο στην απογραφή. Πηγή: Παρατηρητήριο Μετανάστευσης του Πανεπιστημίου της Οξφόρδης.

Τι σημαίνει αυτό για τη δημιουργία αποτελεσματικών οπτικοποιήσεων

Από την έρευνά μας προέκυψε ένα ευρύ φάσμα αντιλήψεων για το τι καθιστά μια οπτικοποίηση αποτελεσματική. Οι οπτικοποιήσεις στα μέσα ενημέρωσης που απευθύνονται σε μη ειδικούς μπορεί να έχουν ως στόχο να πείσουν, για παράδειγμα. Όλες πρέπει να είναι ελκυστικές, για να προσελκύσουν το κοινό και να το πείσουν να αφιερώσει χρόνο, ώστε να γνωρίσει τα δεδομένα στα οποία βασίζεται η οπτικοποίηση. Οι οπτικοποιήσεις μπορεί να διεγείρουν συγκεκριμένα συναισθήματα, τα οποία εμπνέουν τους ανθρώπους να κοιτάξουν περισσότερο, βαθύτερα ή πιο μακριά. Μπορεί να προκαλέσουν ενδιαφέρον ή το αντίθετο. Μια αποτελεσματική οπτικοποίηση μπορεί να:

- εγείρει ερωτήματα/την επιθυμία συμμετοχής σε συζητήσεις με άλλους
- δημιουργήσει συμπάθεια για τους άλλους ανθρώπους που παρουσιάζονται στα δεδομένα
- προκαλέσει αρκετή περιέργεια, ώστε να προσελκύσει τον χρήστη
- ενισχύσει/υποστηρίξει προϋπάρχουσα γνώση
- προκαλέσει έκπληξη
- πείσει ή να αλλάξει τη γνώμη κάποιου
- παρουσιάσει κάτι νέο
- λειτουργήσει ως ένεση αυτοπεποίθησης όσον αφορά στην κατανόηση των οπτικοιποιήσεων
- παρουσιάσει δεδομένα χρήσιμα για τους προσωπικούς σκοπούς κάποιου
- επιτρέψει την τεκμηριωμένη ή κριτική ενασχόληση με ένα θέμα
- είναι μια ευχάριστη εμπειρία
- προκαλέσει μια έντονη συναισθηματική απόκριση.

Αυτό που καθιστά μια οπτικοποίηση αποτελεσματική είναι κάτι το ρευστό –δεν ισχύει ένας και μοναδικός ορισμός για όλες τις οπτικοποιήσεις. Για παράδειγμα, σε ορισμένα πλαίσια, είναι σκόπιμο η οπτικοποίηση να διασκεδάζει το κοινό, αλλά σε άλλα όχι. Οι οπτικοποιήσεις έχουν διάφορους στόχους: να μεταδώσουν νέα δεδομένα· να ενημερώσουν το ευρύ κοινό· να επηρεάσουν τη λήψη αποφάσεων· να επιτρέψουν τη διερεύνηση και την ανάλυση των δεδομένων· να εκπλήξουν και να επηρεάσουν συμπεριφορές. Οι παράγοντες που επηρεάζουν την αλληλεπίδραση, τους οποίους εντοπίσαμε στην έρευνά μας, θα πρέπει να θεωρηθούν διαστάσεις της αποτελεσματικότητας, οι οποίες έχουν διαφορετική βαρύτητα ανάλογα με την οπτικοποίηση, το πλαίσιο και τους σκοπούς. Πολλοί από αυτούς τους παράγοντες βρίσκονται εκτός του ελέγχου των δημιουργών οπτικοποιήσεων δεδομένων, καθώς σχετίζονται με την κατανάλωση και όχι με την παραγωγή οπτικοποιήσεων. Με άλλα λόγια, η αποτελεσματικότητα μιας οπτικοποίησης εξαρτάται σε μεγάλο βαθμό από το πώς, από ποιον, πότε και πού λαμβάνει χώρα η πρόσβαση σε αυτή. Δυστυχώς, η έρευνά μας δεν προτείνει μια απλή λίστα ελέγχου που να εγγυάται την παραγωγή καθολικά αποτελεσματικών οπτικοποιήσεων. Ωστόσο, εάν θέλουμε προσβάσιμες και αποτελεσματικές οπτικοποιήσεις δεδομένων, είναι σημαντικό οι δημοσιογράφοι που εργάζονται στον τομέα να ασχοληθούν με αυτά τα ευρήματα.

Παραπομπές

Appelgren, E. (2017) «An illusion of interactivity: The paternalistic side of data journalism», *Journalism Practice,* 12(3). DOI: https://doi.org/10.1080/17512786.2017.1299032

Baack, S. (2018) *Knowing what counts: How journalists and civic technologists use and imagine data*. PhD thesis. University of Groningen. Διαθέσιμο σε: https://www.rug.nl/research/portal/files/56718534/Complete_thesis.pdf (Προσπελάστηκε: 1 Δεκεμβρίου 2021).

Bloch, M., Byron, L., Carter, S., & Cox, A. (2008) «Ebb and flow of movies: Box office receipts 1986–2008», *The New York Times,* 23 Φεβρουαρίου. Διαθέσιμο σε: http://archive.nytimes.com/www.nytimes.com/interactive/2008/02/23/movies/20080223_REVENUE_GRAPHIC.html?_r=1 (Προσπελάστηκε: 1 Δεκεμβρίου 2021).

Bradshaw, P. (2017) «No, I'm not abandoning the term "storytelling," Alberto—Just the opposite (and here's why)», *Online Journalism Blog,* 14 Σεπτεμβρίου. Διαθέσιμο σε: https://onlinejournalismblog.com/2017/09/14/narrative-storytelling-data-journalism-alberto-cairo/ (Προσπελάστηκε: 1 Δεκεμβρίου 2021).

Burmester, M., Mast, M., Tille, R., & Weber, W. (2010) «How users perceive and use interactive information graphics: An exploratory study», στο E. Banissi (επιμ.), *Proceedings of the 14th International Conference Information Visualisation.* Λονδίνο, Ηνωμένο Βασίλειο, 26-29 Ιουλίου. σσ. 361-368. DOI: https://doi.org/10.1109/IV.2010.57

Espeland, W. N., & Sauder, M. (2007) «Rankings and reactivity: How public measures recreate social worlds» *American Journal of Sociology*, 113(1), σσ. 1-40. DOI: https://doi.org/10.1086/517897

Rusbridger, A. (2010) «*Why journalism matters», Media Standards Trust Series*. British Academy, Λονδίνο, Ηνωμένο Βασίλειο, 15 Ιουλίου. Διαθέσιμο σε: http://mediastandardstrust.org/wp-content/uploads/downloads/2010/08/Why-Journalism-Matters-Alan-Rusbridger.pdf (Προσπελάστηκε: 1 Δεκεμβρίου 2021).

Οι συγγραφείς

Η έρευνα της Έλεν Κένεντι διατρέχει το τοπίο των ψηφιακών μέσων. Σήμερα επικεντρώνεται στις βιωμένες και οπτικοποιημένες εμπειρίες της δεδομενοποίησης και των συναφών φαινομένων (αλ-

γόριθμοι, τεχνητή νοημοσύνη, μηχανική μάθηση), στις ανισότητες και στις καθημερινές προοπτικές σχετικά με τις «δίκαιες» πρακτικές δεδομένων.

Ο Γουίλιαμ Άλεν είναι ερευνητής στο Κέντρο Μετανάστευσης, Πολιτικής και Κοινωνίας (Centre on Migration, Policy, and Society, COMPAS) της Οξφόρδης, όπου η έρευνά του εξετάζει τις διασταυρώσεις της πολιτικής επικοινωνίας και των δημόσιων στάσεων υπό το πρίσμα της μετανάστευσης και της κινητικότητας.

Ο Μάρτιν Ενγκεμπρέτσεν είναι Καθηγητής Γλώσσας και Επικοινωνίας στο Πανεπιστήμιο του Αγκντέρ, με ιδιαίτερη εξειδίκευση στους τομείς της πολυτροπικής ανάλυσης λόγου και των δημοσιογραφικών σπουδών.

Η Ρόζμαρι Λούσι Χιλ ερευνά το φύλο, τη δημοφιλή μουσική και την πολιτική των οπτικοποιήσεων δεδομένων και είναι συγγραφέας του βιβλίου *Gender, Metal and the Media: Women Fans and the Gendered Experience of Music* (Palgrave, 2016).

Ο Άντι Κερκ είναι ειδικός οπτικοποίησης δεδομένων.

Η Βίμπκε Βέμπερ είναι καθηγήτρια Γλωσσολογίας των Μέσων στο Πανεπιστήμιο Εφαρμοσμένων Επιστημών στη Ζυρίχης (Zurich University of Applied Sciences, ZHAW) και μελετά την οπτικοποίηση δεδομένων, τα γραφικά πληροφοριών, την οπτική σημειωτική, τη δημοσιογραφία μέσω κόμικς, την εικονική πραγματικότητα και την πολυτροπικότητα.

5.3
Σκιτσάροντας με τα Δεδομένα

Mona Chalabi, Jonathan Gray

Περίληψη: Μια συνέντευξη με τη διάσημη δημοσιογράφο δεδομένων Μόνα Τσαλαμπί, η οποία διερευνά την ανάπτυξη και την υποδοχή της πρακτικής της σκιτσογραφίας ως έναν τρόπο για να καταστήσει τα δεδομένα προσιτά. Μια συζήτηση για τα δεδομένα ως μέσο δημιουργίας πλαισίου, τις οπτικές πρακτικές που επιτρέπουν τη σύγκριση των πραγμάτων, τον ρόλο του χιούμορ και της αναλογίας στη δουλειά της, τη δημοσιογραφία δεδομένων ως κοινωνικό σχόλιο και τη σημασία της ανάδειξης της αβεβαιότητας των δεδομένων και της προσωρινότητας της ανάλυσης.

Λέξεις-κλειδιά: σκιτσογραφία με δεδομένα, οπτικοποίηση δεδομένων, αβεβαιότητα, κοινό των δεδομένων, δημοσιογραφία δεδομένων, οπτικές πρακτικές

Τζόναθαν Γκρέι (ΤζΓκ): Πώς ξεκινήσατε να σκιτσάρετε με δεδομένα;

Μόνα Τσαλαμπί (ΜΤ): Όταν δούλευα στο *FiveThirtyEight* ένιωθα ότι το μέσο δεν απευθυνόταν σε αναγνώστες σαν εμένα. Απευθυνόταν σε ένα ελαφρώς διαφορετικό είδος αναγνώστη, χρησιμοποιώντας πολύπλοκα διαδραστικά μοντέλα. Εκείνη την περίοδο ξεκίνησα να σκιτσάρω με δεδομένα, κάτι που μπορούσα να κάνω ενώ καθόμουν στο γραφείο μου. Όταν άρχισα να φτιάχνω αυτά τα σκίτσα, συνειδητοποίησα ότι θα μπορούσαν να είναι ένας αρκετά αποτελεσματικός τρόπος, για να γίνει κατανοητή η αβεβαιότητα των έργων δεδομένων. Θα μπορούσαν να υπενθυμίσουν στον κόσμο ότι ένας άνθρωπος είναι υπεύθυνος για τη λήψη όλων αυτών των αποφάσεων σχεδιασμού. Θα μπορούσαν να είναι αρκετά δημοκρατικά, να επικοινωνούν με τα δεδομένα με έναν τρόπο εφικτό για τον καθένα. Είχα μια αυτοσχέδια στήλη στον *Guardian*, στην οποία παρουσίαζα στον κόσμο κάθε βήμα της διαδικασίας μου. Ήταν διασκεδαστικό το γεγονός ότι ως δημοσιογράφος μπορούσες να μιλήσεις στους ανθρώπους όχι μόνο για το πού βρήκες τα δεδομένα σου, πώς ακριβώς τα επεξεργάστηκες και τι τους έκανες, αλλά μπορούσες, επίσης, να τους δώσεις τη δυνατότητα να τα αναπαράγουν, γκρεμίζοντας το τείχος ανάμεσα σε αυτούς και σε εσένα, με την ελπίδα να δημιουργήσεις νέα είδη προσβασιμότητας, συμμετοχής και σχέσεων με τους αναγνώστες.

ΤζΓκ: Στο βιβλίο εξερευνούμε πώς οι δημοσιογράφοι δεδομένων δεν πρέπει απλώς να αντικατοπτρίζουν και να ενισχύουν τις καθιερωμένες μορφές εξειδίκευσης (π.χ. την επιστήμη των δεδομένων και τις προηγμένες στατιστικές μεθόδους), αλλά πώς μπορούν, επίσης, να προωθήσουν άλλα είδη πρακτικών δεδομένων και κουλτούρας δεδομένων. Θεωρείτε ότι η εργασία σας αφορά εν μέρει την εξεύρεση άλλων τρόπων εργασίας με τα δεδομένα και συσχετισμού με αυτά;

ΜΤ: Δεν διαθέτω πραγματικά προχωρημένες στατιστικές δεξιότητες. Ο τρόπος με τον οποίο συχνά ξεκινάω την ανάλυση δεδομένων είναι μέσω σχετικά απλών υπολογισμών που άλλοι άνθρωποι μπορούν να αναπαράγουν. Κατά κάποιον τρόπο, αυτό καθιστά τα δεδομένα που χρησιμοποιώ πολύ πιο αξιόπιστα. Από ένα σημείο και μετά, με άλλες πιο προηγμένες στατιστικές προσεγγίσεις, θέτεις στους αναγνώστες σου ένα τελεσίγραφο: είτε εμπιστεύεσαι την ανάλυση του δημοσιογράφου είτε όχι. Αυτό δεν έχει να κάνει με το αν εμπιστεύεσαι τις στατιστικές της κυβέρνησης και τον βασικό πολλαπλασιασμό ή όχι. Υπάρχει κάποιο όφελος, αν κάνουμε πράγματα με απλούς υπολογισμούς. Αυτό είναι ένα μεγάλο μέρος της εργασίας και της προσέγγισής μου.

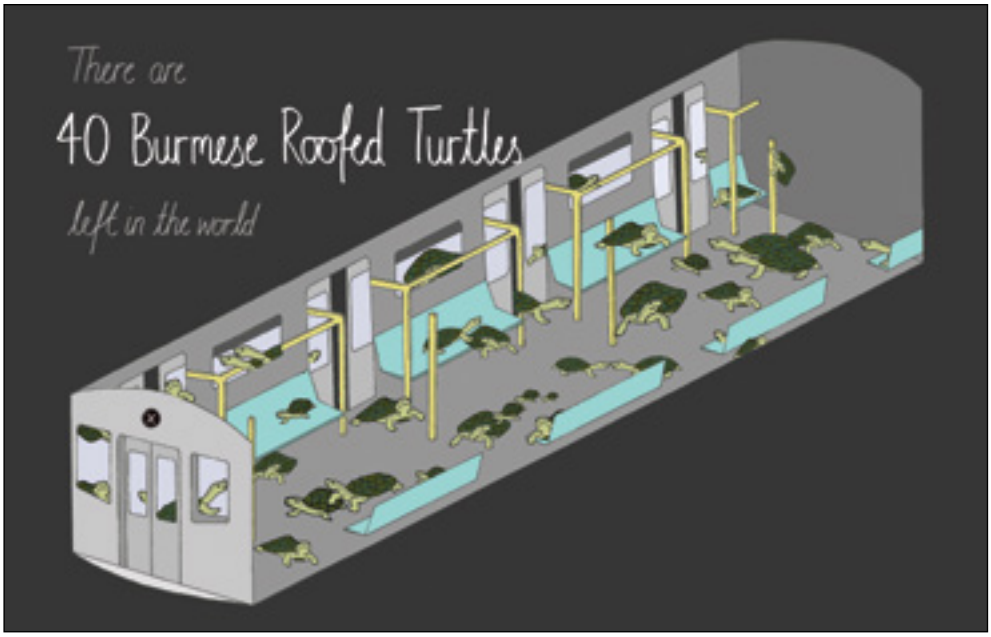
Τα δεδομένα μπορούν να χρησιμοποιηθούν ως ευκαιρία, για να κάνει κανείς δύο διαφορετικά πράγματα: για να εστιάσεις σε κάτι ή για να δείξεις τη «μεγάλη εικόνα». Από τη μία πλευρά, η ευθύνη μου ως δημοσιογράφος δεδομένων είναι να αποεστιάσω από ένα συγκεκριμένο περιστατικό και να δώσω στους αναγνώστες το πλαίσιο, χρησιμοποιώντας δεδομένα. Για παράδειγμα, ας πούμε ότι υπάρχει ένα περιστατικό ή μια επίθεση, θα μπορούσαμε να τους δείξουμε πώς συμβαίνουν αυτές οι επιθέσεις, πού συμβαίνουν, εάν η συχνότητά τους αυξάνεται με την πάροδο του χρόνου και αν υπάρχουν άνθρωποι που στοχοποιούνται περισσότερο από άλλους. Αυτή είναι μια ευκαιρία για τους αναγνώστες να κατανοήσουν ευρύτερες τάσεις, οι οποίες μπορεί να αποδειχθούν πραγματικά ενημερωτικές για εκείνους. Ίσως τους βοηθάει να μη φρίξουν, ή να φρίξουν στον σωστό βαθμό ως αντίδραση στην είδηση.

Εικόνα 5.3.1: Η εικονογράφηση *Average Sentences* («Μέσες Προτάσεις») της Μόνα Τσαλαμπί. Πηγή: *Guardian*.

Από την άλλη πλευρά, μπορούμε να κάνουμε το εντελώς αντίθετο και να εστιάσουμε. Ας υποθέσουμε ότι το BLS [η Στατιστική Υπηρεσία Εργασίας[120] των ΗΠΑ] δημοσιεύει στοιχεία για την ανεργία και ότι τα περισσότερα μέσα ενημέρωσης δημοσιεύουν απλώς το ποσοστό ανεργίας. Εμείς ως δημοσιογράφοι δεδομένων είμαστε σε θέση να εστιάσουμε: μπορούμε να πούμε στους αναγνώστες «ορίστε, να, το εθνικό ποσοστό απασχόλησης και να πώς διαμορφώνεται για τις γυναίκες, τους άνδρες, τις διάφορες ηλικιακές ομάδες, τις διάφορες φυλετικές και εθνοτικές ομάδες». Έτσι, οι αναγνώστες μπορούν να διερευνήσουν τα δεδομένα πιο προσεκτικά.

120. ΣτΜ: Bureau of Labor Statistics

Η δουλειά μου εναλλάσσεται μεταξύ αυτών των δύο τρόπων. Νομίζω ότι μία από τις μεγαλύτερες κριτικές μου για μέσα όπως το *FiveThirtyEight* είναι ότι το έργο τους μπορεί μερικές φορές να αποσκοπεί σε έναν διανοητικό τσαμπουκά: «να τι μπορούμε να κάνουμε». Δεν μου αρέσει αυτό. Σκοπός μου είναι να υπηρετώ τους αναγνώστες και συγκεκριμένα την ευρύτερη κοινότητα των αναγνωστών, όχι μόνο τους λευκούς άνδρες που αυτοπροσδιορίζονται ως geek. Οι αναγνώστες του *FiveThirtyEight* αυτοαποκαλούνται geek και οι δημοσιογράφοι του *FiveThirtyEight* αυτοαποκαλούνται έτσι. Δεν είναι, όμως, αυτός ο λόγος που ασχολήθηκα εγώ με τη δημοσιογραφία.



Εικόνα 5.3.2: Η εικονογράφηση *There are approximately 40 Burmese roofed turtles left in the world* («Έχουν απομείνει περίπου 40 χελώνες Βιρμανίας στον κόσμο») της Μόνα Τσαλαμπί. Πηγή: *Guardian*.

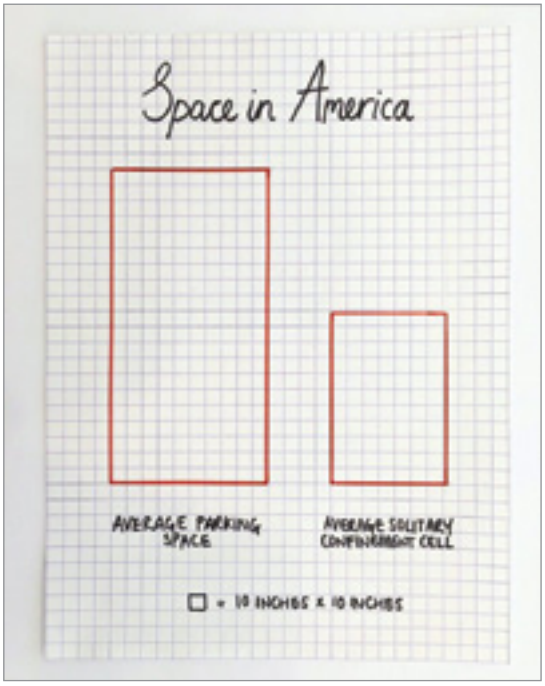
ΤζΓκ: Για να πάρουμε ένα πρόσφατο παράδειγμα της δουλειάς σας, θα μπορούσατε να μας πείτε περισσότερα για το κομμάτι *Endangered Species on a Train* που δημοσιεύτηκε στον *Guardian* (Εικόνα 5.3.2); Πώς ασχοληθήκατε με αυτό το θέμα, πώς προέκυψε το έργο και πώς το προσεγγίσατε;

ΜΤ: Ήταν πραγματικά αρκετά περίεργο. Η έμπνευση στην πραγματικότητα δεν μου ήρθε από την επικαιρότητα. Ήταν περισσότερο η προσωπική μου συνήθεια να κάνω αυτές τις εικονογραφήσεις και η επιθυμία μου να κάνω κάτι πιο φιλόδοξο. Ένας από τους λόγους για τους οποίους άρχισα να κάνω αυτές τις εικονογραφήσεις είναι, επίσης, ότι είναι πολύ αποτελεσματικές και μπορούν να γίνουν μέσα σε λίγες ώρες, εάν χρειαστεί. Ήθελα να δημιουργήσω κάτι μεγαλύτερο που θα απαιτούσε λίγο περισσότερο χρόνο. Ξεκίνησα με ένα πολύ μεγαλύτερο θέμα με το οποίο οι άνθρωποι είναι ήδη εξοικειωμένοι (τα απειλούμενα είδη), για το οποίο η υπάρχουσα οπτική γλώσσα είναι, όμως, ίσως λίγο ανέμπνευστη. Πήρα δεδομένα από τον «Κόκκινο Κατάλογο» της Διεθνούς Ένωσης για τη Διατήρηση της Φύσης («International Union for Conservation of Nature», IUCN)[121]. Για πολλούς από αυτούς τους αριθμούς σχετικά με τα απειλούμενα είδη έδιναν ένα εύρος και εγώ επέλεξα ένα μέσο σημείο για το καθένα από αυτά.

121. https://www.iucnredlist.org/

Κάνοντας ένα βήμα πίσω, θα μπορούσατε να δείτε τις εικονογραφήσεις μου ως διαγράμματα. Αυτό που τις καθιστά διάγραμμα είναι η κλίμακα. Κάθε εικονογράφηση που δημοσιεύω έχει μια αίσθηση κλίμακας, και αυτό κάνουν και τα διαγράμματα. Ένα από τα προβλήματα με την κλίμακα είναι ότι διαφορετικές χώρες και τόποι χρησιμοποιούν διαφορετικές κλίμακες: για παράδειγμα, χιλιοστά στο Ηνωμένο Βασίλειο και ίντσες στις Ηνωμένες Πολιτείες. Οι κλίμακες σημαίνουν διαφορετικά πράγματα για διαφορετικούς ανθρώπους. Πολλοί δημοσιογράφοι δεδομένων το ξεχνούν αυτό. Τι σημαίνει «ένα εκατομμύριο» για κάποιον; Τι σημαίνει το «ένα» για κάποιον; Όλα αυτά εξαρτώνται από το πλαίσιο. Όταν οι αριθμοί είναι μικροί, μπορεί να είναι πιο εύκολο να το καταλάβετε: ξέρετε τι σημαίνει 27. Τι *σημαίνει* όμως αυτό;

Μέρος της ομορφιάς της οπτικοποίησης δεδομένων είναι ότι μπορεί να κάνει τα πράγματα να μοιάζουν πιο ενστικτώδη. Μια άλλη εικονογράφηση για την οποία ήμουν πολύ περήφανη και η οποία τα πήγε πολύ καλά ήταν αυτή όπου συνέκρινα τον μέσο χώρο στάθμευσης με το μέσο κελί απομόνωσης (Εικόνα 5.3.3). Αυτή είναι μια συνήθης πρακτική, όταν πρόκειται για αριθμούς στη δημοσιογραφία: δεν λες «οι τραπεζίτες στο Λονδίνο βγάζουν τόσα», αλλά λες «οι τραπεζίτες στο Λονδίνο βγάζουν 7.000 φορές περισσότερα από όσα βγάζει ένας κοινωνικός λειτουργός». Όλες αυτές οι αναλογίες βοηθούν πραγματικά τους ανθρώπους.



Εικόνα 5.3.3: Η εικονογράφηση *Space in America* («Ο χώρος στην Αμερική»). Πηγή: Μόνα Τσαλαμπί/Instagram.

ΤζΓκ: Φαίνεται ότι μέρος της πρακτικής σας έχει, επίσης, να κάνει με την αντιπαράθεση διαφορετικών στοιχείων (π.χ. το οικείο και το ενοχλητικό). Υπάρχει και εδώ ένα στοιχείο επιμέλειας;

ΜΤ: Το χιούμορ παίζει, επίσης, σημαντικό ρόλο στη δουλειά μου. Όχι ότι τα κομμάτια είναι αστεία, αλλά συχνά υπάρχει κάτι σατιρικό στο στιλ. Η καλύτερη κωμωδία είναι, όταν με τις εικονογραφήσεις ουσιαστικά λέμε «πρόκειται για κωλοκατάσταση». Υπάρχει πάντα κάποιο είδος κοινωνικού σχολιασμού. Εάν μπορείς να εμπλουτίσεις τη δημοσιογραφία των δεδομένων με λίγο από αυτόν, το αποτέλεσμα μπορεί να είναι πραγματικά ισχυρό.

ΤζΓκ: Επιστρέφοντας στο παράδειγμα του *Endangered Species* ως μια περίπτωση που επιτρέπει να γίνουν οι αριθμοί προσιτοί μέσω του χιούμορ και της χρήσης διαφορετικών οπτικών χώρων σύγκρισης, ξεκινήσατε από το βαγόνι (σε αντίθεση με το διάγραμμα);

ΜΤ: Πρώτα σχεδίασα το βαγόνι και μετά σχεδίασα περίπου επτά ή οκτώ ζώα από κάθε είδος. Χρησιμοποίησα το Photoshop, για να διαχωρίσω τα επίπεδα, να τα χρωματίσω και να τα μετρήσω. Για να βεβαιωθώ ότι το έκανα σωστά, το κάθε ζώο βρίσκεται σε ένα διαφορετικό επίπεδο. Η πρώτη μου ιδέα ήταν να ζωγραφίσω απειλούμενα είδη σε διάφορα πράγματα που να είναι όλα οικουμενικά προσιτά. Το μετρό της Νέας Υόρκης δεν είναι τέλειο (είναι μεγαλύτερο ή μικρότερο από το μετρό του Λονδίνου;), αλλά είναι αρκετό για να σου δώσει μια αίσθηση της κλίμακας. Ξεκίνησα με ένα υπολογιστικό φύλλο με διάφορες δυνατότητες που συνδύαζαν απειλούμενα είδη και προσιτούς χώρους. Σκεφτόμουν να δείξω έναν καρχαρία σε μια πισίνα. Με τόσους διαφορετικούς χώρους ήταν, όμως, κάπως δύσκολο να βγάλεις άκρη και, μόλις άρχισα να τους σχεδιάζω, συνειδητοποίησα ότι θα ήταν μια πολύ χρονοβόρα διαδικασία. Αντί να τα ζωγραφίζω όλα σε διαφορετικά μέρη, θα τα έδειχνα όλα στο ίδιο μέρος, το οποίο επίσης λειτουργεί καλύτερα.

Στην πραγματικότητα δεν είναι τέλειο: θα έλεγα ότι είναι λίγο αμφίβολο αν χωράνε όλοι οι ρινόκεροι στην κλίμακα (πολλοί από αυτούς θα έπρεπε να είναι μωρά και όχι ενήλικες!). Σε κάνει, όμως, να νιώθεις κάτι για τους αριθμούς. Και, επίσης, δεν κρύβει τις αδυναμίες του. Όταν κοιτάζετε ένα διάγραμμα του *FiveThirtyEight*, πώς υποτίθεται ότι μπορείτε, ειδικά ως μη ειδικός, να καταλάβετε κατ' ελάχιστον σε ποιο βαθμό είναι ακριβές; Στους αναγνώστες δίνεται απλώς ένα τελεσίγραφο: ή μας εμπιστεύεστε ή όχι. Όταν οι αναγνώστες βλέπουν τις εικονογραφήσεις των απειλούμενων ειδών, μπορούν να κοιτάξουν τους ρινόκερους και να σκεφτούν: «είναι λίγο αλλόκοτο, αλλά το καταλαβαίνω». Έχουν δικαίωμα σε αυτήν την κριτική με τρόπο που δεν συμβαίνει με τα γραφικά που παράγονται από υπολογιστή.

ΤζΓκ: Νωρίτερα αναφέρατε ότι ελπίζατε πως το έργο σας θα μπορούσε να εκδημοκρατίσει τον τρόπο με τον οποίο οι άνθρωποι ασχολούνται με τα δεδομένα. Μπορείτε να μας πείτε κάτι περισσότερο για αυτό;

ΜΤ: Εάν δεν δίνεται στους αναγνώστες η δυνατότητα να συμμετέχουν στη διαδικασία παραγωγής νοήματος από τα δεδομένα, ώστε να σχηματίσουν τη δική τους κρίση, με ποιον τρόπο οι δημοσιογράφοι είναι καλύτεροι από τους πολιτικούς; Υπάρχουν δεξιές και αριστερές εφημερίδες που απλά λένε: «ή μας εμπιστεύεστε ή όχι». Υποτίθεται, όμως, ότι πρέπει να δίνουμε στους ανθρώπους τη δύναμη να λαμβάνουν ενημερωμένες αποφάσεις στην καθημερινή τους ζωή. Η ενδυνάμωση των ανθρώπων δεν είναι μόνο να λες: «αυτά είναι τα γεγονότα, τώρα είναι σαφές ότι πρέπει να πας και να κάνεις αυτό». Είναι να πεις: «αυτά είναι τα γεγονότα· να πώς φτάσαμε ως εδώ». Δεν είναι μόνο η δημοσιογραφία: νομίζω ότι πρέπει να γίνει πολλή δουλειά και στον τομέα της ιατρικής. Θα ήθελα να ασχοληθώ περισσότερο με το πώς μπορούν να αλλάξουν οι ιατρικές συσκευασίες. Αντί για κουτιά που γράφουν «να τι πρέπει να κάνετε», ένας πραγματικά καλός γιατρός θα πρέπει να είναι σε θέση να πει στον ασθενή «αυτοί είναι οι κίνδυνοι αυτού του φαρμάκου. Αυτοί είναι οι κίνδυνοι, εάν δεν το πάρετε. Αυτοί είναι οι κίνδυνοι μιας άλλης φαρμακευτικής αγωγής. Αυτοί είναι οι κίνδυνοι, εάν δεν τη λάβετε», έτσι ώστε οι άνθρωποι να μπορούν να λαμβάνουν οι ίδιοι τις αποφάσεις τους, καθώς κανένας άνθρωπος δεν είναι ίδιος με τον άλλον.

Νομίζω ότι οι καλές οπτικοποιήσεις δεδομένων πρέπει να εκφράζουν την αβεβαιότητα[122]. Η αβεβαιότητα είναι μέρος της όλης διαδικασίας για τη λήψη τεκμηριωμένων αποφάσεων στη ζωή. Ελάχιστοι δημοσιογράφοι δεδομένων αφιερώνουν χρόνο για να εκφράσουν την αβεβαιότητα. Ελάχιστοι δημοσιογράφοι δεδομένων αφιερώνουν χρόνο για να προσεγγίσουν κοινότητες που

122. ΣτΕ του πρωτότυπου έργου: Βλ. επίσης το κεφάλαιο του Άντερσον σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 9.1, «Γενεαλογίες της Δημοσιογραφίας Δεδομένων»), καθώς και το *Apostles of Certainty: Data Journalism and the Politics of Doubt* (Oxford University Press, 2018).

δεν αποτελούνται από geek. Το γεγονός ότι δεν διαθέτετε αυτό το συγκεκριμένο λεξιλόγιο στατιστικών ή υπολογιστικών δεξιοτήτων σημαίνει ότι δεν είστε έξυπνοι, ότι δεν δικαιούστε να κατανοήσετε αυτές τις πληροφορίες; Ασφαλώς και όχι. Και παρ' όλα αυτά, ορισμένοι δημοσιογράφοι δεδομένων αναφέρονται σε τόσους πολλούς από αυτούς τους όρους έτσι αδιάφορα, σαν να λένε: «δεν θα μπαίνω στον κόπο να το εξηγώ κάθε φορά. Ή το καταλαβαίνετε ή δεν το καταλαβαίνετε». Είναι ηλίθιο. Η προσέγγισή μου στη δημοσιογραφία δεδομένων βασίζεται στην ιδέα ότι δεν χρειάζεται απαραίτητα συγκεκριμένο λεξιλόγιο ή εξειδίκευση για να είσαι έξυπνος.

ΤζΓκ: Υπάρχει, επίσης, ένα στοιχείο συμμετοχής των ανθρώπων στη λήψη αποφάσεων για το τι είναι σημαντικό;

ΜΤ: Ένας από τους λόγους για τους οποίους ξεκίνησα τη στήλη συμβουλών *Dear Mona* («Αγαπητή, Μόνα») ήταν για να μπορούν οι άνθρωποι να μου στέλνουν ερωτήσεις. Οι άνθρωποι στέλνουν συνεχώς μηνύματα στο Instagram για ζητήματα που τους απασχολούν και υπάρχουν πολλά πράγματα που δεν θα είχα σκεφτεί ποτέ μόνη μου. Υπάρχουν κάποιοι δρόμοι τους οποίους δεν θέλω να πάρω, όπως η εξέταση της σχέσης μεταξύ ψυχικής υγείας και οπλοκατοχής, η οποία μπορεί να στιγματίσει τα άτομα με προβλήματα ψυχικής υγείας και να ανοίξει τον ασκό του Αιόλου. Ωστόσο, αν λάβει κανείς πολλά μηνύματα από ανθρώπους που θέλουν να μάθουν για αυτό, τότε αναρωτιέται μήπως δεν θα έπρεπε να προσπερνά αυτές τις λεπτές γραμμές, επειδή είναι περίπλοκες, αλλά αντιθέτως θα έπρεπε να προσπαθήσει να τις αντιμετωπίσει κατά μέτωπο. Έτσι, επιζητώ διαρκώς από τους αναγνώστες να μου πουν τι τους ενδιαφέρει. Δεν νομίζω ότι πρόκειται για αποποίηση της δημοσιογραφικής ευθύνης. Είναι κομμάτι του δημοκρατικού ρόλου της δημοσιογραφίας και οι άνθρωποι βλέπουν ότι έχουν μερίδιο στο τελικό προϊόν με κάθε τρόπο: στη διαδικασία της δημιουργίας του, στην κατανόησή του, και δεν είναι κάτι που τους παρέχεται απλά και μόνο με τη λογική «take it or leave it».

ΤζΓκ: Θα μπορούσατε να μας πείτε λίγα λόγια για τις αντιδράσεις στο έργο σας; Έχουν υπάρξει απροσδόκητες ή αξιοσημείωτες αντιδράσεις;

ΜΤ: Εισπράττω έναν σωρό διαφορετικές αντιδράσεις για τη δουλειά μου. Μερικοί άνθρωποι επικεντρώνονται στο θέμα. Έτσι, κάθε φορά που κάνω κάτι σχετικά με τις μισθολογικές διαφορές, για παράδειγμα, έρχονται πολλοί λευκοί άνδρες που λένε: «όχι, οι μαύρες γυναίκες κερδίζουν λιγότερα, επειδή εργάζονται λιγότερο». Τότε, πρέπει να τους μιλήσεις για το πώς οι εικονογραφήσεις βασίζονται σε ομοειδείς συγκρίσεις μεταξύ εργαζομένων πλήρους απασχόλησης και, αν υπάρχουν διαφορές στα επίπεδα στα οποία βρίσκονται (π.χ. ανώτερα διοικητικά στελέχη), αυτό αποτελεί επίσης μέρος του προβλήματος. Είμαι πάντα πρόθυμη να επικεντρωθώ πρώτα στην κριτική.

Ωστόσο, σε γενικές γραμμές λαμβάνω πολύ περισσότερη υποστήριξη παρά κριτική. Μερικές φορές οι άνθρωποι απαντούν στην κριτική στα σχόλια πριν καν φτάσω εγώ σε αυτές. Οι άνθρωποι των οποίων οι ζωές απεικονίζονται στις εικονογραφήσεις παρεμβαίνουν ενίοτε για να πουν: «όχι, η προσωπική μου εμπειρία το επιβεβαιώνει». Ο κόσμος μερικές φορές θέλει να δει επιπλέον δεδομένα. Πολλοί φοιτητές μού γράφουν ότι θέλουν πραγματικά να το κάνουν αυτό (είναι ενδιαφέρον ότι μου γράφουν περισσότερες κοπέλες παρά αγόρια). Πολλές ΜΚΟ και φιλανθρωπικές οργανώσεις μού γράφουν, καθώς θέλουν τα δεδομένα τους να προκαλέσουν συναισθήματα αντί για σκέψεις, και μερικές φορές το έργο μου το επιτυγχάνει αυτό. Γίνεται αναφορά σε ένα από τα έργα μου σε νομοσχέδιο των ΗΠΑ.

Τη δουλειά μου την έχουν δει και την έχουν μοιραστεί πολλοί άνθρωποι στα μέσα κοινωνικής δικτύωσης, οι οποίοι δεν ασχολούνται κατ' ανάγκη με τη δημοσιογραφία δεδομένων, πράγμα που σημαίνει ότι αυτή παρουσιάζεται σε ένα νέο κοινό. Ο Μπέρνι Σάντερς κοινοποίησε την εικονογράφησή μου για την ένοπλη βία, η Μάιλι Σάιρους κοινοποίησε μια άλλη, όπως και η Ιμάν, το μοντέλο, και η Σον Κινγκ, ακτιβίστρια πολιτικών δικαιωμάτων. Δεν πρόκειται για ανθρώπους που γνωρίζω και δεν είναι απαραίτητα άνθρωποι που ακολουθούν τη δουλειά μου, αλλά βλέπουν άλλους ανθρώπους να την κοινοποιούν και, με κάποιον τρόπο, καταλήγει εντός

του οπτικού τους πεδίου. Είναι εκπληκτικό να βλέπεις τον κόσμο να αλληλεπιδρά με τη δουλειά σου. Όταν την κοινοποιήσει κάποιος επιφανής, ορισμένες φορές αποκτά δική της ζωή.

Παραδείγματα των έργων που αναφέρονται σε αυτό το κεφάλαιο μπορείτε να βρείτε στο monachalabi.com και στο Instagram, στον λογαριασμό @monachalabi.

Οι συγγραφείς

Η Μόνα Τσαλαμπί προσπαθεί να ζωντανέψει τους μουδιασμένους αριθμούς.
Ο Τζόναθαν Γκρέι είναι Λέκτορας Σπουδών Κρίσιμων Υποδομών στο Τμήμα Ψηφιακών Ανθρωπιστικών Επιστημών στο King's College του Λονδίνου, συνιδρυτής του Public Data Lab και ερευνητικός συνεργάτης του Digital Methods Initiative στο Πανεπιστήμιο του Άμστερνταμ.

5.4
Ο Παγκόσμιος Ιστός ως Μέσο για την Οπτικοποίηση Δεδομένων

Elliot Bentley

Περίληψη: Εξερευνούμε τα είδη γραφικών που έγιναν εφικτά χάρη στον Ιστό, συμπεριλαμβανομένων διαδραστικών οπτικοποιήσεων, παιχνιδιών και εικονικής πραγματικότητας.

Λέξεις-κλειδιά: διαδραστικά γραφικά, οπτικοποίηση δεδομένων, ανάπτυξη ιστοσελίδων, JavaScript, ενημερωτικά γραφικά, newsgame

Δεν είναι όλα τα μέσα ισότιμα. Μια τηλεοπτική σειρά 20 επεισοδίων μπορεί να αφηγηθεί μια ιστορία διαφορετικά από μια δίωρη ταινία, για παράδειγμα. Με τον ίδιο τρόπο, η ταπεινή ιστοσελίδα μπορεί να παρέχει τις δικές της δυνατότητες για οπτικοποίηση των δεδομένων. Ο Ιστός σχεδιάστηκε αρχικά για απλά έγγραφα με υπερσυνδέσμους που αποτελούνταν κυρίως από κείμενο και στατικές εικόνες. Η προσθήκη της JavaScript και μια αργή είσοδος νέων χαρακτηριστικών και εργαλείων έχει διευρύνει το εύρος της παλέτας που μας διατίθεται για να δουλέψουμε[123].

Αν και σε γραφικά στον Ιστό εξακολουθούν να εφαρμόζονται κυρίως η παραδοσιακή θεωρία και τεχνικές οπτικοποίησης δεδομένων (π.χ. Έντουαρντ Τάφτ, Ζακ Μπερτίν), τα μοναδικά χαρακτηριστικά του Ιστού παρέχουν τεράστιες δυνατότητες για νέες μορφές δημοσιογραφίας δεδομένων. Αυτά τα έργα αναφέρονται συχνά ως «διαδραστικά» («interactives»), μια άχαρη λέξη που αποκρύπτει μερικά από τα μοναδικά πλεονεκτήματα του Διαδικτύου.

Ακολουθεί ένας κατάλογος που αποτυπώνει ορισμένους από τους τρόπους με τους οποίους τα γραφικά στο Διαδίκτυο μπορούν να εκμεταλλευτούν το μέσο που τα φιλοξενεί.

Τεράστια, εξερευνήσιμα σύνολα δεδομένων

Μια κλασική χρήση της διαδραστικότητας είναι να παρουσιάσετε στον αναγνώστη ένα τεράστιο σύνολο δεδομένων και να του επιτρέψετε να «βουτήξει» και να το εξερευνήσει σε όσο βάθος θέλει. Μερικές φορές αυτό παίρνει τη μορφή ενός γιγαντιαίου πίνακα, άλλες φορές ενός μεγάλου διαδραστικού χάρτη.

Στις μέρες μας, συχνά αυτή η μορφή αντιμετωπίζεται αρνητικά, καθώς απαιτεί από τον αναγνώστη να βρει μόνος του τα ενδιαφέροντα σημεία, αλλά μπορεί να παραμείνει πολύτιμη, αν υπάρχει αρκετό «ζουμί» στα δεδομένα. Θεωρώ ότι τα πιο επιτυχημένα έργα τέτοιου τύπου αποδέχονται το γεγονός ότι είναι απλά εργαλεία (σε αντίθεση με τα άρθρα), όπως το εξαιρετικά δημοφιλές College Rankings της *Wall Street Journal* ή οι ενημερωτικές εφαρμογές δημόσιας χρήσης της *ProPublica*[124].

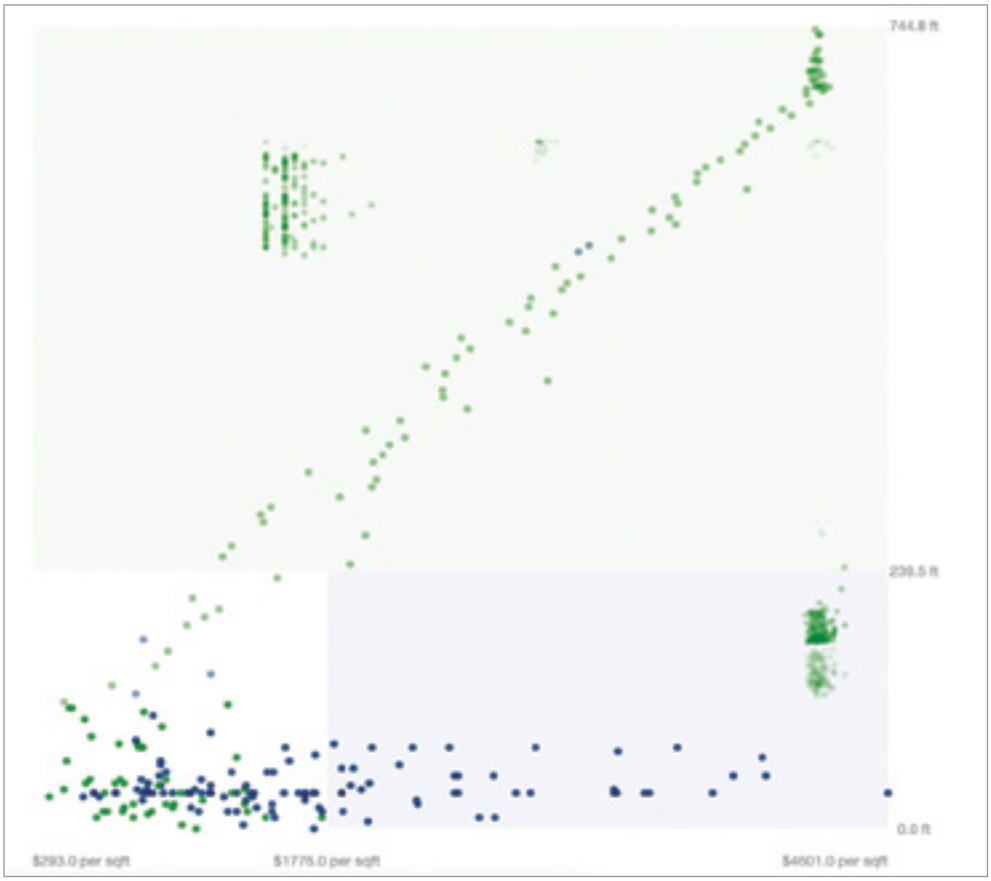
123. Άλλες πλατφόρμες «πολυμέσων», ιδίως το Flash, παρείχαν πληθώρα επιλογών πολύ πριν από το ανοιχτό Διαδίκτυο. Ωστόσο, καλώς ή κακώς, έκτοτε έχουν καταργηθεί σταδιακά. Και ενώ όλες αυτές οι δυνατότητες – και ακόμα περισσότερες μάλιστα– είναι διαθέσιμες σε εγγενείς εφαρμογές, ο Ιστός είναι πολύ πιο εύκολος στη χρήση και η διάθεσή του είναι πρακτικά δωρεάν.

124. https://www.wsj.com/articles/explore-the-full-wsj-the-college-rankings-11567638555, https://www.propublica.org/newsapps/

Καθοδηγώντας τον αναγνώστη μέσα από πολύπλοκα διαγράμματα

Ένα συνηθισμένο φορμάτ πλέον ξεκινά με μόνο ένα διάγραμμα και στη συνέχεια προχωρά στον χειρισμό του (π.χ. κάνοντας ζουμ, ταξιδεύοντας στον χρόνο, εναλλάσσοντας δεδομένα), προκειμένου να μπορέσει κανείς να εξερευνήσει πλήρως το σύνολο των δεδομένων. Αυτό συνδυάζεται εξαιρετικά καλά με το scrollytelling και είναι ιδιαίτερα πολύτιμο στις κινητές συσκευές, όπου μπορεί να μην υπάρχει αρκετός χώρος για να εμφανιστούν όλα τα στοιχεία ενός διαγράμματος ταυτόχρονα[125].

Στο κλασικό πλέον *A Visual Introduction to Machine Learning* («Μια Οπτική Εισαγωγή στη Μηχανική Μάθηση», Εικόνα 5.4.1), τα ίδια σημεία δεδομένων εναλλάσσονται σε πολλαπλά φορμάτ διαγραμμάτων, βοηθώντας τους αναγνώστες να παρακολουθήσουν τον τρόπο με τον οποίο τα ταξινομούν οι αλγόριθμοι μηχανικής μάθησης[126]. Ένα άλλο καλό παράδειγμα είναι το *100 years of tax brackets, in one chart* («100 χρόνια φορολογικές κλίμακες σε ένα διάγραμμα»), ένα κομμάτι του Vox που εστιάζει σε ένα σύνολο δεδομένων το οποίο θα μπορούσε να προκαλέσει αμηχανία στον αναγνώστη, αν παρουσιαζόταν με διαφορετικό τρόπο[127].



Εικόνα 5.4.1: *A Visual Introduction to Machine Learning* («Μια Οπτική Εισαγωγή στη Μηχανική Μάθηση»). Πηγή: R2D3.

125. https://pudding.cool/process/how-to-implement-scrollytelling/
126. http://www.r2d3.us/visual-intro-to-machine-learning-part-1/
127. https://www.vox.com/2015/10/26/9469793/tax-brackets

Ζωντανά δεδομένα μέχρι το τελευταίο δευτερόλεπτο

Γιατί να αρκεστείτε σε ένα στατικό σύνολο δεδομένων, όταν μπορείτε να χρησιμοποιήσετε τους πιο πρόσφατους αριθμούς για οτιδήποτε θέλετε να οπτικοποιήσετε; Οι εκλογές, η κάλυψη αθλητικών γεγονότων, τα καιρικά φαινόμενα και τα οικονομικά δεδομένα είναι προφανείς πηγές ζωντανών δεδομένων, αρκετά ενδιαφέρουσες ώστε να προβάλλονται σε πραγματικό χρόνο. Ακόμα καλύτερη είναι η παρουσίαση του πλαισίου αυτών των ζωντανών στοιχείων με ενδιαφέροντες τρόπους –για παράδειγμα, δείχνοντας ποιες χώρες επωφελούνται από την τρέχουσα τιμή του πετρελαίου (Εικόνα 5.4.2).

Εικόνα 5.4.2: Χώρες που επωφελούνται από την τρέχουσα τιμή του πετρελαίου. Πηγή: *Wall Street Journal*.

Το Ampp3d, ένα βραχύβιο πειραματικό μέσο δημοσιογραφίας δεδομένων, ποπ θεματολογίας, χρησιμοποίησε μετρητές σε πραγματικό χρόνο, για να ζωντανέψει αριθμούς με ενδιαφέροντες τρόπους, όπως τον αριθμό των μεταναστών που εισέρχονται στο Ηνωμένο Βασίλειο και τα κέρδη του ποδοσφαιριστή Γουέιν Ρούνεϊ[128]. Δυστυχώς, έχει κατέβει πια από το Διαδίκτυο.

Τοποθετώντας τον αναγνώστη μέσα σε ένα σύνολο δεδομένων

Μια άλλη εκδοχή της ιδέας των «τεράστιων συνόλων δεδομένων» –που έχω διαπιστώσει ότι είναι απίστευτα ελκυστική για τους αναγνώστες– είναι να δείξετε στον αναγνώστη πού τοποθετείται ο ίδιος μέσα σε ένα σύνολο δεδομένων, συνήθως ζητώντας ορισμένα προσωπικά στοιχεία. Ο χάρτης με το κουίζ διαλέκτων των *New York Times* του 2013 (Εικόνα 5.4.3) αναδείχθηκε ως το πιο δημοφιλές άρθρο της χρονιάς, παρόλο που δημοσιεύτηκε στις 20 Δεκεμβρίου.

128. https://onlinejournalismblog.com/2015/05/13/the-legacy-of-ampp3d-usvsth3m-and-row-zed/

Εικόνα 5.4.3: Ο χάρτης με το κουίζ διαλέκτων των *New York Times* που δημοσιεύτηκε 2013. Πηγή: *New York Times*.

Το BBC φαίνεται να κάνει τέτοιες ενέργειες αρκετά συχνά, πολλές φορές ως εργαλείο προσφοράς δημόσιας υπηρεσίας, με πράγματα όπως ο «μετρητής κλίμακας σωματικού λίπους του Ηνωμένου Βασιλείου»[129]. Μου αρέσει το κομμάτι του Quartz για το πώς σχεδιάζουν κύκλους οι άνθρωποι σε διαφορετικές κουλτούρες, το οποίο ξεκινά ζητώντας από τον αναγνώστη να σχεδιάσει έναν κύκλο, μια συναρπαστική εισαγωγή σε ένα κατά τα άλλα (δυνητικά) βαρετό θέμα[130].

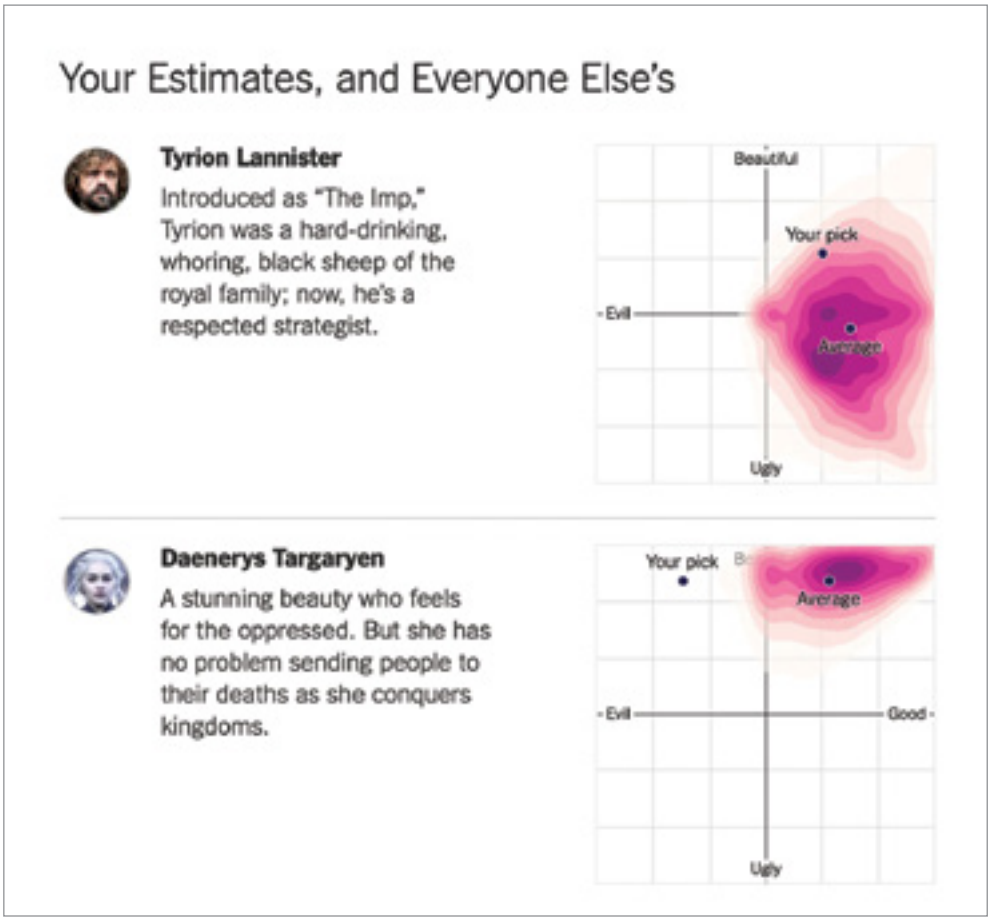
Συγκεντρώνοντας πρωτότυπα σύνολα δεδομένων

Ένα βήμα πέρα από την προηγούμενη κατηγορία είναι τα έργα τα οποία χρησιμοποιούν τα δεδομένα που εισάγουν οι αναγνώστες όχι μόνο για να δώσουν μια άμεση απάντηση, αλλά και για να συγκεντρώσουν ένα νέο σύνολο δεδομένων για περαιτέρω ανάλυση.

Η δημόσια Ραδιοτηλεόραση της Αυστραλίας συνεργάστηκε με πολιτικούς επιστήμονες, για να δημιουργήσει το *Vote Compass* («Πυξίδα Ψήφου») που θα βοηθούσε τους αναγνώστες να κατανοήσουν τη θέση τους στο πολιτικό τοπίο –και στη συνέχεια έγραψε μια σειρά άρθρων με βάση τα δεδομένα[131].

129. https://www.bbc.com/news/health-43697948
130. https://qz.com/994486/the-way-you-draw-circles-says-a-lot-about-you/
131. https://www.abc.net.au/news/nsw-election-2015/vote-compass/, https://www.abc.net.au/news/nsw-election-2015/vote-compass/results/

Πιο πρόσφατα, οι *New York Times* χρησιμοποίησαν την ίδια ιδέα σε ένα πιο χαλαρό θέμα, ζητώντας από τους αναγνώστες να αξιολογήσουν τους χαρακτήρες του *Game of Thrones* και απεικονίζοντας τα αποτελέσματα σε ζωντανά διαγράμματα (Εικόνα 5.4.4).



Εικόνα 5.4.4: Διάγραμμα που αναπαριστά την αξιολόγηση των χαρακτήρων του *Game of Thrones*. Πηγή: The Upshot/*New York Times*.

Ο άπειρος καμβάς

Ο Ιστός είναι άπειρος σε έκταση και χωρητικότητα, αλλά πιο συγκεκριμένα οι ιστοσελίδες μπορούν να είναι όσο φαρδιές ή μακριές θέλουν –ένας «άπειρος καμβάς» πάνω στον οποίο μπορεί κανείς να εργαστεί. Δανείστηκα αυτόν τον όρο από τον καλλιτέχνη Σκοτ Μακλάουντ, ο οποίος υποστηρίζει ότι «δεν υπάρχει κανένας λόγος τα κόμικς μεγάλου μήκους να χωρίζονται σε σελίδες, όταν μεταφέρονται στο Διαδίκτυο»[132]. Και πράγματι, γιατί και τα γραφικά μας να περιορίζονται από τα όρια του χαρτιού;

Στο άρθρο *The Depth of the Problem*[133] («Το Βάθος του Προβλήματος») της *Washington Post*, χρησιμοποιείται ένα γραφικό ύψους 16K πίξελ, για να δείξει το βάθος της περιοχής του ωκε-

132. http://scottmccloud.com/4-inventions/canvas/index.html
133. ΣτΕ: Διακρίθηκε στα βραβεία Information is Beautiful το 2014.

ανού υπό έρευνα για την εξαφανισμένη πτήση MH370 (Εικόνα 5.4.5)[134]. Βεβαίως, αυτές οι πληροφορίες θα μπορούσαν να είχαν στριμωχτεί σε μόνο μία οθόνη, αλλά έτσι θα έλειπε το επίπεδο λεπτομέρειας και ο συναισθηματικός αντίκτυπος αυτού του εξαιρετικά εκτεταμένου γραφικού.

Στο άρθρο του *Guardian*, *How the List Tallies Europe's Migrant Bodycount*, («Πώς ο Κατάλογος Απαριθμεί τον Αριθμό των Νεκρών Μεταναστών στην Ευρώπη»), δεκάδες χιλιάδες θάνατοι μεταναστών αποδίδονται δυναμικά ως μεμονωμένες κουκκίδες που εμφανίζονται μία προς μία καθώς ο αναγνώστης «σκρολάρει» προς τα κάτω στη σελίδα[135].

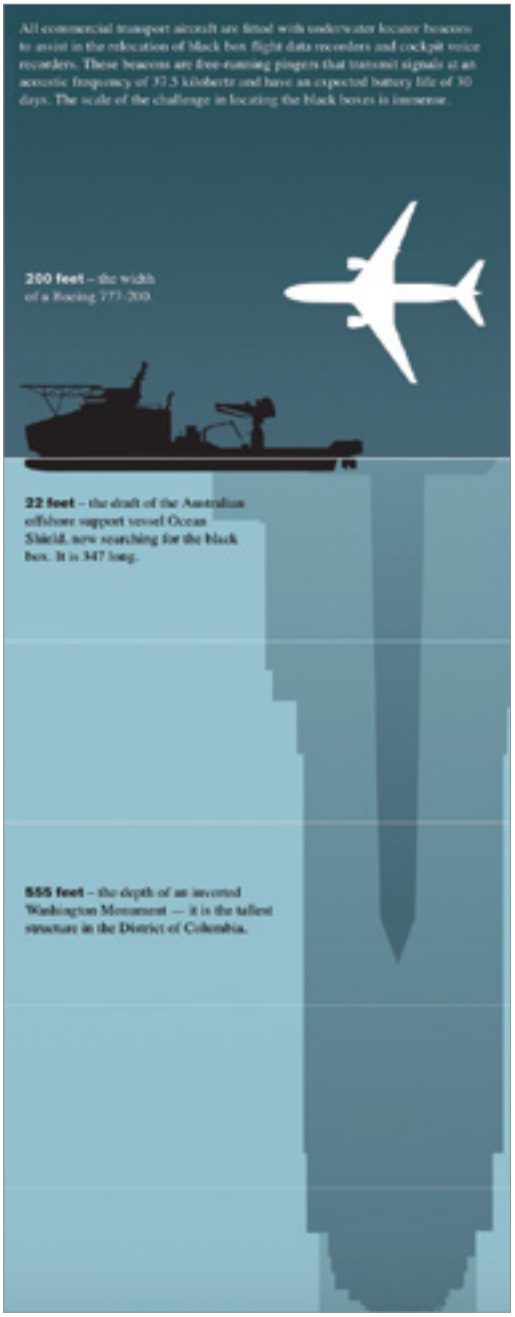
Παιχνίδια με γνώμονα τα δεδομένα

Τα «newsgames», οι διαδραστικές εμπειρίες που δανείζονται μηχανισμούς από τα βιντεοπαιχνίδια για να εξερευνήσουν ειδησεογραφικά θέματα, υπάρχουν εδώ και καιρό, με διαφορετικά επίπεδα επιτυχίας.

Η σειρά *You Draw It* του Upshot (Εικόνα 5.4.6) δοκιμάζει τις υποθέσεις των αναγνωστών, ζητώντας τους να συμπληρώσουν ένα κενό διάγραμμα, πριν να αποκαλύψουν την απάντηση και εξερευνήσουν το θέμα σε μεγαλύτερο βάθος.

Ορισμένα παιχνίδια είναι πιο απαιτητικά, ζητώντας ίσως από τον αναγνώστη να λύσει μια απλοποιημένη εκδοχή ενός πραγματικού προβλήματος –όπως το πώς να χρηματοδοτηθεί το BBC–, για να αποδείξει πόσο δύσκολο είναι[136].

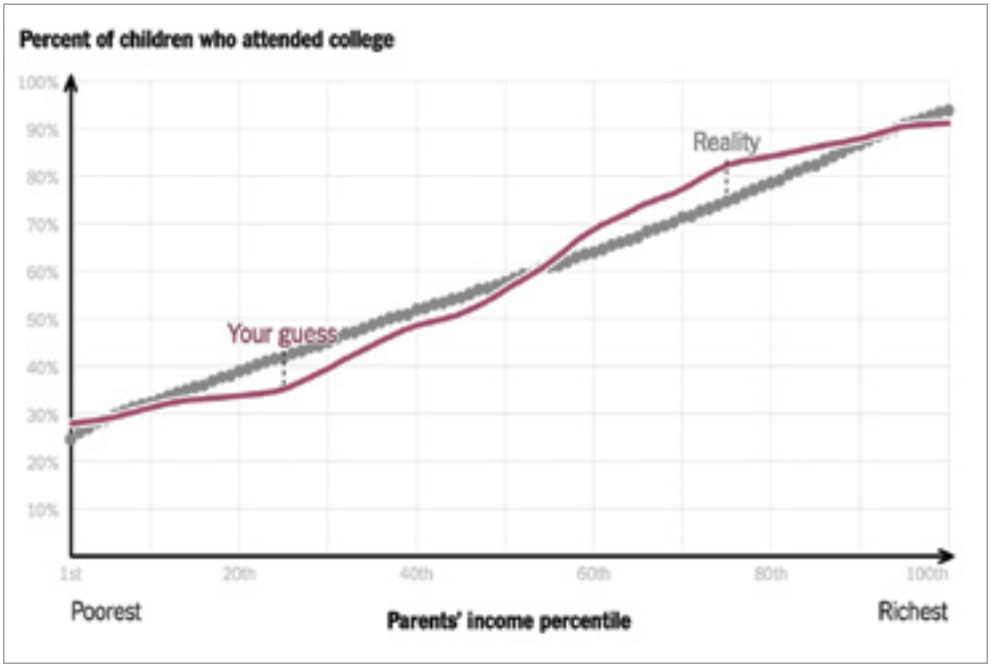
Αυτά θα μπορούσαν να θεωρηθούν παιχνίδια που παρουσιάζουν στον αναγνώστη μόνο επιφανειακές πληροφορίες, αλλά μπορούν να προσφέρουν μια νέα οπτική γωνία σε θέματα που έχουν εξαντληθεί, αν γίνουν σωστά. Το *How to Win a Trade War* («Πώς Κερδίζεται Ένας Εμπορικός Πόλεμος») του *FiveThirtyEight*, στο



Εικόνα 5.4.5: Γραφικό που δείχνει το βάθος της περιοχής του ωκεανού υπό έρευνα για την εξαφανισμένη πτήση MH370. Πηγή: *Washington Post*.

134. https://www.washingtonpost.com/apps/g/page/world/the-depth-of-the-problem/931/
135. https://www.theguardian.com/world/2018/jun/20/the-list-europe-migrant-bodycount
136. https://ig.ft.com/sites/2015/bbc/

οποίο ο αναγνώστης επιλέγει μια εμπορική στρατηγική και ανταγωνίζεται έναν προηγούμενο επισκέπτη της σελίδας, ζωντανεύει την κατά τα άλλα πιθανώς ανιαρή οικονομική θεωρία[137].



Εικόνα 5.4.6: Διάγραμμα από τη σειρά *You Draw It* του Upshot. Πηγή: The Upsot/*New York Times*.

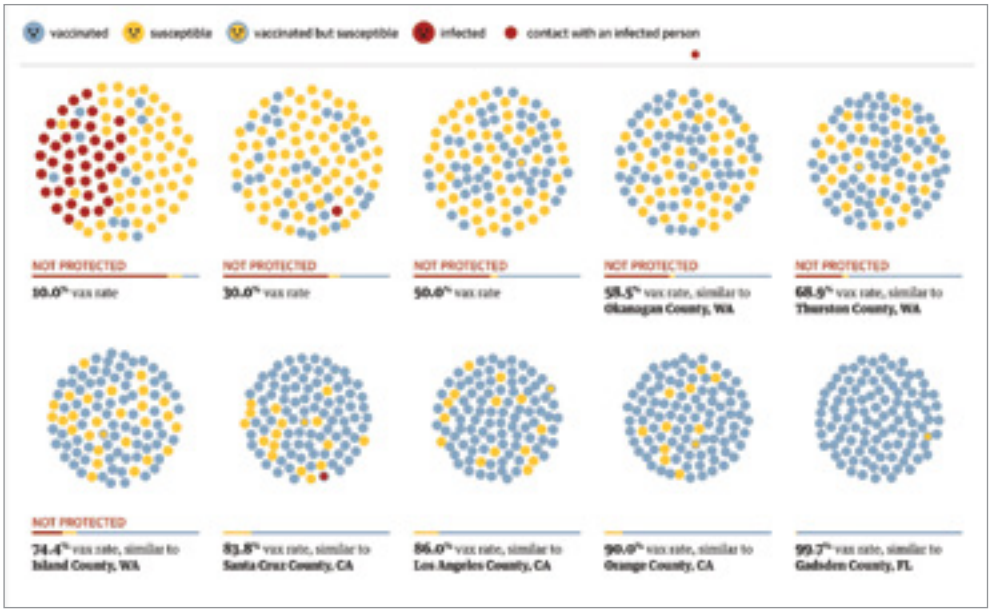
Ζωντανά, τυχαιοποιημένα πειράματα

Σε ένα παρόμοιο φορμάτ, επιτρέπεται στον αναγνώστη να εκτελεί μια ζωντανή προσομοίωση στο πρόγραμμα περιήγησής του. Ως κάτι περισσότερο από μια απλή επεξήγηση κινουμένων σχεδίων, εισάγεται ένας βαθμός τυχαιότητας που οδηγεί σε ένα μοναδικό αποτέλεσμα κάθε φορά και είναι ένας πολύ καλός τρόπος για να δώσει κανείς ζωή στις αφηρημένες στατιστικές πιθανότητες.

Το κομμάτι του *Guardian*, στην Εικόνα 5.4.7, προσομοιώνει ένα ξέσπασμα ιλαράς σε δέκα πληθυσμούς με διαφορετικά ποσοστά εμβολιασμού. Τα διαδικτυακά γραφικά καθιστούν τα αποτελέσματα απολύτως ξεκάθαρα με τρόπο που τα ποσοστά από μόνα τους δεν θα μπορούσαν να το καταφέρουν. Στο έργο του Νέιθαν Γιάου με τίτλο *Years You Have Left to Live, Probably* («Τα Χρόνια Ζωής που Σας Απομένουν, Ίσως»), ένα απλό γραμμικό διάγραμμα (*Probability of living to next year* –«Πιθανότητα να ζήσετε μέχρι την επόμενη χρονιά») γίνεται πιο οδυνηρό με «ζωές» που χάνονται με τυχαίο τρόπο και στη συνέχεια συσσωρεύονται[138].

Αυτές οι προσομοιώσεις δεν χρειάζεται να χρησιμοποιήσουν φανταστικά δεδομένα. Το έργο *The Birthday Paradox* («Το Παράδοξο των Γενεθλίων») ελέγχει την πιθανότητα ίδιων γενεθλίων, χρησιμοποιώντας δεδομένα από προηγούμενους επισκέπτες της σελίδας[139].

137. https://fivethirtyeight.com/features/how-to-win-a-trade-war/
138. https://flowingdata.com/2015/09/23/years-you-have-left-to-live-probably/
139. https://pudding.cool/2018/04/birthday-paradox/



Εικόνα 5.4.7: Προσομοίωση ξεσπάσματος ιλαράς σε δέκα πληθυσμούς με διαφορετικά ποσοστά εμβολιασμού. Πηγή: *Guardian*.

3D, VR και AR

Τα τρισδιάστατα γραφικά (3D) και η εικονική πραγματικότητα («virtual reality», VR) είναι δύσκολο να αξιοποιηθούν στην υπηρεσία της δημοσιογραφίας δεδομένων, εκτός από τους χάρτες μορφολογίας εδάφους.

Δύο αξιοσημείωτα πειράματα, και τα δυο από το 2015 με θέμα τα οικονομικά δεδομένα, το *Is the Nasdaq in Another Bubble?* («Βρίσκεται το Nasdaq σε άλλη μια Φούσκα;») και το *A 3D View of a Chart That Predicts the Economic Future: The Yield Curve* («Μια Τρισδιάστατη Άποψη ενός Διαγράμματος που Προβλέπει το Οικονομικό Μέλλον: Η Καμπύλη Απόδοσης»), είναι έξυπνες καινοτομίες, αλλά δεν κατάφεραν να εκτοξεύσουν τη χρήση των τρισδιάστατων διαγραμμάτων[140]. Ίσως είναι καλύτερα έτσι.

Οι δυνατότητες της επαυξημένης πραγματικότητας («augmented reality», AR), κατά την οποία μια εικόνα του πραγματικού κόσμου (από κάμερα) επικαλύπτεται από γραφικά, δεν έχουν ακόμη εξακριβωθεί.

Συμπέρασμα: Πώς προκύπτουν νέα φορμάτ

Ορισμένα από τα προαναφερθέντα διαδικτυακά γραφικά είναι νέα φορμάτ που εμφανίστηκαν μόλις τα τελευταία χρόνια. Ορισμένα παρέμειναν, όπως η καθοδηγούμενη περιήγηση σε ένα πολύπλοκο διάγραμμα (που συνήθως χρησιμοποιεί ένα μοτίβο αλληλεπίδρασης με scrollytelling). Άλλες, όπως τα τρισδιάστατα διαγράμματα, ήταν απλά πυροτεχνήματα.

Ωστόσο, δεν είναι μόνο το γούστο που καθορίζει ποιοι τύποι γραφικών είναι της μόδας στον

140. http://graphics.wsj.com/3d-nasdaq/, https://www.nytimes.com/interactive/2015/03/19/upshot/3d-yield-curve-economic-growth.html

Ιστό: η διαθέσιμη τεχνολογία και οι καταναλωτικές συνήθειες των αναγνωστών, επίσης, διαμορφώνουν τις τάσεις.

Πάρτε, για παράδειγμα, τον ευρέως χρησιμοποιούμενο διαδραστικό χάρτη. Εκτός του ότι είναι ένα οπτικά ελκυστικό και εύκολα κατανοητό φορμάτ, η διάδοση και η εξοικείωση με αυτό το φορμάτ βοηθήθηκε αναμφίβολα από εργαλεία που καθιστούν εύκολη τη δημιουργία και τον χειρισμό του –Google Maps και Leaflet είναι δύο από τα πιο συνηθισμένα.

Χωρίς να διαθέτουμε αδιάσειστα στοιχεία, τουλάχιστον φαίνεται ότι στις μέρες μας δημοσιεύονται λιγότεροι διαδραστικοί χάρτες. Αν και θα ήταν εύκολο να αποδοθεί αυτή η τάση στην αυξανόμενη συνειδητοποίηση των δημοσιογράφων πως αυτή η διαδραστικότητα (ή ακόμη και ο ίδιος ο χάρτης) μπορεί συχνά να είναι περιττή, οι νέες τεχνολογίες πιθανότατα συνέβαλαν, επίσης, σε αυτήν την πτώση.

Ένα μεγάλο ποσοστό αναγνωστών έχει πλέον πρόσβαση στο Διαδίκτυο μέσω κινητών τηλεφώνων και οι διαδραστικοί χάρτες αποτελούν ιδιαίτερα κακή εμπειρία στις μικρές οθόνες αφής. Επιπλέον, υπάρχει μια νέα τεχνολογική λύση που από πολλές απόψεις είναι ανώτερη: το ai2html, ένα script ανοιχτού κώδικα που διατίθεται από τους *New York Times*, το οποίο παράγει «responsive»[141] .html από αρχεία του Adobe Illustrator[142]. Οι χάρτες που έχουν δημιουργηθεί με το ai2html μπορούν να αξιοποιήσουν τις δεξιότητες ενός παραδοσιακού χαρτογράφου και ταυτόχρονα να περιέχουν ευκρινές, μηχαναγνώσιμο κείμενο. Η έλλειψη διαδραστικότητας σε τέτοιους χάρτες είναι συχνά ευλογία, ακόμη και αν είναι από πολλές απόψεις περιοριστική.

Αυτό είναι μόνο ένα παράδειγμα για το πώς οι δημοσιογράφοι δεδομένων πρέπει να είναι προσεκτικοί στη χρήση των μοναδικών χαρακτηριστικών του Διαδικτύου. Με τόσες πολλές διαθέσιμες δυνατότητες, είναι σημαντικό να τις αξιολογούμε προσεκτικά και να τις χρησιμοποιούμε μόνο όταν είναι πραγματικά απαραίτητες.

Ο συγγραφέας

Ο Έλιοτ Μπέντλεϊ εργάζεται στο τμήμα γραφικών της *Wall Street Journal* από το 2014 και είναι, επίσης, ο δημιουργός της εφαρμογής μεταγραφής ανοιχτού κώδικα oTranscribe.

141. ΣτΜ: Σελίδες .html οι οποίες προσαρμόζονται ανάλογα με τις διαστάσεις της οθόνης του χρήστη.
142. https://github.com/newsdev/ai2html

5.5
Τέσσερις Πρόσφατες Εξελίξεις στα Ενημερωτικά Γραφήματα

Gregor Aisch, Lisa Charlotte Rost

Περίληψη: Αυτό το κεφάλαιο διερευνά τέσσερις εξελίξεις που είδαμε να συμβαίνουν πρόσφατα στα ενημερωτικά γραφήματα: το μότο «mobile first» γίνεται ολοένα και πιο σημαντικό, η σημασία της διαδραστικότητας μετατοπίζεται, αναπτύσσονται περισσότερα (εσωτερικά) εργαλεία σχεδιασμού διαγραμμάτων και οι διαδικτυακές δημοσιεύσεις που επικεντρώνονται στα δεδομένα βρίσκονται σε άνθηση.

Λέξεις-κλειδιά: ενημερωτικά γραφήματα, κινητές συσκευές, εργαλεία σχεδιασμού διαγραμμάτων, διαδραστικότητα, οπτικοποίηση δεδομένων, δημοσιογραφία δεδομένων

Ο τομέας των ενημερωτικών γραφημάτων είναι ακόμη νέος και προσπαθεί να απαντήσει σε ερωτήματα όπως: πώς παρουσιάζουμε τη μεροληψία και την αβεβαιότητα στα (δημοσκοπικά) δεδομένα (Cairo & Schlossberg, 2019); Πώς συνεργαζόμαστε με τους ρεπόρτερ; Πώς μεταδίδουμε πολύπλοκα δεδομένα στα ταχυκίνητα μέσα κοινωνικής δικτύωσης (Segger, 2018); Εδώ, προσπαθούμε να καλύψουμε τέσσερις βασικές εξελίξεις που θεωρούμε ότι θα έχουν σημασία για τα επόμενα χρόνια.

Το μότο «mobile first» αρχίζει να λαμβάνεται σοβαρά υπόψη

Το «mobile first» («πρώτα για κινητές συσκευές») είναι ένα ευρέως χρησιμοποιούμενο τσιτάτο, αλλά στον ταχυκίνητο κόσμο των ενημερωτικών γραφημάτων, οι εμπειρίες σε κινητές συσκευές συχνά περνούσαν σε δεύτερη μοίρα. Τώρα επιτέλους τις βλέπουμε να σκαρφαλώνουν στη λίστα των προτεραιοτήτων. Αυτό έχει δύο συνέπειες.

Πρώτον, υπάρχει μεγαλύτερη μέριμνα, ώστε τα γραφικά να λειτουργούν σε κινητές συσκευές. Ένα μήνυμα που λέει στους χρήστες κινητών συσκευών ότι «αυτή η υπηρεσία λειτουργεί καλύτερα σε σταθερό υπολογιστή» θεωρείται πλέον ατόπημα. Ένα διάγραμμα πρέπει να προσαρμόζεται αποτελεσματικά, έτσι ώστε να μην αναγκάζει περισσότερους από τους μισούς χρήστες να το εγκαταλείπουν. Το να σκέφτεσαι, όμως, μέσα στα λιγοστά πίξελ μιας κινητής συσκευής μπορεί να είναι εκνευριστικό για τους ρεπόρτερ γραφικών («graphics reporters»), πολλοί από τους οποίους έχουν συνηθίσει στην «πολυτέλεια» να γεμίζουν ολόκληρες σελίδες σε έντυπες εφημερίδες και να σχεδιάζουν εμπειρίες πλήρους οθόνης για σταθερούς υπολογιστές. Στην καλύτερη περίπτωση, τα όρια της μικρής οθόνης παρακινούν τους ρεπόρτερ γραφικών να σκεφτούν «έξω από το κουτί» και να γίνουν δημιουργικοί. Αυτό το βλέπουμε ήδη να συμβαίνει: για παράδειγμα, οι *Financial Times* έστρεψαν το διάγραμμα κοινοβουλευτικών εδρών κατά 90 μοίρες, δημιουργώντας ουσιαστικά έναν νέο τύπο διαγράμματος[143].

143. https://ig.ft.com/italy-poll-tracker/

Η δεύτερη συνέπεια της «mobile first» οπτικοποίησης δεδομένων είναι ότι οι προγραμματιστές ενημερωτικών εφαρμογών και οι ρεπόρτερ θα βλέπουν πια το «κινητό» όχι απλά ως μια μικροσκοπική οθόνη αλλά και ως μια συσκευή που είναι γεμάτη με αισθητήρες. Αυτό μπορεί να οδηγήσει σε νέες εμπειρίες δεδομένων. Ο *Guardian* δημιούργησε μια εφαρμογή με την οποία μπορείτε να κάνετε μια εικονική ακουστική περιήγηση στο Ρίο ντε Τζανέιρο, καλύπτοντας το ίδιο μήκος με τον μαραθώνιο που διεξήχθη εκεί το 2016[144]. «Η πρόκληση που σας θέτουμε: συμπληρώστε τα 42,2 χλμ. της διαδρομής μέσα στις επόμενες τρεις εβδομάδες», γράφουν. Η επαυξημένη πραγματικότητα («augmented reality», AR) και η εικονική πραγματικότητα («virtual reality», VR) χρησιμοποιούν τα smartphone μας με παρόμοιο τρόπο και τις βλέπουμε πλέον να εμφανίζονται και στις ειδήσεις.

Η διαδραστικότητα είναι νεκρή, εκτός από όταν δεν είναι

Τα τελευταία χρόνια βλέπουμε τη διαδραστικότητα να χρησιμοποιείται ολοένα και λιγότερο για απλά διαγράμματα. Πλέον προορίζεται για τα μεγαλύτερα έργα που θα δημοσιεύει ένας δημοσιογραφικός οργανισμός κάθε χρόνο. Πάντως, η διαδραστικότητα δεν είναι πλέον απαραίτητη για την επιτυχία. Αίθουσες σύνταξης όπως αυτές των *Financial Times*, του *FiveThirtyEight* και του *National Geographic* έχουν επανειλημμένα δημοσιεύσει διαγράμματα που έγιναν viral, χωρίς να επιτρέπουν στους χρήστες να αλληλεπιδρούν μαζί τους.

Βλέπουμε δύο βασικούς λόγους μείωσης των διαδραστικών γραφικών. Πρώτον, είναι λιγότεροι οι άνθρωποι που αλληλεπιδρούν με τα διαγράμματα απ' ό,τι υποθέταμε στο παρελθόν[145]. Οι φιλοπερίεργοι, έμπειροι στο Διαδίκτυο άνθρωποι –όπως οι ρεπόρτερ γραφικών– θα προσπαθούν πάντα να περάσουν το ποντίκι τους πάνω από μια οπτικοποίηση, περιμένοντας να εμφανιστεί κάτι. Και οι δημοσιογράφοι θέλουν τα άρθρα τους να αποπνέουν μια πιο ζωντανή αίσθηση. Ωστόσο, δημιουργούμε για ένα κοινό που προτιμά την παθητική κατανάλωση· ειδικά στις κινητές συσκευές. Οι περισσότεροι άνθρωποι θα παραλείψουν το περιεχόμενο, αν αυτό είναι κρυμμένο πίσω από τη διαδραστικότητα, γεγονός που οδήγησε πολλούς ρεπόρτερ γραφικών στην απόφαση να μην κρύβουν τίποτα εξαρχής.

Δεύτερον, τα γραφικά έφτασαν στον κύκλο των έκτακτων ειδήσεων. Οι ρεπόρτερ γραφικών έχουν γίνει όλο και πιο γρήγοροι στη δημιουργία οπτικοποιήσεων και σε μια έκτακτη είδηση, για παράδειγμα, θα εμφανιστεί γρήγορα ένας χάρτης εντοπισμού της τοποθεσίας όπου έλαβε χώρα ένα γεγονός. Ωστόσο, η καλοφτιαγμένη διαδραστικότητα παραμένει χρονοβόρα. Συχνά, παραλείπεται για λόγους ταχύτερης δημοσίευσης του άρθρου.

Βλέπουμε ακόμα διαδραστικά ενημερωτικά γραφικά, αλλά η σημασία τους έχει μετατοπιστεί. Αντί να προσθέτει σε μια ιστορία, η διαδραστικότητα γίνεται η ίδια η ιστορία. Έχουμε δει σπουδαία παραδείγματα επεξηγήσεων με δυνατότητα εξερεύνησης, όπου οι αναγνώστες μπορούν να εισάγουν τα προσωπικά τους δεδομένα, όπως την τοποθεσία, το εισόδημα ή τη γνώμη τους, για να δουν στη συνέχεια πώς εντάσσονται σε ένα ευρύτερο σύστημα. Τέτοια παραδείγματα είναι το *You Draw It: How Family Income Predicts Children's College Chances* («Σχεδιάστε το: Πώς το Οικογενειακό Εισόδημα Προβλέπει τις Προοπτικές των Παιδιών στο Κολέγιο») και το *Is It Better to Rent or Buy?* («Καλύτερα το Ενοίκιο ή η Αγορά;») από τους *New York Times*[146]. Και τα δύο κομμάτια δεν έχουν καμία αξία για τους αναγνώστες, αν δεν εισάγουν δεδομένα: η αξία έρχεται μέσα από την αλληλεπίδραση.

144. https://www.theguardian.com/sport/2016/aug/06/rio-running-app-marathon-course-riorun
145. https://vimeo.com/182590214, https://medium.com/@dominikus/the-end-of-interactive-visualizations-52c585dcafcb
146. https://www.nytimes.com/interactive/2015/05/28/upshot/you-draw-it-how-family-income-affects-childrens-college-chances.html, https://www.nytimes.com/interactive/2014/upshot/buy-rent-calculator.html

Οι αίθουσες σύνταξης χρησιμοποιούν περισσότερα (εσωτερικά) εργαλεία οπτικοποίησης

Περισσότερο παρά ποτέ, οι δημοσιογράφοι πιέζονται να κάνουν τα άρθρα τους ξεχωριστά. Η προσθήκη ενός διαγράμματος είναι μια λύση, αλλά οι ομάδες γραφικών δυσκολεύονται να διαχειριστούν τον ολοένα και μεγαλύτερο αριθμό εισερχόμενων αιτημάτων. Αυτός είναι ο λόγος για τον οποίο βλέπουμε όλο και περισσότερες αίθουσες σύνταξης να αποφασίζουν να χρησιμοποιήσουν εργαλεία που διευκολύνουν τη δημιουργία διαγραμμάτων, χαρτών και πινάκων με λίγα κλικ. Μια αίθουσα σύνταξης έχει δύο επιλογές, όταν πρόκειται για εργαλεία οπτικοποίησης: τη χρήση ενός εξωτερικού εργαλείου, όπως το Datawrapper ή το Infogram, ή τη δημιουργία ενός εσωτερικού εργαλείου, προσαρμοσμένου στις εσωτερικές απαιτήσεις και ενσωματωμένου στο σύστημα διαχείρισης περιεχομένου.

Αν και η δεύτερη επιλογή ακούγεται σαν μια εξαιρετική ιδέα, πολλές αίθουσες σύνταξης διαπιστώνουν ότι απαιτεί περισσότερους πόρους από τους αναμενόμενους. Τα εξωτερικά εργαλεία οπτικοποίησης δημιουργούνται από εξειδικευμένες ομάδες που συντηρούν το εργαλείο και προσφέρουν την απαραίτητη εκπαίδευση. Μέσα σε μια αίθουσα σύνταξης, όλα αυτά συχνά γίνονται από την ομάδα γραφικών ή διαδραστικών εφαρμογών, με αποτέλεσμα να τους μένει λιγότερος χρόνος για τα πραγματικά ενημερωτικά έργα. Ένα εσωτερικό εργαλείο οπτικοποίησης μπορεί να είναι επιτυχές μόνο αν αποτελεί προτεραιότητα. Η *Neue Zürcher Zeitung*, για παράδειγμα, διαθέτει τρεις προγραμματιστές που αφιερώνουν τον χρόνο τους αποκλειστικά στην ανάπτυξη και συντήρηση του εργαλείου οπτικοποιήσεων Q.

Οι εκδόσεις με επίκεντρο τα δεδομένα προωθούν την καινοτομία και τον οπτικό αλφαβητισμό

Ενώ η προσέγγιση ενός θέματος με βάση τα δεδομένα θεωρούνταν χρήσιμη μόνο για μεμονωμένες ιστορίες πριν από μερικά χρόνια, τώρα βλέπουμε ολόκληρες (επιτυχημένες!) εκδόσεις να βασίζονται σε αυτήν την ιδέα. Συχνά, αυτοί οι ιστότοποι χρησιμοποιούν τα δεδομένα ως μέσο επικοινωνίας για θέματα που αφορούν συγκεκριμένες δημοσιεύσεις –για παράδειγμα, το *FiveThirtyEight* για την πολιτική και τον αθλητισμό, το *Pudding* για την ποπ κουλτούρα και το *Our World in Data* για τη μακροπρόθεσμη ανάπτυξη της ανθρωπότητας. Ίσως η μεγαλύτερη διαφορά μεταξύ αυτών των εκδόσεων και άλλων για τα ίδια θέματα είναι το κοινό τους: είναι ένα φιλοπερίεργο και προσανατολισμένο στα δεδομένα κοινό, ένα κοινό που δεν φοβάται να δει ένα διάγραμμα. Κατά συνέπεια, οι δημοσιεύσεις με επίκεντρο τα δεδομένα μπορούν να παρουσιάσουν στο αναγνωστικό κοινό τους πιο δυσνόητους τύπους διαγραμμάτων, όπως τα συνδεδεμένα διαγράμματα διασποράς. Εάν χρησιμοποιηθούν σωστά, δίνουν μια πιο σύνθετη, λιγότερο συγκεντρωτική, εικόνα του κόσμου και καθιστούν τις συγκρίσεις ορατές με τρόπο που δεν μπορεί να κάνει ένα ραβδόγραμμα.

Ένα κεφάλαιο που εξετάζει τις πρόσφατες εξελίξεις μπορεί γρήγορα να καταστεί παρωχημένο. Ωστόσο, οι τέσσερις εξελίξεις που καλύψαμε κυριαρχούν στις συζητήσεις εδώ και μερικά χρόνια και αναμένουμε ότι θα παραμείνουν επίκαιρες. Αυτό οφείλεται στο γεγονός ότι στηρίζονται σε ερωτήματα που δεν έχουν μόνο μία σωστή απάντηση στην καθημερινή ειδησεογραφική εργασία: «σχεδιάζουμε ένα έργο πρώτα για κινητές συσκευές ή επιλέγουμε μια πιο σύνθετη λύση που λειτουργεί μόνο σε επιτραπέζιο υπολογιστή;», «καταβάλλουμε προσπάθεια για να κάνουμε αυτήν την οπτικοποίηση διαδραστική και ενδεχομένως πιο ενδιαφέρουσα για τους αναγνώστες (ακόμη και αν μόνο ένα εκτιμώμενο 10-20% από αυτούς θα χρησιμοποιήσει τα διαδραστικά χαρακτηριστικά);», «φτιάχνουμε την οπτικοποίηση από το μηδέν ή χρησιμοποιούμε ένα εργαλείο οπτικοποίησης;», «δημιουργούμε μια οπτικοποίηση για ένα ευρύτερο κοινό ή για ένα κοινό εξοικειωμένο με τα δεδομένα;».

Οι απαντήσεις μπορεί να διαφέρουν ανάλογα με τις αίθουσες σύνταξης, τις ομάδες γραφικών και

τα έργα. Ωστόσο, όλο και περισσότερο, πιστεύουμε ότι οι απαντήσεις θα συγκλίνουν σε «mobile first» διαγράμματα και οπτικοποιήσεις που δεν θα είναι διαδραστικές και θα δημιουργούνται με εργαλεία οπτικοποίησης, για ένα κοινό αυξανόμενα αλφαβητισμένο ως προς τα δεδομένα.

Παραπομπές

Cairo, A., & Schlossberg, T. (2019) «Those hurricane maps don't mean what you think they mean», *The New York Times*, 29 Αυγούστου. Διαθέσιμο σε: https://www.nytimes.com/interactive/2019/08/29/opinion/hurricane-dorian-forecast-map.html (Προσπελάστηκε: 5 Δεκεμβρίου 2021).

Segger, M. (2018) «Lessons for showcasing data journalism on social media», *Medium*, 28 Ιουνίου. Διαθέσιμο σε: https://medium.com/severe-contest/lessons-for-showcasing-data-journalism-on-social-media-17e6ed03a868 (Προσπελάστηκε: 5 Δεκεμβρίου 2021).

Οι συγγραφείς

Ο Γκρέγκορ Άις είναι δημοσιογράφος δεδομένων, μηχανικός λογισμικού και πρώην υπεύθυνος γραφικών στους *New York Times*. Τώρα εργάζεται ως επικεφαλής του τεχνολογικού τμήματος του εργαλείου οπτικοποίησης δεδομένων Datawrapper.
Η Λίζα Σαρλότε Ροστ είναι ντιζάινερ που οπτικοποιούσε δεδομένα για πολλές αίθουσες σύνταξης (*Der Spiegel*, NPR, *Bloomberg*, *ZEIT Online*) πριν να ενταχθεί στο Datawrapper.

5.6
Βάσεις Δεδομένων με Δυνατότητα Αναζήτησης ως Δημοσιογραφικό Προϊόν

Zara Rahman, Stefan Wehrmeyer

Περίληψη: Εξερεύνηση των προκλήσεων για υπευθυνότητα στα δεδομένα και των ευκαιριών για διαφάνεια που προσφέρει η χρήση βάσεων δεδομένων, προσβάσιμων στο κοινό, με δυνατότητα αναζήτησης, στο πλαίσιο μιας έρευνας δημοσιογραφίας δεδομένων.

Λέξεις-κλειδιά: βάσεις δεδομένων, υπευθυνότητα στα δεδομένα, crowdsourcing, αλληλεπίδραση, δημοσιογραφία δεδομένων, διαφάνεια

Ένα, ακόμη αναδυόμενο, δημοσιογραφικό φορμάτ είναι η διαδικτυακή βάση δεδομένων, με δυνατότητα αναζήτησης –μια διαδικτυακή διεπαφή μέσω της οποίας παρέχεται πρόσβαση σε ένα σύνολο δεδομένων, από τις αίθουσες σύνταξης. Το φορμάτ αυτό δεν είναι νέο, αλλά η χρήση του σε έργα δημοσιογραφίας δεδομένων εξακολουθεί να είναι σχετικά σπάνια (Holovaty, 2006).

Σε αυτό το κεφάλαιο, εξετάζουμε ένα εύρος τύπων βάσεων δεδομένων, από αυτές που καλύπτουν θέματα τα οποία επηρεάζουν άμεσα τη ζωή του αναγνώστη, μέχρι διεπαφές που δημιουργούνται στην υπηρεσία περαιτέρω ερευνητικών δραστηριοτήτων. Το έργο μας αντλεί πληροφορίες από το έργο ενός εκ των συγγραφέων στην έρευνα *Euros für Ärzte* («Ευρώ για τους Γιατρούς») του *Correctiv*, η οποία περιγράφεται παρακάτω ως ενδεικτική περιπτωσιολογική μελέτη[147]. Αξίζει, επίσης, να σημειωθεί ότι, αν και έχει καθιερωθεί ως καλή πρακτική η διάθεση των ανεπεξέργαστων δεδομένων μετά από μια έρευνα βασισμένη σε δεδομένα, το βήμα της δημιουργίας μιας διεπαφής με δυνατότητα αναζήτησης στα δεδομένα αυτά είναι σημαντικά λιγότερο συνηθισμένο.

Εξετάζουμε τις ιδιαίτερες δυνατότητες δημιουργίας βάσεων δεδομένων στη δημοσιογραφία, αλλά σημειώνουμε, επίσης, ότι αυτές δημιουργούν μια σειρά από ζητήματα ιδιωτικότητας και ηθικής σχετικά με τον τρόπο χρήσης, πρόσβασης, τροποποίησης και κατανόησης των δεδομένων. Στη συνέχεια, εξετάζουμε ποιες εκτιμήσεις περί υπευθυνότητας στα δεδομένα προκύπτουν ως συνέπεια της χρήσης δεδομένων με αυτόν τον τρόπο, μελετώντας τη δυναμική της εξουσίας που ενυπάρχει, καθώς και τις συνέπειες της τοποθέτησης αυτού του είδους των πληροφοριών στο Διαδίκτυο. Καταλήγουμε προτείνοντας ένα σύνολο βέλτιστων πρακτικών, οι οποίες πιθανότατα θα εξελιχθούν στο μέλλον.

Παραδείγματα δημοσιογραφικών βάσεων δεδομένων

Οι βάσεις δεδομένων μπορούν να αποτελέσουν μέρος της δημόσιας πτυχής της ερευνητικής δημοσιογραφίας με διάφορους τρόπους.

147. https://correctiv.org/recherchen/euros-fuer-aerzte/ (στα γερμανικά)

Ένας τύπος βάσης δεδομένων που έχει ισχυρό το στοιχείο εξατομίκευσης είναι το *Dollars for Docs* της *ProPublica*. Συγκεντρώνει στοιχεία σχετικά με τις πληρωμές προς τους γιατρούς και τα πανεπιστημιακά νοσοκομεία από τις φαρμακευτικές εταιρείες και τις εταιρείες ιατροτεχνολογικών προϊόντων[148]. Αυτό το θέμα και η προσέγγιση υιοθετήθηκε από το *Correctiv* και το *Der Spiegel* για τη δημιουργία του έργου *Euros für Ärzte*, μιας βάσης δεδομένων με δυνατότητα αναζήτησης για τους αποδέκτες πληρωμών από φαρμακευτικές εταιρείες, όπως εξηγείται λεπτομερέστερα παρακάτω. Και οι δύο αυτές προσεγγίσεις περιλάμβαναν τη συγκέντρωση δεδομένων από ήδη διαθέσιμες πηγές. Ο στόχος ήταν να αυξηθεί η προσβασιμότητα στα εν λόγω δεδομένα, ώστε οι αναγνώστες να μπορούν να τα αναζητήσουν οι ίδιοι για να δουν, για παράδειγμα, αν ο δικός τους γιατρός ήταν αποδέκτης πληρωμών. Και τα δύο συνοδεύονταν από ρεπορτάζ και συνεχιζόμενες έρευνες.

Σε παρόμοια κατεύθυνση, η *Berliner Morgenpost* δημιούργησε το *Schul Finder*, για να βοηθήσει τους γονείς να βρουν σχολεία στην περιοχή τους. Σε αυτήν την περίπτωση, η ίδια η διεπαφή της βάσης δεδομένων είναι το κύριο προϊόν[149].

Σε αντίθεση με το είδος της βάσης δεδομένων όπου τα δεδομένα συλλέγονται και προετοιμάζονται από την αίθουσα σύνταξης, ένα άλλο είδος είναι αυτό όπου οι αναγνώστες μπορούν να συνεισφέρουν στα δεδομένα, μερικές φορές γνωστό ως «δεδομένα που παράγονται από τους πολίτες» («citizen-generated data») ή απλά ως πληθοπορισμός («crowdsourcing»). Αυτό είναι ιδιαίτερα αποτελεσματικό, όταν τα απαιτούμενα δεδομένα δεν συλλέγονται από επίσημες πηγές, όπως στη συμμετοχική βάση δεδομένων *The Counted* του *Guardian* που συγκέντρωσε πληροφορίες για τους ανθρώπους οι οποίοι σκοτώθηκαν από την αστυνομία στις Ηνωμένες Πολιτείες το 2015-2016[150]. Η βάση δεδομένων τους χρησιμοποίησε διαδικτυακές δημοσιεύσεις καθώς και πληροφορίες από αναγνώστες.

Ένας άλλος τύπος βάσης δεδομένων περιλαμβάνει τη λήψη ενός υπάρχοντος συνόλου δεδομένων και τη δημιουργία μιας διεπαφής που επιτρέπει στον αναγνώστη να παράγει μια έκθεση αναφοράς με βάση τα κριτήρια που έχει ορίσει κάθε φορά. Για παράδειγμα, τα *Nauru Files* επιτρέπουν στους αναγνώστες να προβάλλουν μια περίληψη των αναφορών περιστατικών που συντάχθηκαν από το προσωπικό του κέντρου κράτησης του Ναουρού στην Αυστραλία μεταξύ 2013 και 2015[151]. Ο οργανισμός *The Bureau of Investigative Journalism*, με έδρα το Ηνωμένο Βασίλειο, συγκεντρώνει δεδομένα από διάφορες πηγές που συλλέγονται μέσω των ερευνών του, σε μια βάση δεδομένων που ονομάζεται *Drone Warfare*[152]. Η βάση δεδομένων επιτρέπει στους αναγνώστες να επιλέξουν συγκεκριμένες χώρες που καλύπτονται και το χρονικό πλαίσιο, προκειμένου να δημιουργήσουν μια έκθεση με οπτικοποιήσεις που συνοψίζουν τα δεδομένα.

Τέλος, βάσεις δεδομένων μπορούν, επίσης, να δημιουργηθούν για την περαιτέρω εξυπηρέτηση της δημοσιογραφίας, ως εργαλείο για την υποστήριξη της έρευνας. Η Διεθνής Κοινοπραξία Ερευνητών Δημοσιογράφων («International Consortium of Investigative Journalists», ICIJ) δημιούργησε και διατηρεί τη βάση δεδομένων *Offshore Leaks*, η οποία αντλεί δεδομένα από τα *Panama Papers*, τα *Paradise Papers* και άλλες έρευνες[153]. Παρομοίως, το Organized Crime and Corruption Reporting Project (OCCRP) διατηρεί και ενημερώνει τα δεδομένα, τα οποία επιτρέπουν στους επισκέπτες να αναζητήσουν περισσότερα από 19 εκατομμύρια δημόσια αρχεία[154].

148. https://projects.propublica.org/docdollars/
149. https://interaktiv.morgenpost.de/schul-finder-berlin/#/
150. https://www.theguardian.com/us-news/ng-interactive/2015/jun/01/the-counted-police-killings-us-database
151. https://www.theguardian.com/australia-news/ng-interactive/2016/aug/10/the-nauru-files-the-lives-of-asylum-seekers-in-detention-detailed-in-a-unique-database-interactive
152. https://www.thebureauinvestigates.com/projects/drone-war/
153. https://offshoreleaks.icij.org/
154. https://data.occrp.org/

Και στις δύο περιπτώσεις, ο κύριος χρήστης των εργαλείων δεν προβλέπεται να είναι ο μέσος αναγνώστης αλλά, αντιθέτως, δημοσιογράφοι και ερευνητές που προβλέπεται να διεξάγουν περαιτέρω έρευνα σχετικά με τις όποιες πληροφορίες προκύπτουν με τη χρήση αυτών των εργαλείων.

Η παρακάτω λίστα συνοψίζει τις διάφορες εκτιμήσεις για τη δημιουργία βάσεων δεδομένων ως ειδησεογραφικό προϊόν:

- **Κοινό:** απευθύνεται απευθείας στους αναγνώστες ή ως ερευνητική βάση δεδομένων για άλλους δημοσιογράφους
- **Επικαιρότητα:** ενημέρωση σε συνεχή ροή ή ως μεμονωμένη δημοσίευση
- **Πλαίσιο:** μέρος μιας έρευνας ή μιας ιστορίας ή η ίδια η βάση δεδομένων ως το κύριο προϊόν
- **Διαδραστικότητα:** οι αναγνώστες ενθαρρύνονται να συνεισφέρουν ενεργά στη βελτίωση της βάσης δεδομένων ή οι αναγνώστες θεωρούνται κυρίως θεατές των δεδομένων.
- **Πηγές:** χρήση ήδη δημόσιων δεδομένων ή δημοσιοποίηση νέων πληροφοριών μέσω της βάσης δεδομένων

Περιπτωσιολογική μελέτη: *Euros für Ärzte* («Ευρώ για τους Γιατρούς»)

Η Ευρωπαϊκή Ομοσπονδία Φαρμακευτικών Βιομηχανιών και Ενώσεων («European Federation of Pharmaceutical Industries and Associations», EFPIA) είναι μια επαγγελματική ένωση που αριθμεί 33 εθνικές ενώσεις και 40 φαρμακευτικές εταιρείες μεταξύ των μελών της. Το 2013, η ομοσπονδία αποφάσισε ότι, αρχής γενομένης από τον Ιούλιο του 2016, οι εταιρείες-μέλη θα πρέπει να δημοσιεύουν τις πληρωμές προς τους επαγγελματίες υγείας και τους οργανισμούς στις χώρες όπου δραστηριοποιούνται (EFPIA, 2013). Εμπνευσμένη από την εφαρμογή *Dollars for Docs* της *ProPublica*, ο μη κερδοσκοπικός γερμανικός ερευνητικός ειδησεογραφικός οργανισμός *Correctiv* αποφάσισε να συλλέξει αυτές τις δημοσιεύσεις από τους ιστοτόπους των γερμανικών φαρμακευτικών εταιρειών και να δημιουργήσει μια κεντρική, αναζητήσιμη βάση δεδομένων των αποδεκτών πληρωμών από φαρμακευτικές εταιρείες, για δημόσια χρήση. Ονόμασαν την έρευνα *Euros für Ärzte* («Ευρώ για τους Γιατρούς»).

Σε συνεργασία με τον γερμανικό εθνικής εμβέλειας μέσο *Der Spiegel*, συγκεντρώθηκαν έγγραφα και δεδομένα από περίπου 50 ιστοτόπους και μετατράπηκαν από διαφορετικά φορμάτ σε συνεπή, πινακοποιημένα δεδομένα. Τα δεδομένα αυτά υποβλήθηκαν σε επιπλέον καθαρισμό και ταυτοποιήθηκαν οι παραλήπτες πληρωμών από πολλαπλές εταιρείες. Ο συνολικός χρόνος για τον καθαρισμό των δεδομένων ήταν περίπου δέκα ημέρες και στη διαδικασία συμμετείχαν έως και πέντε άτομα. Το *Correctiv* σχεδίασε και δημοσίευσε μια προσαρμοσμένη διεπαφή αναζήτησης σε βάση δεδομένων, με μεμονωμένες διευθύνσεις URL ανά παραλήπτη[155]. Η βάση δεδομένων ενημερώθηκε το 2017 με μια παρόμοια διαδικασία. Το *Correctiv* χρησιμοποίησε, επίσης, την ίδια μεθοδολογία και τη διαδικτυακή διεπαφή για να δημοσιεύσει δεδομένα από την Αυστρία, σε συνεργασία με το derStandard.at και το δίκτυο ORF, και δεδομένα από την Ελβετία σε συνεργασία με το Beobachter.ch.

Ο δημοσιογραφικός στόχος ήταν να επισημανθεί η συστηματική επιρροή της φαρμακευτικής βιομηχανίας στους επαγγελματίες υγείας μέσω εκδηλώσεων και οργανώσεων, καθώς και οι σχετικές συγκρούσεις συμφερόντων. Η βάση δεδομένων με δυνατότητα αναζήτησης είχε ως στόχο να ενθαρρύνει τους αναγνώστες να ξεκινήσουν μια συζήτηση με τον γιατρό τους πάνω σε αυτό το θέμα και να επιστήσει την προσοχή στο ίδιο το γεγονός ότι συμβαίνει αυτό.

155. https://correctiv.org/thema/aktuelles/euros-fuer-aerzte/

Σε ένα διαφορετικό επίπεδο, η πρωτοβουλία ανέδειξε, επίσης, την ανεπάρκεια των κανόνων εθελοντικής δημοσιοποίησης. Επειδή η προϋπόθεση δημοσίευσης ήταν μια πρωτοβουλία του κλάδου, και όχι μια νομική υποχρέωση, η βάση δεδομένων ήταν ελλιπής –και είναι απίθανο αυτό να αλλάξει χωρίς νομικά επιβεβλημένη δημοσιοποίηση.

Όπως περιγράφηκε παραπάνω, η βάση δεδομένων ήταν ελλιπής, πράγμα που σημαίνει ότι ένας αριθμός ατόμων που είχαν λάβει πληρωμές από φαρμακευτικές εταιρείες έλειπε από τη βάση δεδομένων. Κατά συνέπεια, όταν οι χρήστες αναζητούν τον γιατρό τους, ένα κενό αποτέλεσμα μπορεί να σημαίνει είτε ότι ο γιατρός δεν έλαβε καμία πληρωμή είτε ότι αρνήθηκε τη δημοσίευση –δύο εντελώς διαφορετικά συμπεράσματα. Η κριτική σημειώνει πως αυτή η διαδικασία τοποθετεί στο επίκεντρο του ενδιαφέροντος όσους συνεργάζονται με διαφάνεια, αφήνοντας στο σκοτάδι πιθανώς πιο σκανδαλώδεις χρηματικές συναλλαγές. Για να το αντισταθμίσει αυτό, το *Correctiv* παρείχε τη δυνατότητα επιλογής για τους γιατρούς που δεν είχαν λάβει πληρωμές να εμφανίζονται επίσης στη βάση δεδομένων, γεγονός που παρέχει σημαντικό πλαίσιο στην αφήγηση, αλλά εξακολουθεί να αφήνει αβεβαιότητα στο αποτέλεσμα της αναζήτησης.

Μετά τη δημοσίευση, τόσο το *Correctiv* όσο και το *Der Spiegel* έλαβαν δεκάδες καταγγελίες και νομικές απειλές από γιατρούς που εμφανίζονταν στη βάση δεδομένων. Καθώς τα δεδομένα προέρχονταν από δημόσιες, αν και δυσεύρετες, πηγές, η νομική ομάδα του *Der Spiegel* αποφάσισε να παραπέμψει τις περισσότερες καταγγελίες στις φαρμακευτικές εταιρείες και να προσαρμόσει τη βάση δεδομένων μόνο σε περίπτωση αλλαγών στην πηγή.

Τεχνικές παράμετροι της δημιουργίας βάσεων δεδομένων

Για μια αίθουσα σύνταξης που μελετά τον τρόπο με τον οποίο θα καταστήσει ένα σύνολο δεδομένων διαθέσιμο και προσβάσιμο στους αναγνώστες, υπάρχουν διάφορα κριτήρια που πρέπει να εξετάσει, όπως το μέγεθος και η πολυπλοκότητα του συνόλου δεδομένων, η εσωτερική τεχνική δυναμικότητα της αίθουσας σύνταξης και ο τρόπος με τον οποίο οι αναγνώστες θα πρέπει να είναι σε θέση να αλληλεπιδρούν με τα δεδομένα.

Όταν μια αίθουσα σύνταξης αποφασίζει ότι μια βάση δεδομένων θα μπορούσε να είναι το κατάλληλο προϊόν μιας έρευνας, η δημιουργία μιας τέτοιας βάσης δεδομένων απαιτεί ανάπτυξη και εγκατάσταση κατά παραγγελία –ένα όχι αμελητέο μέγεθος πόρων. Η πρόσβαση σε αυτά τα δεδομένα μέσω τρίτων υπηρεσιών συνήθως είναι απλούστερη και απαιτεί λιγότερους πόρους.

Για παράδειγμα, στην περίπτωση του *Correctiv*, η ανάγκη αναζήτησης και καταγραφής περίπου 20.000 παραληπτών και των οικονομικών τους συνδέσεων με φαρμακευτικές εταιρείες απαιτούσε μια προσαρμοσμένη λύση λογισμικού. Το *Correctiv* ανέπτυξε το λογισμικό για τη βάση δεδομένων σε ξεχωριστό αποθετήριο δεδομένων από τον κύριο ιστότοπό του, αλλά με τρόπο που θα μπορούσε να συνδεθεί με το σύστημα διαχείρισης περιεχομένου. Η απόφαση αυτή ελήφθη για να επιτραπεί η οπτική και εννοιολογική ενσωμάτωση στον κύριο δικτυακό τόπο και στην ενότητα των ερευνών. Για τον διαχωρισμό ενδιαφερόντων («separation of concerns»)[156], τα δεδομένα αποθηκεύτηκαν σε σχεσιακή βάση δεδομένων ξεχωριστά από τη βάση δεδομένων περιεχομένου. Σε αυτήν την περίπτωση, η ύπαρξη μιας διαδικασίας και μιας διεπαφής για την προσαρμογή των καταχωρίσεων στη βάση δεδομένων ήταν ζωτικής σημασίας, καθώς δεκάδες διορθώσεις των πρωτογενών δεδομένων ήρθαν μετά τη δημοσίευση.

Ωστόσο, μικρότερα σύνολα δεδομένων με απλές δομές μπορούν να καταστούν προσβάσιμα χωρίς δαπανηρά έργα ανάπτυξης λογισμικού. Ορισμένα τρίτα εργαλεία υπολογιστικών φύλλων (π.χ. Google Sheets) επιτρέπουν την ενσωμάτωση πινάκων. Υπάρχουν επίσης πολυάριθμες

156. ΣτΕ: Όρος που στην ανάπτυξη λογισμικών και συστημάτων δηλώνει ότι διαφορετικά τμήματα του συστήματος/διαφορετικές τεχνολογίες εστιάζουν στην επίλυση διαφορετικών ζητημάτων.

βιβλιοθήκες JavaScript για το frontend, για τον εμπλουτισμό των πινάκων .html με λειτουργίες αναζήτησης, φιλτραρίσματος και ταξινόμησης, οι οποίες συχνά είναι αρκετές, για να καταστήσουν μερικές εκατοντάδες σειρές προσβάσιμες στους αναγνώστες.

Μια ελκυστική μέση λύση για την πρόσβαση σε μεγαλύτερα σύνολα δεδομένων είναι οι διαδικτυακές εφαρμογές που βασίζονται σε JavaScript και έχουν πρόσβαση στο σύνολο δεδομένων μέσω API. Ένα τέτοιου είδους «σετάρισμα» προσφέρεται για την εκτέλεση διεπαφών αναζήτησης με ενσωμάτωση iframe, χωρίς να αναλαμβάνει τη δέσμευση για μια ολοκληρωμένη διαδικτυακή εφαρμογή. Το API μπορεί, στη συνέχεια, να εκτελεστεί μέσω τρίτων υπηρεσιών, ενώ εξακολουθεί να έχει τον πλήρη έλεγχο της διαμόρφωσης του frontend.

Προσφερόμενες δυνατότητες των βάσεων δεδομένων

Οι βάσεις δεδομένων εντός μιας ιστορίας ή στο πλαίσιο αυτής παρέχουν μια σειρά από δυνατότητες τόσο για τους αναγνώστες όσο και για τις αίθουσες σύνταξης.

Από την πλευρά του αναγνώστη, η διάθεση μιας διαδικτυακής βάσης δεδομένων του επιτρέπει να αναζητήσει τη δική του πόλη, τον δικό του πολιτικό ή γιατρό και να συνδέσει την ιστορία με τη δική του ζωή. Προσφέρει έναν διαφορετικό δίαυλο εμπλοκής με μια ιστορία, σε ένα πιο προσωπικό επίπεδο. Εφόσον υπάρχουν αναλυτικά στοιχεία σχετικά με τις αναζητήσεις των χρηστών, αυτό παρέχει, επίσης, στην αίθουσα σύνταξης περισσότερα δεδομένα για το τι ενδιαφέρει τους αναγνώστες –πιθανώς προσφέροντας περισσότερα στοιχεία για μελλοντικά έργα.

Από την πλευρά της αίθουσας σύνταξης, εάν η βάση δεδομένων θεωρείται μια μακροχρόνια επένδυση για την έρευνα, μπορεί να χρησιμοποιηθεί για την αυτόματη διασταύρωση καταχωρίσεων με άλλες βάσεις δεδομένων ή σύνολα εγγράφων για τη δημιουργία νέων στοιχείων. Παρομοίως, εάν ή όταν άλλες αίθουσες σύνταξης αποφασίσουν να διαθέσουν στο κοινό παρόμοιες βάσεις δεδομένων, η συνεργασία και η αύξηση της κάλυψης γίνεται πολύ πιο εύκολη, ενώ παράλληλα επαναχρησιμοποιούνται οι υπάρχουσες υποδομές και μεθοδολογίες.

Οι βάσεις δεδομένων προσφέρουν, επίσης, δυνητικά αυξημένη βελτιστοποίηση για τις μηχανές αναζήτησης, φέρνοντας έτσι μεγαλύτερη επισκεψιμότητα στον ιστότοπο του μέσου. Όταν η βάση δεδομένων παρέχει μεμονωμένες διευθύνσεις URL για τις καταχωρήσεις που περιέχει, οι μηχανές αναζήτησης θα εντοπίσουν αυτές τις σελίδες και θα τις κατατάξουν ψηλά στα αποτελέσματά τους για αναζητήσεις με σπάνιες λέξεις-κλειδιά που σχετίζονται με αυτές τις πολυάριθμες καταχωρήσεις (οι λεγόμενες «long tail» αναζητήσεις)[157], αυξάνοντας έτσι την επισκεψιμότητα στον ιστότοπο του οργανισμού.

Η βελτιστοποίηση για τις μηχανές αναζήτησης μπορεί να θεωρηθεί μια ανούσια πρακτική στο πλαίσιο της δημοσιογραφίας. Ωστόσο, η παροχή στους αναγνώστες δημοσιογραφικών πληροφοριών κατά την αναζήτηση συγκεκριμένων θεμάτων μπορεί, επίσης, να θεωρηθεί ως μέρος της επιτυχημένης αλληλεπίδρασης με το κοινό. Αν και ο στόχος της δημόσιας βάσης δεδομένων δεν θα πρέπει να είναι ο ανταγωνισμός ως προς τις λέξεις-κλειδιά για αναζήτηση, είναι πιθανό να αποτελέσει ένα ευπρόσδεκτο πλεονέκτημα που θα φέρει οργανική επισκεψιμότητα και μπορεί με τη σειρά του να προσελκύσει νέο αναγνωστικό κοινό.

Προβληματισμοί σχετικά με την υπευθυνότητα στα δεδομένα

Με βάση την προσέγγιση των υπεύθυνων μελών της κοινότητας των δεδομένων, δηλαδή ανθρώπων οι οποίοι εργάζονται για την ανάπτυξη βέλτιστων πρακτικών, λαμβάνοντας υπόψη

157. ΣτΕ: Συνήθως πρόκειται για μακροσκελείς όρους αναζήτησης, με χαμηλό όγκο αναζήτησης, που κερδίζουν μια θέση στα πρώτα αποτελέσματα της εκάστοτε μηχανής αναζήτησης.

τις προκλήσεις σχετικά με την ηθική και την προστασία της ιδιωτικής ζωής που προκύπτουν από τη χρήση δεδομένων με νέους και διαφορετικούς τρόπους, μπορούμε να εξετάσουμε τους πιθανούς κινδύνους[158].

Πρώτον, προκύπτει το ερώτημα του τρόπου με τον οποίο κατανέμεται η εξουσία σε μια τέτοια συνθήκη, όταν μια αίθουσα σύνταξης αποφασίζει να δημοσιεύσει μια βάση δεδομένων που περιέχει δεδομένα τα οποία αφορούν ανθρώπους. Συνήθως, αυτοί οι άνθρωποι δεν έχουν καμία αρμοδιότητα ή δυνατότητα να ασκήσουν βέτο ή να διορθώσουν αυτά τα δεδομένα πριν από τη δημοσίευση. Η δύναμη που διαθέτουν αυτοί οι άνθρωποι εξαρτάται σε μεγάλο βαθμό από το ποιοι είναι –για παράδειγμα, ένα πολιτικά εκτεθειμένο πρόσωπο που περιλαμβάνεται σε μια τέτοια βάση δεδομένων πιθανότατα θα ανέμενε μια τέτοια εξέλιξη και θα είχε στη διάθεσή του τους κατάλληλους πόρους για να αναλάβει δράση, ενώ ένας επαγγελματίας υγείας πιθανώς δεν θα περίμενε να εμπλακεί σε μια έρευνα. Από τη στιγμή που μια βάση δεδομένων θα δημοσιευθεί, ο τρόπος με τον οποίο προβάλλονται τα άτομα που περιλαμβάνονται σε αυτή μπορεί να αλλάξει γρήγορα –για παράδειγμα, οι γιατροί στη βάση δεδομένων *Euros für Ärzte* ανέφεραν ότι ένα από τα κορυφαία αποτελέσματα της διαδικτυακής αναζήτησης για το όνομά τους ήταν τώρα η σελίδα τους σε αυτή τη βάση δεδομένων.

Η δυναμική της εξουσίας από την πλευρά του αναγνώστη ή του θεατή αξίζει, επίσης, να εξεταστεί. Για ποιον θα μπορούσε να είναι πιο χρήσιμη η βάση δεδομένων; Διαθέτουν οι χρήστες τα εργαλεία και την ικανότητα που απαιτούνται για να μπορέσουν να αξιοποιήσουν τη βάση δεδομένων ή οι πληροφορίες αυτές θα χρησιμοποιηθούν από τους ήδη ισχυρούς για την προώθηση των συμφερόντων τους; Αυτό μπορεί να σημαίνει ότι πρέπει να υπάρξει διεύρυνση των δοκιμών από τους χρήστες πριν από τη δημοσίευση, ώστε να διασφαλιστεί ότι δίνεται αρκετό περιεχόμενο για να εξηγηθεί σωστά η βάση δεδομένων στο επιθυμητό κοινό, ή να συμπεριληφθούν ορισμένα χαρακτηριστικά που θα έκαναν τη διεπαφή της βάσης δεδομένων πιο προσιτή σε αυτήν την ομάδα.

Η παραδοχή ότι τα περισσότερα δεδομένα οδηγούν σε αποφάσεις που είναι καλύτερες για την κοινωνία έχει αμφισβητηθεί σε πολλαπλά επίπεδα τα τελευταία χρόνια. Η ακαδημαϊκός στον τομέα της εκπαίδευσης, Κλερ Φοντέν (2017), επεκτείνει το θέμα αυτό, σημειώνοντας ότι, στις Ηνωμένες Πολιτείες, τα σχολεία γίνονται όλο και πιο διαχωρισμένα, παρά την αύξηση των διαθέσιμων δεδομένων σχετικά με τις «σχολικές επιδόσεις» (ή ίσως εξαιτίας της). Σημειώνει ότι «η αιτιώδης σχέση μεταξύ της σχολικής επιλογής και του ανεξέλεγκτου διαχωρισμού δεν έχει ακόμη τεκμηριωθεί», αλλά η ίδια και άλλοι εργάζονται περισσότερο για την κατανόηση αυτής της σχέσης, εξετάζοντας την ίσως υπερβολικά απλουστευμένη άποψη ότι οι περισσότερες πληροφορίες οδηγούν σε καλύτερες αποφάσεις και αμφισβητώντας τι μπορεί να σημαίνει «καλύτερο» (Fontaine, 2017).

Δεύτερο είναι το ερώτημα της ίδιας της βάσης δεδομένων. Μια βάση δεδομένων από μόνη της περιέχει πολλές ανθρώπινες αποφάσεις: τι συγκεντρώθηκε και τι παραλείφθηκε και πώς κατηγοριοποιήθηκε, ταξινομήθηκε ή αναλύθηκε, για παράδειγμα. Κανένα δεδομένο δεν είναι αντικειμενικό, αν και ο αλφαβητισμός και η κατανόηση των περιορισμών των δεδομένων βρίσκονται σε σχετικά χαμηλά επίπεδα, πράγμα που σημαίνει ότι οι αναγνώστες θα μπορούσαν κάλλιστα να παρεξηγήσουν τα συμπεράσματα που εξάγονται.

Για παράδειγμα, η απουσία μιας οργάνωσης από μια βάση δεδομένων των πολιτικών οργανώσεων που εμπλέκονται στο οργανωμένο έγκλημα μπορεί να μη σημαίνει ότι η ίδια η οργάνωση δεν συμμετέχει στο οργανωμένο έγκλημα· σημαίνει απλώς ότι δεν υπήρχαν διαθέσιμα στοιχεία σχετικά με τη δράση της. Οι Μίχαελ Γκολεμπιέφσκι και Ντάνα Μπόιντ αναφέρονται σε αυτήν την απουσία δεδομένων ως «κενό δεδομένων» (Golebiewski & Boyd, 2018), σημειώνοντας ότι

158. https://responsibledata.io/what-is-responsible-data/

σε ορισμένες περιπτώσεις το κενό δεδομένων μπορεί να «αντανακλά παθητικά τη μεροληψία ή την προκατάληψη στην κοινωνία». Αυτός ο τύπος απουσίας δεδομένων σε έναν κατά τα άλλα κορεσμένο από δεδομένα χώρο, επίσης, αντιστοιχεί σε αυτό που η καλλιτέχνιδα και ερευνήτρια Μίμι Ονουόχα, η οποία έχει τη βάση της στο Μπρούκλιν, αναφέρει ως «χαμένο σύνολο δεδομένων» και αναδεικνύει τις κοινωνικές επιλογές που αφορούν τη συλλογή και τη συγκέντρωση δεδομένων (Onuoha, 2016).

Τρίτον είναι η κατεύθυνση της προσοχής. Οι βάσεις δεδομένων μπορούν να αλλάξουν την εστίαση του δημόσιου ενδιαφέροντος από ένα ευρύτερο συστημικό ζήτημα στις ενέργειες μεμονωμένων ατόμων και αντίστροφα. Οι οικονομικές συναλλαγές μεταξύ των φαρμακευτικών εταιρειών και των επαγγελματιών του τομέα της υγειονομικής περίθαλψης είναι, σαφώς, ένα ζήτημα δημοσίου ενδιαφέροντος –αλλά, σε ατομικό επίπεδο, οι γιατροί μπορεί να μην θεωρούν τον εαυτό τους πρόσωπο δημοσίου ενδιαφέροντος. Παραμένει, ωστόσο, γεγονός ότι, για να καταδειχθεί ένα ζήτημα ως ευρύτερο και συστημικό (ως μοτίβο και όχι ως μεμονωμένο περιστατικό), απαιτούνται δεδομένα από πολλά άτομα. Ορισμένες βάσεις δεδομένων, όπως η προαναφερθείσα περιπτωσιολογική μελέτη, *Euros für Ärzte*, μεταβάλλουν, επίσης, τα όρια του τι ή ποιος συνιστά δημόσιο ενδιαφέρον.

Ακόμα και όταν τα άτομα συμφωνούν με τη δημοσίευση των δεδομένων τους, οι δημοσιογράφοι πρέπει να αποφασίσουν για πόσο καιρό τα δεδομένα αυτά είναι δημόσιου ενδιαφέροντος αλλά και αν και πότε θα πρέπει να αποσυρθούν. Ο Γενικός Κανονισμός για την Προστασία Δεδομένων («General Data Protection Regulation», GDPR) πιθανότατα θα επηρεάσει τον τρόπο με τον οποίο οι δημοσιογράφοι θα πρέπει να διαχειρίζονται αυτού του είδους τα προσωπικά δεδομένα και τους μηχανισμούς που θα έχουν στη διάθεσή τους τα άτομα, για να ανακαλέσουν τη συγκατάθεσή τους για την καταχώριση των δεδομένων τους.

Με όλες αυτές τις προκλήσεις, η προσέγγισή μας είναι να εξετάσουμε πώς επηρεάζονται τα δικαιώματα των ανθρώπων τόσο από τη διαδικασία όσο και από το τελικό αποτέλεσμα της έρευνας ή του προϊόντος. Στο επίκεντρο βρίσκεται η κατανόηση του γεγονότος ότι οι υπεύθυνες πρακτικές δεδομένων αποτελούν συνεχείς προσεγγίσεις και όχι λίστες πραγμάτων που πρέπει να «τσεκάρεις» σε συγκεκριμένα στάδια. Υποστηρίζουμε ότι οι προσεγγίσεις που δίνουν προτεραιότητα στα δικαιώματα των ανθρώπων οι οποίοι απεικονίζονται στα δεδομένα σε όλη τη διάρκεια της έρευνας, από τη συλλογή δεδομένων έως τη δημοσίευση, αποτελούν βασικό μέρος της βελτιστοποίησης της δημοσιογραφίας (δεδομένων) ως προς την εμπιστοσύνη (Rosen, 2018).

Βέλτιστες πρακτικές

Για τους δημοσιογράφους που σκέφτονται να δημιουργήσουν μια βάση δεδομένων για να μοιραστούν την έρευνά τους με το κοινό, ακολουθούν ορισμένες βέλτιστες πρακτικές και προτάσεις. Οραματιζόμαστε ότι θα εξελιχθούν με την πάροδο του χρόνου και καλωσορίζουμε οποιεσδήποτε συστάσεις.

Πρώτον, πριν από τη δημοσίευση, αναπτύξτε μια διαδικασία για το πώς να διορθώνετε τα λάθη στη βάση δεδομένων. Οι καλές πρακτικές τεκμηρίωσης των δεδομένων μπορούν να σας βοηθήσουν να εντοπίσετε από πού πηγάζουν τα σφάλματα. Δεύτερον, δημιουργήστε ένα κανάλι ανατροφοδότησης. Ιδιαίτερα όταν φυσικά πρόσωπα αναφέρονται χωρίς να το περιμένουν σε μια έρευνα, είναι πιθανό να υπάρξει ανατροφοδότηση (ή παράπονα). Η παροχή μιας καλής εμπειρίας στον χρήστη, για να εκφράσει αυτά τα παράπονα, μπορεί να συμβάλλει στην εμπειρία. Τρίτον, είτε να διατηρείτε τη βάση δεδομένων ενημερωμένη, είτε να επισημαίνετε ρητά ότι αυτή δεν ενημερώνεται πλέον. Εντός του δημοσιογραφικού πλαισίου, η δημοσίευση μιας βάσης δεδομένων απαιτεί υψηλότερο επίπεδο συντήρησης από ό,τι η δημοσίευση ενός άρθρου. Το επίπεδο διαδραστικότητας που παρέχει μια βάση δεδομένων σημαίνει ότι υπάρχουν διαφορετικές προσδοκίες σχετικά με το πόσο ενημερωμένη είναι σε σύγκριση με ένα άρθρο.

Τέταρτον, διαθέστε αρκετούς πόρους για συντήρηση σε βάθος χρόνου. Η επικαιροποίηση των δεδομένων και του λογισμικού της βάσης δεδομένων απαιτεί σημαντικούς πόρους. Για παράδειγμα, η προσθήκη δεδομένων της επόμενης χρονιάς σε μια βάση δεδομένων απαιτεί τη συγχώνευση νεότερων δεδομένων με παλαιότερα δεδομένα και την προσθήκη μιας επιπλέον χρονικής διάστασης στη διεπαφή χρήστη. Πέμπτον, παρατηρήστε πώς χρησιμοποιούν οι αναγνώστες τη βάση δεδομένων. Οι τάσεις στις αναζητήσεις ή στη χρήση μπορεί να σας δώσουν στοιχεία για μελλοντικά θέματα και έρευνες. Τέλος, να είστε διαφανείς: σπανίως μια βάση δεδομένων θα είναι 100% «πλήρης» και κάθε βάση δεδομένων θα έχει ορισμένες επιλογές ενσωματωμένες σε αυτήν. Αντί να αποσιωπάτε αυτές τις επιλογές, κάντε τις ορατές, ώστε οι αναγνώστες να γνωρίζουν τι βλέπουν.

Παραπομπές

EFPIA (2013) «About the EFPIA Disclosure Code», *European Federation of Pharmaceutical Industries and Associations*. Διαθέσιμο σε: https://efpia.eu/media/25046/efpia_about_disclosure_code_march-2016.pdf (Προσπελάστηκε: 5 Δεκεμβρίου 2021).

Fontaine, C. (2017) «Driving school choice», *Data & Society* [Medium], 20 Απριλίου. Διαθέσιμο σε: https://points.datasociety.net/driving-school-choice-16f014d8d4df (Προσπελάστηκε: 5 Δεκεμβρίου 2021).

Golebiewski, M., & Boyd, D. (2018) «Data voids: Where missing data can be easily exploited», *Data & Society* [Medium], Μάιος. Διαθέσιμο σε: https://datasociety.net/wp-content/uploads/2018/05/Data_Society_Data_Voids_Final_3.pdf (Προσπελάστηκε: 5 Δεκεμβρίου 2021).

Holovaty, A. (2006) «A fundamental way newspaper sites need to change», *Adrian Holovaty,* 6 Σεπτεμβρίου. Διαθέσιμο σε: http://www.holovaty.com/writing/fundamental-change/ (Προσπελάστηκε: 5 Δεκεμβρίου 2021).

Onuoha, M. (2016) *On missing data sets* [GitHub], 3 Φεβρουαρίου. Διαθέσιμο σε: https://github.com/MimiOnuoha/missing-datasets (Προσπελάστηκε: 5 Δεκεμβρίου 2021).

Rosen, J. (2018) «Optimizing journalism for trust», *The Correspondent* [Medium], 14 Μαΐου. Διαθέσιμο σε: https://medium.com/de-correspondent/optimizing-journalism-for-trust-1c67e81c123 (Προσπελάστηκε: 5 Δεκεμβρίου 2021).

Οι συγγραφείς

Η Ζαρά Ραχμάν είναι αναπληρώτρια διευθύντρια του Engine Room και συνεργάτης του Εργαστηρίου Ψηφιακής Κοινωνίας των Πολιτών στο Κέντρο Φιλανθρωπίας και Κοινωνίας των Πολιτών στο Πανεπιστήμιο του Στάνφορντ.

Ο Στέφαν Βιερμάγιερ είναι ο ιδρυτής της FragDenStaat.de, της γερμανικής πύλης για την Ελευθερία της Πληροφόρησης, και εργάζεται ως δημοσιογράφος και προγραμματιστής δεδομένων.

5.7

Η Αφήγηση Συγκρούσεων για τους Υδάτινους Πόρους Μέσω Δεδομένων και Διαδραστικών Κόμικς

Nelly Luna Amancio

Περίληψη: Πώς αναπτύξαμε ένα διαδραστικό κόμικ, για να αφηγηθούμε τα ευρήματα μιας δημοσιογραφικής έρευνας σχετικά με τον πόλεμο για το νερό στο Περού ενάντια σε μια μεγάλη εταιρεία εξόρυξης.

Λέξεις-κλειδιά: συγκρούσεις για τους υδάτινους πόρους, δημοσιογραφία δεδομένων, περιβάλλον, κόμικς, διαδραστικότητα, Περού

Όλα όσα περιγράφονται στο κόμικ *La guerra por el agua* («Ο πόλεμος για το νερό») είναι αληθινά (Εικόνα 5.7.1). Οι κύριοι χαρακτήρες –ο Μάουρο Απάσα και η Μελτσόρα Τάκουρε– υπάρχουν, μαζί με τους φόβους και τις αβεβαιότητές τους. Τους βρήκαμε μια ζεστή μέρα του Σεπτεμβρίου του 2016. Ήταν μεσημέρι, χωρίς σκιά και με άπνοια. Εκείνη ξεχορτάριαζε το χώμα με τα χέρια, εκείνος έσκαβε αυλάκια στο τραχύ έδαφος. Για πάνω από 70 χρόνια καλλιεργούν τρόφιμα σε ένα μικρό χωράφι στην κοιλάδα Τάμπο, μια γεωργική περιοχή στο νότιο Περού, όπου έχουν γίνει προτάσεις για ένα έργο εξόρυξης. Η ιστορία αυτού του ζευγαριού, όπως και η ιστορία χιλιάδων αγροτών και αυτοχθόνων κοινοτήτων, μιλάει για διαμάχες μεταξύ αγροτών και ισχυρών βιομηχανιών που εργάζονται για την εξόρυξη ενός από τους πιο στρατηγικούς πόρους στον κόσμο: του νερού.

Πώς να αφηγηθεί κανείς αυτήν την αντιπαράθεση σε μια χώρα όπως το Περού, όπου υπάρχουν περισσότερες από 200 περιβαλλοντικές συγκρούσεις και ο εθνικός προϋπολογισμός εξαρτάται σε μεγάλο βαθμό από τα έσοδα του εν λόγω τομέα; Πώς να προσεγγίσεις μια ιστορία για τις εντάσεις μεταξύ των ευάλωτων αγροτών, των συμφερόντων των πολυεθνικών εταιρειών και εκείνων μιας κυβέρνησης που πρέπει να αυξήσει την είσπραξη των φόρων της; Ποια αφήγηση μπορεί να μας βοηθήσει να το καταλάβουμε; Πώς είναι δυνατόν να κινητοποιηθεί ο κόσμος γύρω από αυτό το επιτακτικό ζήτημα; Αυτά τα ερωτήματα αποτέλεσαν το έναυσμα για το έργο *La guerra por el agua* –το πρώτο διαδραστικό κόμικ στο Περού, που δημιουργήθηκε από το *OjoPúblico*. Το κομμάτι ενσωματώνει δεδομένα και απεικονίσεις σε μια αφήγηση για αυτή τη σύγκρουση[159].

Γιατί ένα διαδραστικό κόμικ;

Το έργο ξεκίνησε τον Ιούλιο του 2016. Ξεκινήσαμε σκοπεύοντας να αφηγηθούμε τη σύγκρουση από οικονομική άποψη, αλλά προσεγγίζοντας τον αναγνώστη από την οπτική γωνία δύο αγροτών, μέσα από μια διαδρομή που μιμείται ένα οικείο ταξίδι σε μία από τις πιο εμβληματικές περιοχές της σύγκρουσης. Το διαδραστικό φορμάτ επιτρέπει στο κοινό να ανακαλύψει τους ήχους και τους διαλόγους της σύγκρουσης, πέρα και πάνω από τα σκίτσα.

159. https://laguerraporelagua.ojo-publico.com/en/

Εικόνα 5.7.1: Η αρχική οθόνη του διαδραστικού κόμικ *La guerra por el agua*. Πηγή: *OjoPúblico*.

Επιλέξαμε την ιστορία του μεταλλευτικού έργου Tía María της Southern Copper Corporation –μίας από τις μεγαλύτερες εταιρείες εξόρυξης στον κόσμο, η οποία ανήκει σε έναν από τους πλουσιότερους ανθρώπους στο Μεξικό και παγκοσμίως, τον Χερμάν Λαρέα. Η αντίθεση της τοπικής κοινωνίας σε αυτό το έργο οδήγησε σε βίαιη αστυνομική καταστολή, κατά την οποία σκοτώθηκαν έξι πολίτες.

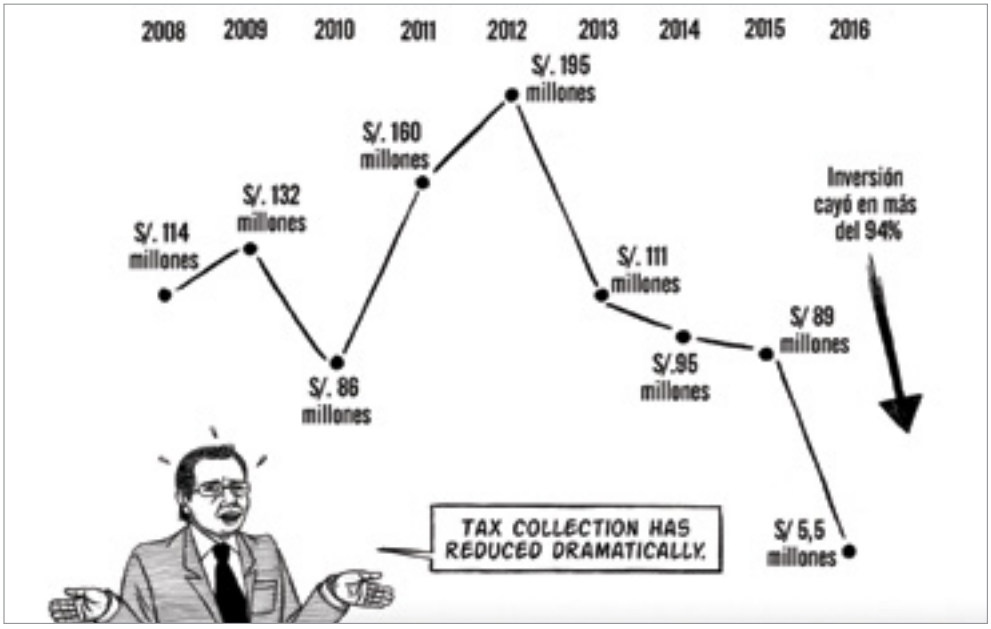
Την ομάδα που δημιούργησε αυτό το κόμικ αποτελούσαν μια δημοσιογράφος (εγώ), ο σκιτσογράφος Χεσούς Κοσίο και ο προγραμματιστής για την ανάπτυξη της ιστοσελίδας, Χασόν Μαρτίνεθ. Οι τρεις μας ταξιδέψαμε στην κοιλάδα Τάμπο στην Αρεκίπα, στην καρδιά της σύγκρουσης, για να πάρουμε συνεντεύξεις από ηγέτες, αγρότες και Αρχές και να καταγράψουμε τη διαδικασία. Κρατήσαμε σημειώσεις, τραβήξαμε φωτογραφίες και κάναμε σχέδια που αργότερα θα γίνονταν τα πρώτα σκίτσα του κόμικ. Επιστρέφοντας στη Λίμα, διαμορφώσαμε μια πρώτη εκδοχή του κόμικ, πάνω στην οποία γράψαμε το τελικό σενάριο, επεξεργαστήκαμε τις διαδραστικές δυνατότητες και αρχίσαμε να αναπτύσσουμε το έργο.

Ειλικρίνεια με κόμικς

Επιλέξαμε ως μέσο το κόμικ, επειδή πιστεύουμε ότι οι δημοσιογράφοι δεν πρέπει –όπως το θέτει ο σκιτσογράφος Τζο Σάκο (2012)– «να ευνουχίζουν την αλήθεια στο όνομα της ισοκατανομής του χρόνου». Ο Σάκο συμμετείχε στην παρουσίαση του πρώτου κεφαλαίου του έργου και ήταν μία από τις δικές του δουλειές που μας ενέπνευσε: το *Srebrenica*, ένα «webcomic» για τη σφαγή κατά την οποία έχασαν τη ζωή τους περισσότεροι από 8.000 Βόσνιοι μουσουλμάνοι το 1995.

Το έργο *La guerra por el agua* χρειάστηκε οκτώ μήνες για να ολοκληρωθεί. Βασίζεται σε αληθινά γεγονότα και έχει μια αφηγηματική δομή που επιτρέπει στο κοινό να βιώσει την καθημερινή ζωή των χαρακτήρων και να φέρει στην επιφάνεια ένα από τα μεγαλύτερα διλήμματα στην οικονομία του Περού: γεωργία ή εξόρυξη; Υπάρχει αρκετό νερό για να γίνουν και τα δύο; Αφηγηθήκαμε την ιστορία αυτής της σύγκρουσης μέσα από τα μάτια και τις αναμνήσεις του Μάουρο και της Μελτσόρα. Η ιστορία συνοδεύεται από οπτικοποιήσεις δεδομένων που δείχνουν την οικονομική εξάρτηση της περιοχής, καθώς και τα φορολογικά προνόμια που έχουν οι εταιρείες εξόρυξης. Όλες οι σκηνές και οι διάλογοι στο κόμικ είναι πραγματικές, προϊόντα του ρεπορτάζ

μας στην περιοχή, των συνεντεύξεων με τις Αρχές και τους ντόπιους αλλά και των ερευνών σχετικά με τα οικονομικά της Southern Copper. Στόχος μας ήταν να συνθέσουμε σκηνές από διαλόγους, φιγούρες, συνεντεύξεις και τοποθεσίες με ειλικρίνεια και ακρίβεια.



Εικόνα 5.7.2: Οπτικοποίηση δεδομένων που δείχνει τη μείωση των φορολογικών εσόδων από το 2008 στο Περού, ως αποτέλεσμα της σύγκρουσης για το νερό. Πηγή: *OjoPúblico*.

Εικόνα 5.7.3: Ο τρόπος με τον οποίο οι δημοσιογράφοι και ο εικονογράφος του *OjoPúblico* ανέπτυξαν το διαδραστικό σενάριο του κόμικ *La guerra por el agua*. Πηγή: *OjoPúblico*.

Από το χαρτί στον Ιστό

Για τον σκιτσογράφο Χεσούς Κοσίο, η πρόκληση ήταν να επανεξετάσει τη διαχείριση του χρόνου σε ένα διαδραστικό κόμικ: «Ενώ σε ένα έντυπο καρτούν ή σε ένα στατικό ψηφιακό στριπ η ιδέα είναι να κάνεις τον αναγνώστη να σταθεί στον αντίκτυπο των εικόνων, σε ένα διαδραστικό κόμικ η σύνθεση και οι εικόνες έπρεπε να προσαρμοστούν στην πιο ευέλικτη και δυναμική ροή της ανάγνωσης».

Από τεχνολογική άποψη, το έργο ήταν μια πρόκληση για την ομάδα του *OjoPúblico*, καθώς δεν είχαμε αναπτύξει ποτέ άλλοτε ένα διαδραστικό κόμικ. Χρησιμοποιήσαμε την GreenSock Animation Platform (GSAP), μια βιβλιοθήκη που μας επέτρεψε να φτιάξουμε κινούμενα σχέδια και μεταβάσεις, καθώς και να τυποποιήσουμε τις σκηνές και το χρονοδιάγραμμα. Συμπληρωματικά χρησιμοποιήθηκαν οι γλώσσες JavaScript, CSS και HTML5.

Το κόμικ έχει 42 σκηνές και περισσότερα από 120 σχέδια. Ο Χεσούς Κοσίο σχεδίασε καθέναν από τους χαρακτήρες, τις σκηνές και τα τοπία του σεναρίου με μολύβι και μελάνι. Στη συνέχεια, οι εικόνες αυτές ψηφιοποιήθηκαν και διαχωρίστηκαν ανά επίπεδα: το φόντο, το περιβάλλον, οι χαρακτήρες και τα στοιχεία του σχεδίου που έπρεπε να αλληλεπιδρούν μεταξύ τους.

Επιστροφή από τον Ιστό στο χαρτί

Το έργο *La guerra por el agua* είναι μια διαμεσική εμπειρία. Έχουμε, επίσης, δημοσιεύσει και μια έντυπη έκδοση. Μέσα από τα δύο φορμάτ του, το κόμικ επιδιώκει να προσεγγίσει διαφορετικό κοινό. Ένα από τα μεγαλύτερα ενδιαφέροντα της ομάδας του *OjoPúblico* είναι η εξερεύνηση αφηγήσεων και φορμάτ για την αφήγηση (συχνά πολύπλοκων) ιστοριών δημόσιου ενδιαφέροντος. Έχουμε κερδίσει στο παρελθόν βραβεία για τις έρευνες δεδομένων μας.

Σε άλλα έργα έχουμε, επίσης, χρησιμοποιήσει το φορμάτ του κόμικ, για να αφηγηθούμε το θέμα της βίας. Στο *Proyecto Memoria* («Έργο Μνήμης»), οι εικόνες αφηγούνται τη φρίκη των εσωτερικών συγκρούσεων που αντιμετώπισε το Περού μεταξύ 1980 και 2000. Τα κόμικς παρέχουν μια ισχυρή γλώσσα για την αφήγηση ιστοριών με δεδομένα. Αυτή είναι η πρότασή μας: οι ερευνητές δημοσιογράφοι θα πρέπει να δοκιμάζουν όλες τις πιθανές γλώσσες, για να αφηγούνται ιστορίες για διαφορετικά ακροατήρια. Κυρίως θέλουμε, όμως, να καταγγείλουμε τις ανισορροπίες εξουσίας –στην προκείμενη περίπτωση, τη διαχείριση των φυσικών πόρων στο Περού.

Παραπομπές

Sacco, J. (2012). *Journalism*. Νέα Υόρκη: Henry Holt and Co.

Η συγγραφέας

Η Νέλι Λούνα Αμάνθιο είναι ερευνήτρια δημοσιογράφος, εκδότρια και ιδρύτρια του *OjoPúblico*, ενός περουβιανού μέσου ενημέρωσης που ερευνά την εξουσία και διεξάγει διασυνοριακές έρευνες στη Λατινική Αμερική.

5.8
Η Δημοσιογραφία Δεδομένων Πρέπει να Εστιάζει στους Ανθρώπους και στις Ιστορίες

Winny de Jong

Περίληψη: Η ιστορία και οι άνθρωποι στους οποίους αναφέρεται αυτή η ιστορία πρέπει να είναι ο ήλιος γύρω από τον οποίο περιστρέφεται η δημοσιογραφία, συμπεριλαμβανομένης της δημοσιογραφίας δεδομένων.

Λέξεις-κλειδιά: αφήγηση, δημοσιογραφία δεδομένων, ραδιόφωνο, τηλεόραση, δημόσια δεδομένα, οπτικοποίηση δεδομένων

Όπως συμβαίνει με τους ανθρώπους, η δημοσιογραφία δεδομένων και η δημοσιογραφία έχουν περισσότερα κοινά παρά διαφορές[160]. Παρόλο που το ρεπορτάζ με γνώμονα τα δεδομένα βασίζεται σε διαφορετικούς τύπους πηγών που απαιτούν άλλες δεξιότητες για την εξέτασή τους, ο τρόπος σκέψης είναι σε μεγάλο βαθμό ο ίδιος. Στην πραγματικότητα, εάν το εξετάσετε αρκετά ψύχραιμα, θα διαπιστώσετε ότι οι διαδικασίες είναι σχεδόν απαράλλακτες.

Εγνωσμένοι άγνωστοι

Στον πυρήνα της, η δημοσιογραφία είναι η δουλειά της μετατροπής των εγνωσμένων αγνώστων παραγόντων σε εγνωσμένους γνωστούς. Οι έννοιες των γνωστών και αγνώστων υιοθετήθηκαν από τον υπουργό Άμυνας των ΗΠΑ, Ντόναλντ Ράμσφελντ, το 2002. Εκείνη την εποχή δεν υπήρχαν αποδείξεις ότι η ιρακινή κυβέρνηση είχε προμηθεύσει όπλα μαζικής καταστροφής σε τρομοκρατικές ομάδες. Κατά τη διάρκεια ενημέρωσης Τύπου για το θέμα, ο Ράμσφελντ είπε:

> Με ενδιαφέρουν πάντα οι αναφορές που λένε ότι κάτι δεν έχει συμβεί, διότι όπως ξέρουμε, υπάρχουν εγνωσμένοι γνωστοί παράγοντες –υπάρχουν πράγματα που ξέρουμε ότι τα ξέρουμε. Γνωρίζουμε, επίσης, ότι υπάρχουν εγνωσμένοι άγνωστοι παράγοντες –δηλαδή, γνωρίζουμε ότι υπάρχουν κάποια πράγματα που δεν τα γνωρίζουμε. Υπάρχουν, όμως, και οι μη εγνωσμένοι άγνωστοι παράγοντες –αυτά που δεν ξέρουμε ότι δεν τα ξέρουμε. Και αν μελετήσει κανείς την ιστορία της χώρας μας και άλλων ελεύθερων χωρών, η τελευταία κατηγορία είναι αυτή που τείνει να είναι η πιο δύσκολη (US Department of Defense, 2002).

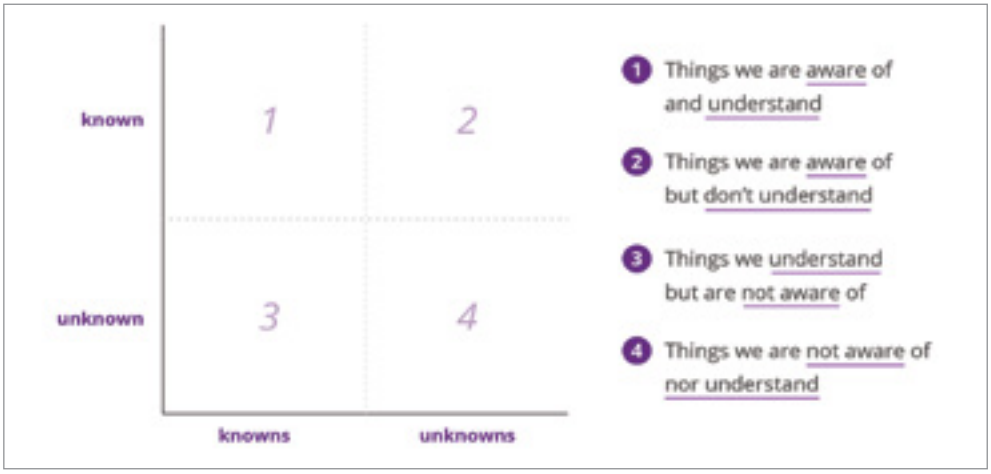
Κάθε δημοσιογραφική διαδικασία καταλήγει να μετακινεί τα πιόνια πάνω στο πλέγμα των γνωστών και των αγνώστων. Όλη η δημοσιογραφία ξεκινά με μια ερώτηση –ή, σύμφωνα με το εν λόγω πλέγμα, με έναν εγνωσμένο άγνωστο. (Ξέρετε ότι υπάρχει κάτι που δεν γνωρίζετε, εξού και η ερώτηση.) Όταν ξεκινάμε από το ερώτημα ή το προαίσθημα, για να καταλήξουμε σε μια έτοιμη για

160. Δεδομένου ότι οι ιδέες είναι νέοι συνδυασμοί παλαιών στοιχείων, αυτό το δοκίμιο βασίζεται στην πρόβλεψη της Γουίνι ντε Γιόνγκ για το Nieman Lab το 2019, σε μια ομιλία στο συνέδριο Smart News Design στο Άμστερνταμ και στο alshetongeveermaarklopt.nl, έναν ολλανδικό ιστότοπο που διδάσκει μαθηματικά σε δημοσιογράφους.

δημοσίευση ιστορία, η ιδανική διαδρομή είναι να μετακινήσουμε «απλά» όλα τα πιόνια από τους εγνωσμένους αγνώστους στους εγνωσμένους γνωστούς. Αλλά, όπως θα σας πει κάθε δημοσιογράφος, η πραγματικότητα τείνει να διαφέρει. Κατά την έρευνα –είτε παίρνοντας συνεντεύξεις από ανθρώπους είτε εξετάζοντας έγγραφα ή σύνολα δεδομένων– είναι πιθανό να βρείτε πράγματα που δεν γνωρίζατε ότι δεν γνωρίζατε (μη εγνωσμένοι άγνωστοι), τα οποία ζητούν, επίσης, απαντήσεις. Εάν είστε τυχεροί, μπορεί να πέσετε πάνω σε κάποια πράγματα που δεν γνωρίζατε ότι γνωρίζατε (άγνωστοι γνωστοί). Όσο πλησιάζετε την προθεσμία σας για την παράδοση του δημοσιογραφικού έργου, μετατρέπετε τρεις κατηγορίες γνώσεων σε εγνωσμένους γνωστούς παράγοντες: τους εγνωσμένους αγνώστους (δηλαδή, τα ερωτήματα που σας έκαναν να ξεκινήσετε), τους μη εγνωσμένους αγνώστους (δηλαδή, τα ερωτήματα που δεν ξέρατε ότι έπρεπε να είχατε θέσει) και τους άγνωστους γνωστούς (απαντήσεις που δεν ξέρατε ότι είχατε). Σε αντίθεση με τις κυβερνήσεις μας, οι δημοσιογράφοι μπορούν να αναλάβουν δράση ή να δημοσιεύσουν μόνο ό,τι είναι εγνωσμένα γνωστό.

Στέρεη δημοσιογραφία

Με το ρεπορτάζ που βασίζεται σε δεδομένα και το κλασικό δημοσιογραφικό εγχείρημα που βασίζεται σε ίδιες δυνάμεις να είναι τόσο δυσδιάκριτα, σίγουρα τα δύο θα πρέπει να πληρούν τα ίδια πρότυπα. Όπως η δημοσιογραφία, έτσι και η δημοσιογραφία δεδομένων πρέπει πάντα να είναι ειλικρινής, ανεξάρτητη και απαλλαγμένη από προκαταλήψεις. Όπως όλα τα άλλα γεγονότα, τα δεδομένα πρέπει να επιβεβαιώνονται. Έτσι, πριν να προσπαθήσετε να δημιουργήσετε εγνωσμένους γνωστούς, ρωτήστε τον εαυτό σας: είναι τα δεδομένα αληθινά; Τι σημαίνει στην πραγματικότητα ο κάθε αριθμός; Ποια είναι η πηγή; Γιατί συλλέχθηκαν τα δεδομένα; Ποιος έφτιαξε το σύνολο δεδομένων; Πώς δημιουργήθηκε ο πίνακας δεδομένων; Υπάρχουν ακραίες τιμές στα δεδομένα; Βγάζουν νόημα; Και, κάτι που συχνά ξεχνιέται, αλλά, όπως σε κάθε συνέντευξη, έχει ιδιαίτερη σημασία: τι δεν λέει η πηγή; Ενώ οι απαιτήσεις και, συνεπώς, οι ερωτήσεις είναι οι ίδιες, οι ενέργειες στις οποίες καταλήγουν διαφέρουν ελαφρώς.

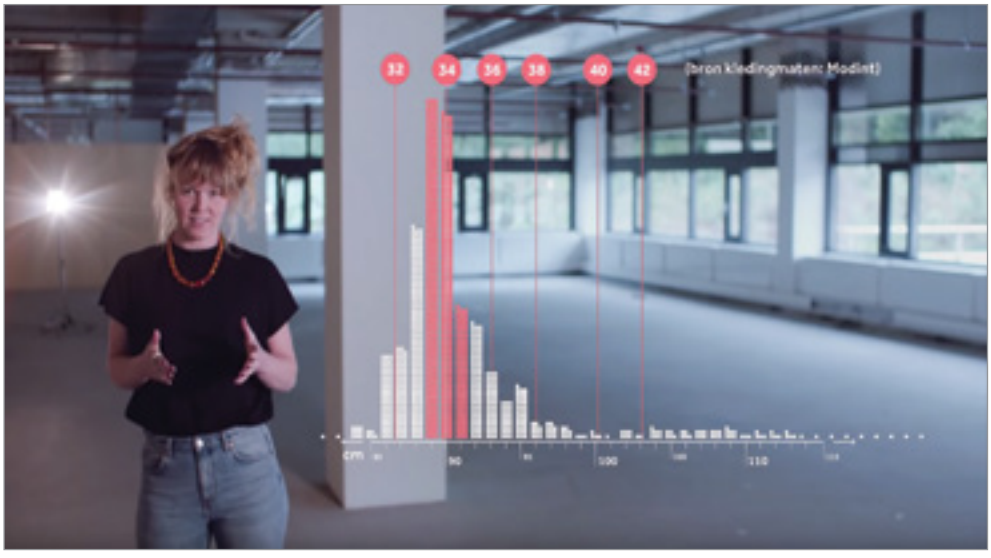


Εικόνα 5.8.1: Το πλέγμα των γνωστών και των αγνώστων. Το πλέγμα διαφέρει ελαφρώς από το παράθυρο Johari[161], το οποίο μερικές φορές χρησιμοποιείται στη γνωστική ψυχολογία, για να βοηθήσει τους ανθρώπους να κατανοήσουν καλύτερα τη σχέση τους με τους άλλους. Πηγή: Lars Boogaard.

161. ΣτΕ: Πρόκειται για εργαλείο αυτεπίγνωσης, που χρησιμοποιείται για την κατανόηση των ανθρώπινων ικανοτήτων στη διαπροσωπική επικοινωνία. Η ονομασία του, «παράθυρο Johari», προέρχεται από τα ονόματα των επινοητών του, Τζόζεφ Λαφ (Joseph Luff) και Χάρι Ίνγκαμ (Harry Ingham).

Εικόνα 5.8.2: Πλοηγούμενοι στο πλέγμα των γνωστών και των αγνώστων στη δημοσιογραφία. Πηγή: Lars Boogaard.

Όπως περιγράφουν οι Μπιλ Κόβατς και Τομ Ρόζενστιλ (2007) στο βιβλίο *The Elements of Journalism*, το πρώτο καθήκον του δημοσιογράφου είναι «να εξακριβώσει ποιες πληροφορίες είναι αξιόπιστες και, στη συνέχεια, να τις ταξινομήσει, έτσι ώστε οι άνθρωποι να μπορούν να τις κατανοήσουν με αποτελεσματικό τρόπο». Για τους δημοσιογράφους δεδομένων –ιδιαίτερα εκείνους που εργάζονται στην τηλεόραση ή στο ραδιόφωνο–, αυτό σημαίνει ότι οι αριθμοί που έφτασαν να αγαπούν δεν έχουν απαραίτητα θέση στο τελικό προϊόν.



Εικόνα 5.8.3: Καρέ από βίντεο του NOS σχετικά με το πόσο αδύνατος χρειάζεται να είσαι, για να γίνεις μοντέλο. Πηγή: NOS.

Σπασίκλες με μέτρο

Προφανώς θα πρέπει να είστε ακριβείς κατά την ανάλυση δεδομένων. Προκειμένου η ιστορία σας να παραμείνει, όμως, «αποτελεσματικά κατανοητή», πρέπει να υπάρχει ένα όριο στην ακρίβεια –για παράδειγμα, στον αριθμό των δεκαδικών ψηφίων που χρησιμοποιούνται στην τελική δημοσίευση. Το «4 στα 10 άτομα» είναι μάλλον καλύτερο από το «41,8612%». Με βάση την εμπειρία μου, η σωστή δόση ακρίβειας είναι πολύ κοντά στην ακρίβεια που θα χρησιμοποιούσατε αν μιλούσατε για την ιστορία σας σε φίλους που δεν είναι κολλημένοι με τα δεδομένα, ένα απόγευμα Σαββάτου.

Εκτός εάν το κοινό σας πρέπει να γνωρίζει τις μεθόδους και τα εργαλεία που χρησιμοποιήθηκαν για να μπορέσει να κατανοήσει την ιστορία, θα πρέπει μάλλον να αφήσετε τα καλούδια του σπασίκλα για τη δημοσίευση όπου θα εξηγείτε τη μεθοδολογία σας. Διότι όταν το κοινό σας διαβάζει, ακούει ή παρακολουθεί το προϊόν που βασίζεται στα δεδομένα, θα πρέπει να σκέφτεται την ιστορία –και όχι τα δεδομένα, την ανάλυση ή την τεχνολογία που διατηρούν την ιστορία ζωντανή. Αυτό σημαίνει ότι η καλύτερη δημοσιογραφία δεδομένων μπορεί δύσκολα να διακρίνεται ως τέτοια –κάτι που καθιστά τη δημοσιογραφία δεδομένων μια αόρατη τέχνη. Εφόσον αυτή η αορατότητα διευκολύνει την ιστορία, κάνοντας τη δημοσιογραφία σας πιο «αποτελεσματική στην κατανόηση», είναι για καλό. Άλλωστε, η δημοσιογραφία δημιουργεί διαφορετικούς χάρτες με τους οποίους οι πολίτες μπορούν να πλεύσουν στην κοινωνία. Επομένως, θα πρέπει να διασφαλίσουμε ότι οι χάρτες μας είναι ευανάγνωστοι για όλους και ότι διαβάζονται από πολλούς.

Ραδιόφωνο και τηλεόραση

Όταν δημοσιεύετε ιστορίες δημοσιογραφίας δεδομένων για το ραδιόφωνο ή την τηλεόραση, το λίγο αρκεί –«less is more», που λέμε στα αγγλικά. Στην αίθουσα σύνταξης του NOS, του μεγαλύτερου ειδησεογραφικού οργανισμού της Ολλανδίας, οι δημοσιογράφοι μιλούν για τον αριθμό των δευτερολέπτων που έχουν στη διάθεσή τους, για να αφηγηθούν τις ιστορίες τους. Αυτό σημαίνει ότι δεν υπάρχει χρόνος για να ασχοληθούμε με το πώς δημιουργήθηκε μια ιστορία ή για ποιον λόγο αποφασίσαμε να χρησιμοποιήσουμε τη μία πηγή δεδομένων και όχι την άλλη, αν αυτό δεν συμβάλλει στην ιστορία ή στην κατανόηση της εν λόγω ιστορίας από το κοινό. Σε ένα διαδικτυακό βίντεο σχετικά με το πόσο αδύνατος πρέπει να είναι κανείς, για να μπορέσει να γίνει μοντέλο της υψηλής μόδας, αφιερώσαμε 20 δευτερόλεπτα, για να εξηγήσουμε τις μεθόδους μας[162]. Όταν έχεις στη διάθεσή σου 90 δευτερόλεπτα για να πεις μια ιστορία σε εθνικό τηλεοπτικό δίκτυο, τα 20 δευτερόλεπτα είναι πολλά. Σε αυτήν την περίπτωση, το «less is more» σημαίνει ότι δεν μένει χρόνος για να εξηγήσουμε πώς διεξαγάγαμε την έρευνα. Όταν ο χρόνος και ο χώρος είναι περιορισμένοι, η ιστορία υπερισχύει όλων των άλλων.

Μετριοπαθή γραφικά

Φυσικά, το ρητό «less is more» ισχύει και για τις οπτικοποιήσεις δεδομένων. Η δημοσιογραφία δεδομένων μοιάζει πολύ με το εφηβικό σεξ: όλοι μιλάνε γι' αυτό, αλλά σχεδόν κανείς δεν το κάνει στην πραγματικότητα. Όταν οι αίθουσες σύνταξης προσθέτουν, τελικά, τα δεδομένα στην εργαλειοθήκη τους, ορισμένες έχουν την τάση να φλυαρούν φτιάχνοντας οπτικές απεικονίσεις δεδομένων για τα πάντα. Σίγουρα, κι εμένα μου αρέσουν τα γραφικά, ειδικά τα καινοτόμα, υψηλής ποιότητας γραφικά, αλλά μόνο αν προσθέτουν στην ιστορία. Οι οπτικοποιήσεις μπορούν να προσθέσουν αξία στη δημοσιογραφία με πολλούς τρόπους. Μεταξύ άλλων, μπορούν να το κάνουν εμβαθύνοντας στην κατανόηση της συγκεκριμένης ιστορίας από το κοινό και διευρύνο-

162. https://www.youtube.com/watch?v=DWRGqmywNY

ντας την κατανόηση του κοινού με το να παρέχουν επιπλέον πληροφορίες, για παράδειγμα, σε περιφερειακό επίπεδο. Φερθείτε, λοιπόν, ευγενικά και μη φλυαρείτε. Περιοριστείτε σε οπτικοποιήσεις δεδομένων που προσθέτουν αξία και βοηθούν στην κατανόηση της ιστορίας. Σήμερα, οι περισσότεροι άνθρωποι ακούν ραδιόφωνο ή παρακολουθούν τηλεόραση συνδυαστικά με κάποια άλλη δραστηριότητα. Αυτό περιορίζει τη δυνατότητα πρόσληψης πληροφοριών: το ίδιο ισχύει και για την παρακολούθηση τηλεόρασης ενώ μαγειρεύετε. Προσοχή, λοιπόν, να μη ζητάτε πολλά από το κοινό σας. Και πάλι, αυτό μπορεί να καθιστά τη δουλειά μας αόρατη, αλλά είμαστε εδώ για να μεταδίδουμε ειδήσεις και να λέμε ιστορίες –όχι για να κάνουμε εξάσκηση στην οπτικοποίηση δεδομένων.

Σχετικά με τους ανθρώπους

Όλα αυτά σημαίνουν πως ό,τι πραγματικά έχει σημασία –στην ιστορία σας, στη δημοσιογραφία και στη ζωή γενικότερα– δεν χωράει σε ένα σύνολο δεδομένων. Ποτέ δεν χωρούσε και ποτέ δεν θα χωρέσει. Στο τέλος, το θέμα είναι πάντα οι άνθρωποι. Επομένως, ό,τι και αν κάνετε, όπου και αν δημοσιεύετε, να μιλάτε για ανθρώπους και όχι για δεδομένα. Και όταν μπαίνετε στον πειρασμό να χρησιμοποιήσετε περισσότερα δεδομένα, τεχνολογία ή ειδησεογραφική εξειδίκευση από ό,τι χρειάζεται, να θυμάστε ότι είστε ένας από τους λίγους τεχνίτες σε αυτόν τον τομέα, κάτι από μόνο του υπέροχο: δεν συντρέχει λόγος να υπογραμμιστεί το προφανές. Επομένως, απλά τηρήστε την ιεραρχική σειρά που απαντάται στα καλύτερα δείγματα της δημοσιογραφίας δεδομένων: η φόρμα εξυπηρετεί τα δεδομένα, εξυπηρετεί την ιστορία. Όλα και όλοι πρέπει να περιστρέφονται γύρω από την ιστορία –αυτή είναι ο ήλιος μας. Η ιστορία είναι ο βασιλιάς.

Παραπομπές

Kovach, B., & Rosenstiel, T. (2007). T*he elements of journalism: What newspeople should know and the public should expect.* Νέα Υόρκη: Three Rivers Press.

US Department of Defense (2002) DoD *News Briefing—Secretary Rumsfeld and Gen.* Myers [Απομαγνητοφώνηση συνέντευξης], 12 Φεβρουαρίου.

Η συγγραφέας

Η Γουίνι ντε Γιόνγκ εργάζεται ως δημοσιογράφος δεδομένων στον ολλανδικό εθνικό ραδιοτηλεοπτικό οργανισμό NOS και δημοσιεύει το εβδομαδιαίο *Data Journalism Newsletter* στο ddj.news.

6.

Διερευνώντας τα Δεδομένα, τις Πλατφόρμες και τους Αλγορίθμους

6.1
Το Ρεπορτάζ των Αλγορίθμων: Οπτικές και Μέθοδοι Έρευνας

Nicholas Diakopoulos

Περίληψη: Μια ειδίκευση ρεπορτάζ αρχίζει και δομείται γύρω από τους αλγορίθμους, καθώς οι δημοσιογραφικές δεξιότητες συνδυάζονται με τις τεχνικές για να προσφέρουν τον κατάλληλο έλεγχο για τους αλγορίθμους.

Λέξεις-κλειδιά: αλγόριθμοι, αλγοριθμική λογοδοσία, υπολογιστική δημοσιογραφία, ερευνητική δημοσιογραφία, σπουδές αλγορίθμων, ελευθερία πληροφόρησης

Η σειρά *Machine Bias* («Μεροληψία των Μηχανών») της *ProPublica* ξεκίνησε τον Μάιο του 2016 ως μια προσπάθεια διερεύνησης των αλγορίθμων στην κοινωνία[163]. Ίσως το πιο εντυπωσιακό κομμάτι της σειράς ήταν μια έρευνα και ανάλυση που εξέθεσε τη φυλετική προκατάληψη που επιδεικνύουν οι αλγόριθμοι αξιολόγησης του κινδύνου υποτροπής εγκληματιών που χρησιμοποιούνται στις αποφάσεις της ποινικής δικαιοσύνης (Angwin κ.α., 2016). Αυτοί οι αλγόριθμοι βαθμολογούν τα άτομα με βάση τον πιθανό κίνδυνο να επαναλάβουν την όποια εγκληματική τους δραστηριότητα −εάν είναι υψηλός ή χαμηλός. Πολιτείες και άλλοι δήμοι χρησιμοποιούν ποικιλοτρόπως τις βαθμολογίες για τη διαχείριση της προφυλάκισης, της αναστολής, της αποφυλάκισης υπό όρους και μερικές φορές ακόμη και της καταδίκης. Οι δημοσιογράφοι της *ProPublica* υπέβαλαν αίτημα για τα δημόσια αρχεία σχετικά με τις βαθμολογίες από την κομητεία Μπρόγουαρντ στη Φλόριντα και, στη συνέχεια, σύγκριναν τις βαθμολογίες αυτές με το πραγματικό ποινικό ιστορικό, για να δουν εάν ένα άτομο είχε πράγματι επαναλάβει την εγκληματική του δράση (δηλαδή, αν είχε υποπέσει ξανά σε αξιόποινη πράξη) εντός δύο ετών. Η ανάλυση των δεδομένων έδειξε ότι οι μαύροι κατηγορούμενοι έτειναν να λαμβάνουν υψηλότερες βαθμολογίες κινδύνου από ό,τι οι λευκοί κατηγορούμενοι και ήταν πιο πιθανό να χαρακτηρίζονται εσφαλμένα ως κατηγορούμενοι υψηλού κινδύνου, ενώ στην πραγματικότητα μετά από δύο χρόνια δεν είχαν συλληφθεί εκ νέου (Larson κ.α., 2016).

Η βαθμολόγηση στο πλαίσιο του συστήματος ποινικής δικαιοσύνης είναι, φυσικά, μόνο ένας τομέας όπου οι αλγόριθμοι χρησιμοποιούνται στην κοινωνία. Η σειρά *Machine Bias* έχει έκτοτε πληθώρα θεμάτων, από το σύστημα στόχευσης διαφημίσεων του Facebook μέχρι τις γεωγραφικές διακρίσεις στις τιμές ασφάλισης αυτοκινήτων και τις αθέμιτες πρακτικές τιμολόγησης στο Amazon.com. Η λήψη αλγοριθμικών αποφάσεων εξαπλώνεται ολοένα και περισσότερο τόσο στον δημόσιο όσο και στον ιδιωτικό χώρο. Το βλέπουμε σε τομείς όπως η βαθμολόγηση πιστωτικού και ασφαλιστικού κινδύνου, τα συστήματα απασχόλησης, η διαχείριση της κοινωνικής πρόνοιας, η κατάταξη εκπαιδευτικών και καθηγητών και η επιμέλεια διαδικτυακών μέσων, μεταξύ πολλών άλλων (Eubanks, 2018· O'Neil, 2016· Pasquale, 2015). Λειτουργώντας σε κλίμακα και επηρεάζοντας συχνά μεγάλες ομάδες ανθρώπων, οι αλγόριθμοι μπορούν να παίρνουν αποφάσεις που έχουν συνέπειες, μερικές φορές αμφισβητήσιμες, υπολογισμού, κατάταξης, ταξινόμησης, συσχέτισης και διαλογής. Οι αλγόριθμοι, τροφοδοτούμενοι από στοίβες δεδομένων, αποτελούν έναν ισχυρό νέο τρόπο άσκησης της εξουσίας στην κοινωνία.

163. https://www.propublica.org/series/machine-bias

Όπως μαρτυρά η σειρά *Machine Bias* της *ProPublica*, ένα νέο σκέλος της υπολογιστικής δημοσιογραφίας («computational jounralism») και της δημοσιογραφίας δεδομένων αναδύεται, για να διερευνήσει και να αποδώσει ευθύνες για τον τρόπο με τον οποίο ασκείται η εξουσία μέσω αλγορίθμων. Αυτό το ονομάζω ρεπορτάζ αλγοριθμικής λογοδοσίας και είναι ένας επαναπροσανατολισμός της παραδοσιακής δημοσιογραφικής λειτουργίας στον έλεγχο της εξουσίας που ασκείται μέσω των αλγορίθμων (Diakopoulos, 2015)[164]. Παρά τη φαινομενική αντικειμενικότητά τους, οι αλγόριθμοι μπορούν να κάνουν λάθη και να ενσωματώνουν προκαταλήψεις που απαιτούν προσεκτικότερη διερεύνηση. Σιγά σιγά μια ειδίκευση ρεπορτάζ αρχίζει και δομείται γύρω από τους αλγορίθμους, καθώς οι δημοσιογραφικές δεξιότητες συνδυάζονται με τις τεχνικές δεξιότητες, για να προσφέρουν τον απαιτούμενο έλεγχο των αλγορίθμων.

Υπάρχουν, βέβαια, ποικίλες μορφές αλγοριθμικής λογοδοσίας που μπορεί να λαμβάνουν χώρα σε διάφορα φόρουμ πέραν της δημοσιογραφίας, όπως σε πολιτικά, νομικά, ακαδημαϊκά, ακτιβιστικά ή καλλιτεχνικά πλαίσια (Brain & Mattu, άνευ ημερομηνίας· Bucher, 2018)[165]. Ωστόσο, στο παρόν κεφάλαιο επικεντρώνομαι αποκλειστικά στο ρεπορτάζ αλγοριθμικής λογοδοσίας ως ανεξάρτητη δημοσιογραφική προσπάθεια που συμβάλλει στη λογοδοσία μέσω της άσκησης δημόσιας πίεσης. Αυτό μπορεί να θεωρηθεί συμπληρωματικό άλλων τρόπων οι οποίοι, τελικά, μπορούν να συμβάλλουν επίσης στη λογοδοσία, όπως η ανάπτυξη κανονισμών και νομικών προτύπων, η δημιουργία ελεγκτικών θεσμών στην κοινωνία των πολιτών, η χάραξη αποτελεσματικών πολιτικών διαφάνειας, η έκθεση αναστοχαστικών καλλιτεχνικών επιδείξεων και η δημοσίευση ακαδημαϊκών κριτικών.

Προκειμένου να αποφασιστεί τι συνιστά ρεπορτάζ ειδίκευσης στη δημοσιογραφία, είναι χρήσιμο να οριστεί πρώτα τι είναι αυτό που αποτελεί είδηση σχετικά με τους αλγορίθμους. Τεχνικά μιλώντας, ένας αλγόριθμος είναι μια ακολουθία βημάτων που εφαρμόζονται, προκειμένου να επιλυθεί ένα συγκεκριμένο πρόβλημα ή να επιτευχθεί ένα καθορισμένο αποτέλεσμα. Σε όρους διαδικασιών πληροφόρησης, τα αποτελέσματα των αλγορίθμων είναι συνήθως αποφάσεις. Η ουσία της ισχύος των αλγορίθμων συχνά έγκειται στην ικανότητα των υπολογιστών να λαμβάνουν τέτοιες αποφάσεις πολύ γρήγορα και σε κλίμακα, επηρεάζοντας δυνητικά μεγάλο αριθμό ανθρώπων. Ωστόσο, στην πράξη, η αλγοριθμική λογοδοσία δεν αφορά μόνο την τεχνική πλευρά των αλγορίθμων –οι αλγόριθμοι θα πρέπει να κατανοούνται ως σύνθετα τεχνολογικά σύνολα που συνυφαίνονται με ανθρώπους, όπως οι σχεδιαστές, οι χειριστές, οι ιδιοκτήτες και οι υπεύθυνοι συντήρησης σε πολύπλοκα κοινωνικοτεχνικά συστήματα (Ananny, 2015· Seaver, 2017). Η αλγοριθμική λογοδοσία αφορά στην κατανόηση του τρόπου με τον οποίο οι άνθρωποι αυτοί ασκούν εξουσία εντός και μέσω του συστήματος και είναι, τελικά, υπεύθυνοι για τις αποφάσεις που παίρνει το σύστημα. Συχνά ένας αλγόριθμος γίνεται ενδιαφέρων για την επικαιρότητα, όταν παίρνει μια «κακή» απόφαση. Αυτό μπορεί να περιλαμβάνει έναν αλγόριθμο ο οποίος κάνει κάτι που δεν έπρεπε να κάνει ή ίσως δεν κάνει κάτι που έπρεπε να κάνει. Για τη δημοσιογραφία, η δημόσια σημασία και οι συνέπειες μιας κακής απόφασης είναι βασικοί παράγοντες. Ποια είναι η πιθανή ζημία για ένα άτομο ή για την κοινωνία; Οι κακές αποφάσεις μπορεί να επηρεάζουν άμεσα τα φυσικά πρόσωπα ή συνολικά την κοινωνία, ενισχύοντας ζητήματα όπως η δομική προκατάληψη. Οι κακές αποφάσεις μπορεί, επίσης, να είναι κοστοβόρες. Ας δούμε πώς διάφορες κακές αποφάσεις μπορούν να οδηγήσουν σε ειδήσεις.

164. Ο όρος αλγοριθμική λογοδοσία καθιερώθηκε αρχικά στο: Diakopoulos, N. (2013) «Sex, violence, and autocomplete algorithms», *Slate Magazine*, 2 Αυγούστου. Διαθέσιμο σε: https://slate.com/technology/2013/08/words-banned-from-bing-and-googles-autocomplete-algorithms.html (Προσπελάστηκε: 7 Δεκεμβρίου 2021). Ο όρος αναλύεται στο: Diakopoulos, N. (2013) «Rage against the algorithms», *The Atlantic*, 3 Οκτωβρίου. Διαθέσιμο σε: https://www.theatlantic.com/technology/archive/2013/10/rage-against-the-algorithms/280255/ (Προσπελάστηκε: 7 Δεκεμβρίου 2021).

165. Για ένα ακτιβιστικό/καλλιτεχνικό πλαίσιο, βλ.: Brain, T., & Mattu, S. (άνευ ημερομηνίας). *Algorithmic disobedience*. Διαθέσιμο σε: https://samatt.github.io/algorithmic-disobedience/#/. Για μια ακαδημαϊκή προσέγγιση που εξετάζει την αλγοριθμική ισχύ, βλ.: Bucher, T. (2018). *If . . . then: Algorithmic power and politics*. Νέα Υόρκη: Oxford University Press. Για μια ευρύτερη επιλογή των ακαδημαϊκών μελετών σχετικά με τις κριτικές μελέτες αλγορίθμων, μπορείτε να δείτε: https://socialmediacollective.org/reading-lists/critical-algorithm-studies/

Οπτικές των αλγορίθμων

Παρατηρώντας την ειδίκευση στο ρεπορτάζ των αλγορίθμων που αναπτύχθηκε τα τελευταία χρόνια στη δημοσιογραφία, καθώς και μέσω των δικών μου ερευνών για τους αλγορίθμους, έχω εντοπίσει τουλάχιστον τέσσερις κινητήριες δυνάμεις που φαίνεται να διέπουν πολλές ιστορίες αλγοριθμικής λογοδοσίας: (α) διακρίσεις και αδικίες, (β) σφάλματα ή λάθη στις προβλέψεις ή στις ταξινομήσεις, (γ) παραβιάσεις νομικών ή κοινωνικών κανόνων και (δ) κατάχρηση των αλγορίθμων από τους ανθρώπους, είτε αθέλητα είτε ηθελημένα. Παραθέτω ενδεικτικά παραδείγματα για καθεμία από αυτές στις ακόλουθες υποενότητες.

Διακρίσεις και αδικίες. Η αποκάλυψη των διακρίσεων και της αδικίας αποτελεί κοινό θέμα στο ρεπορτάζ για την αλγοριθμική λογοδοσία. Η ιστορία της *ProPublica* που άνοιξε αυτό το κεφάλαιο είναι ένα εντυπωσιακό παράδειγμα του τρόπου με τον οποίο ένας αλγόριθμος μπορεί να οδηγήσει σε συστηματικές ανισότητες στη μεταχείριση διαφορετικών ομάδων ανθρώπων. Η Northpointe, η εταιρεία που σχεδίασε τις βαθμολογίες αξιολόγησης κινδύνου (έκτοτε έχει μετονομαστεί σε Equivant), υποστήριξε ότι οι βαθμολογίες ήταν εξίσου ακριβείς για όλες τις φυλές και επομένως δίκαιες. Όμως, ο ορισμός τους για τη δικαιοσύνη απέτυχε να λάβει υπόψη του τον δυσανάλογο όγκο των σφαλμάτων που επηρέαζαν τους μαύρους πολίτες. Οι ιστορίες διακρίσεων και αδικίας εξαρτώνται από τον ορισμό της δικαιοσύνης που χρησιμοποιείται, ο οποίος μπορεί να αντανακλά διαφορετικές πολιτικές τοποθετήσεις (Lepri κ.α., 2018).

Έχω, επίσης, εργαστεί για θέματα που αποκαλύπτουν αδικίες οι οποίες οφείλονται σε αλγοριθμικά συστήματα –συγκεκριμένα, εξετάζοντας πώς η δυναμική τιμολόγησης της Uber μπορεί να επηρεάσει διαφορετικά τις γειτονιές στην Ουάσινγκτον (Stark & Diakopoulos, 2016). Με βάση τις αρχικές παρατηρήσεις των διαφορετικών χρόνων αναμονής και του τρόπου με τον οποίο οι εν λόγω χρόνοι αναμονής μετατοπίζονταν με βάση τον αλγόριθμο τιμολόγησης της Uber, υποθέσαμε ότι οι διάφορες γειτονιές θα είχαν διαφορετικά επίπεδα ποιότητας υπηρεσιών (δηλαδή χρόνο αναμονής). Με τη συστηματική δειγματοληψία των χρόνων αναμονής σε διάφορες απογραφικές περιοχές με την πάροδο του χρόνου, δείξαμε ότι οι απογραφικές περιοχές με περισσότερους έγχρωμους τείνουν να έχουν μεγαλύτερους χρόνους αναμονής για ένα αυτοκίνητο, ακόμη και όταν λαμβάνονται υπόψη άλλοι παράγοντες όπως το εισόδημα, το ποσοστό φτώχειας και η πυκνότητα του πληθυσμού στη γειτονιά. Είναι δύσκολο να αποδώσουμε το άδικο αποτέλεσμα απευθείας στον τεχνικό αλγόριθμο της Uber, επειδή το σύστημα διέπεται και από άλλους ανθρώπινους παράγοντες, όπως η συμπεριφορά και οι πιθανές προκαταλήψεις των οδηγών της. Ωστόσο, τα αποτελέσματα υποδηλώνουν ότι όταν το σύστημα εξετάζεται στο σύνολό του, παρουσιάζει ανισότητες που σχετίζονται με τα δημογραφικά χαρακτηριστικά.

Σφάλματα και λάθη. Οι αλγόριθμοι μπορούν, επίσης, να έχουν ειδησεογραφική αξία, όταν κάνουν συγκεκριμένα σφάλματα ή λάθη στις αποφάσεις ταξινόμησης, πρόβλεψης ή διαλογής. Σκεφτείτε την περίπτωση πλατφορμών, όπως το Facebook και το Google, που χρησιμοποιούν αλγοριθμικά φίλτρα, για να μειώσουν την έκθεση σε επιβλαβές περιεχόμενο, όπως η ρητορική μίσους, η βία και η πορνογραφία. Αυτό μπορεί να είναι σημαντικό για την προστασία συγκεκριμένων ευάλωτων πληθυσμών, όπως τα παιδιά, ιδίως σε προϊόντα (όπως το YouTube Kids της Google) που προωθούνται ρητά ως ασφαλή για τα παιδιά. Τα σφάλματα στον αλγόριθμο διαλογής για την εφαρμογή έχουν ειδησεογραφική αξία, διότι σημαίνουν ότι μερικές φορές τα παιδιά συναντούν ακατάλληλο ή βίαιο περιεχόμενο (Maheshwari, 2017). Τυπικά, υπάρχουν δύο είδη λαθών στα αποτελέσματα των αλγορίθμων: ψευδώς θετικά και ψευδώς αρνητικά. Στην περίπτωση του YouTube Kids, ένα ψευδώς θετικό αποτέλεσμα θα ήταν ένα βίντεο που λανθασμένα ταξινομείται ως ακατάλληλο, ενώ στην πραγματικότητα είναι απολύτως κατάλληλο για παιδιά. Ψευδώς αρνητικό είναι ένα βίντεο που χαρακτηρίζεται ως κατάλληλο, ενώ στην πραγματικότητα δεν είναι κάτι που θα θέλατε να δουν τα παιδιά.

Οι αποφάσεις ταξινόμησης επηρεάζουν τους ανθρώπους, όταν αυξάνουν ή μειώνουν τη θετική ή αρνητική μεταχείριση που δέχεται ένα άτομο. Όταν ένας αλγόριθμος επιλέγει λανθασμένα

ένα άτομο να λάβει δωρεάν παγωτό (αυξημένη θετική μεταχείριση), δεν θα ακούσετε αυτό το άτομο να διαμαρτύρεται (αν και όταν το μάθουν άλλοι, μπορεί να πουν ότι είναι άδικο). Σε γενικές γραμμές, τα λάθη έχουν ειδησεογραφική αξία, όταν οδηγούν σε αυξημένη αρνητική μεταχείριση ενός ατόμου, όπως για παράδειγμα αν ένα παιδί εκτεθεί σε ένα ακατάλληλο βίντεο. Επίσης, τα σφάλματα παρουσιάζουν ειδησεογραφικό ενδιαφέρον, όταν οδηγούν σε μείωση της θετικής μεταχείρισης ενός ατόμου, όπως όταν ένα άτομο χάνει μια ευκαιρία. Φανταστείτε έναν αγοραστή που πληροί τις προϋποθέσεις για να λάβει κάποια ειδική προσφορά, ωστόσο δεν την λαμβάνει ποτέ, επειδή ένας αλγόριθμος τον αποκλείει από λάθος. Τέλος, τα σφάλματα μπορεί να έχουν ειδησεογραφική αξία όταν προκαλούν μείωση της αιτιολογημένης αρνητικής προσοχής. Σκεφτείτε ότι ένας αλγόριθμος αξιολόγησης ποινικού κινδύνου χαρακτηρίζει λανθασμένα ένα άτομο υψηλού κινδύνου ως χαμηλού κινδύνου –ένα ψευδώς αρνητικό αποτέλεσμα. Αν και αυτό είναι υπέροχο για το συγκεκριμένο άτομο, δημιουργεί, ωστόσο, μεγαλύτερο κίνδυνο για τη δημόσια ασφάλεια, αποφυλακίζοντας κάποιον που μπορεί να εγκληματήσει ξανά.

Παραβάσεις νομικών και κοινωνικών κανόνων. Οι αλγόριθμοι πρόβλεψης μπορούν μερικές φορές να δοκιμάσουν τα όρια των καθιερωμένων νομικών ή κοινωνικών κανόνων, οδηγώντας σε άλλες ευκαιρίες και οπτικές γωνίες για κάλυψη. Σκεφτείτε για μια στιγμή την πιθανότητα αλγοριθμικής δυσφήμισης (Diakopoulos, 2013· Lewis κ.α., 2019). Η συκοφαντική δυσφήμιση ορίζεται ως «ψευδής δήλωση γεγονότων που εκθέτει ένα πρόσωπο σε μίσος, γελοιοποίηση ή περιφρόνηση, το μειώνει στην εκτίμηση των συναδέλφων του, προκαλεί την αποφυγή του ή το βλάπτει στην επιχείρηση ή στο επάγγελμά του»[166]. Τα τελευταία αρκετά χρόνια έχουν υπάρξει πολλές ιστορίες και νομικές διαμάχες σχετικά με άτομα που θεωρούν ότι έχουν δυσφημιστεί από τον αλγόριθμο αυτόματης συμπλήρωσης του Google. Μια αυτόματη συμπλήρωση μπορεί να συνδέσει το όνομα ενός ατόμου ή μιας εταιρείας με τα πάντα, από το έγκλημα και την απάτη μέχρι την πτώχευση ή τη σεξουαλική συμπεριφορά, γεγονός που ενδέχεται, στη συνέχεια, να έχει συνέπειες για την υπόληψή τους. Οι αλγόριθμοι μπορούν, επίσης, να αποτελέσουν αντικείμενο ειδησεογραφίας, όταν παραβιάζουν κοινωνικούς κανόνες, όπως η ιδιωτικότητα. Για παράδειγμα, το *Gizmodo* έχει καλύψει εκτενώς τον αλγόριθμο «People You May Know» (ΡΥΜΚ, «Άτομα που ίσως γνωρίζετε») του Facebook, ο οποίος προτείνει πιθανούς «φίλους» στην πλατφόρμα που μερικές φορές είναι ακατάλληλοι ή ανεπιθύμητοι (Hill, 2017b). Σε ένα άρθρο, οι δημοσιογράφοι εντόπισαν μια περίπτωση όπου το ΡΥΜΚ αποκάλυψε την πραγματική ταυτότητα μιας σεξεργάτριας στους πελάτες της (Hill, 2017a). Αυτό είναι προβληματικό όχι μόνο λόγω του πιθανού στίγματος το οποίο συνδέεται με τη σεξεργασία, αλλά και λόγω του κινδύνου από τους πελάτες που θα μπορούσαν να γίνουν παρενοχλητικοί.

Η δυσφήμιση και η παραβίαση της ιδιωτικής ζωής είναι μόνο δύο πιθανές οπτικές γωνίες της ιστορίας εδώ. Οι δημοσιογράφοι θα πρέπει να είναι σε επιφυλακή για μια σειρά από άλλες παραβιάσεις νομικών ή κοινωνικών κανόνων που θα μπορούσαν να δημιουργήσουν οι αλγόριθμοι σε διάφορα κοινωνικά πλαίσια. Δεδομένου ότι οι αλγόριθμοι βασίζονται αναγκαστικά σε μια ποσοτικοποιημένη εκδοχή της πραγματικότητας που ενσωματώνει μόνο ό,τι είναι μετρήσιμο ως δεδομένα, μπορεί να χάσουν πολλά από τα κοινωνικά και νομικά συμφραζόμενα που διαφορετικά θα ήταν απαραίτητα για τη λήψη μιας ορθής απόφασης. Κατανοώντας τι ποσοτικοποιεί στην πραγματικότητα ένας συγκεκριμένος αλγόριθμος για τον κόσμο –πώς «βλέπει» τα πράγματα–, οι δημοσιογράφοι μπορούν να ασκήσουν τεκμηριωμένη κριτική, φωτίζοντας τα κομμάτια που λείπουν και που θα υποστήριζαν μια απόφαση με την πληρότητα του ολοκληρωμένου πλαισίου της.

Ανθρώπινη κατάχρηση. Οι αλγοριθμικές αποφάσεις συχνά ενσωματώνονται σε ευρύτερες διαδικασίες λήψης αποφάσεων που περιλαμβάνουν ένα σύνολο ανθρώπων και αλγορίθμων οι οποίοι συνυπάρχουν σε ένα κοινωνικοτεχνικό σύστημα. Παρά την έλλειψη προσβασιμότητας σε ορισμένα από τα ευαίσθητα τεχνικά στοιχεία τους, η κοινωνικοτεχνική φύση των αλγορίθμων ανοίγει νέες ευκαιρίες για τη διερεύνηση των σχέσεων που μπορεί να έχουν οι χρήστες,

166. http://www.dmlp.org/legal-guide/defamation

οι σχεδιαστές, οι ιδιοκτήτες και άλλοι ενδιαφερόμενοι με το σύστημα εν συνόλω (Trielli & Diakopoulos, 2017). Εάν οι αλγόριθμοι χρησιμοποιούνται καταχρηστικά από τους ανθρώπους του κοινωνικοτεχνικού συνόλου, αυτό μπορεί, επίσης, να αποτελέσει είδηση. Οι σχεδιαστές των αλγορίθμων μπορούν μερικές φορές να προβλέψουν και να διατυπώσουν κατευθυντήριες γραμμές για ένα εύλογο σύνολο πλαισίων χρήσης ενός συστήματος και, έτσι, αν οι άνθρωποι τις αγνοήσουν στην πράξη, αυτό θα μπορούσε να οδηγήσει σε μια ιστορία αμέλειας ή κατάχρησης. Το θέμα για την αξιολόγηση κινδύνου από την *ProPublica* αποτελεί ένα χαρακτηριστικό παράδειγμα. Η Northpointe είχε πράγματι δημιουργήσει δύο εκδόσεις και βαθμονομήσεις του εργαλείου, μία για άνδρες και μία για γυναίκες. Τα στατιστικά μοντέλα πρέπει να εκπαιδεύονται σε δεδομένα που αντικατοπτρίζουν τον πληθυσμό στον οποίο θα χρησιμοποιηθούν και το φύλο αποτελεί σημαντικό παράγοντα στην πρόβλεψη της επανάληψης της εγκληματικής δράσης. Η κομητεία Μπρόγουαρντ έκανε, όμως, κατάχρηση της βαθμολογίας κινδύνου που σχεδιάστηκε και βαθμονομήθηκε για τους άνδρες, χρησιμοποιώντας τη και για τις γυναίκες (Larson, 2016).

Πώς να ερευνήσετε έναν αλγόριθμο

Υπάρχουν διάφορες οδοί για τη διερεύνηση της αλγοριθμικής δύναμης και καμία προσέγγιση δεν είναι υποχρεωτικά η πιο κατάλληλη. Ωστόσο, υπάρχουν σταθερές μέθοδοι προς επιλογή, που ολοένα και αυξάνονται και οι οποίες περιλαμβάνουν τα πάντα: από τεχνικές υψηλής τεχνολογίας αντίστροφης μηχανικής και επιθεώρησης κώδικα έως έλεγχο με τη χρήση αυτοματοποιημένης ή συμμετοχικής συλλογής δεδομένων, ή ακόμη και πιο χαμηλού τεχνολογικού επιπέδου προσεγγίσεις για την προτροπή και την κριτική, με βάση αλγοριθμικές αντιδράσεις (Diakopoulos, 2017· Diakopoulos, 2019)[167]. Το κάθε θέμα μπορεί να απαιτεί διαφορετική προσέγγιση ανάλογα με την οπτική γωνία και το συγκεκριμένο πλαίσιο, συμπεριλαμβανομένου του βαθμού πρόσβασης στον αλγόριθμο, στα δεδομένα του και στον κώδικά του. Για παράδειγμα, ένα αποκαλυπτικό άρθρο σχετικά με τις συστηματικές διακρίσεις μπορεί να στηριχθεί σε μεγάλο βαθμό σε μια μέθοδο ελέγχου με τη χρήση δεδομένων που συλλέγονται στο Διαδίκτυο, ενώ μια αναθεώρηση κώδικα μπορεί να είναι απαραίτητη για την επαλήθευση της ορθής εφαρμογής μιας επιδιωκόμενης πολιτικής (Lecher, 2018). Η παραδοσιακή δημοσιογραφική αναζήτηση πηγών για να μιλήσει κανείς με τους ανθρώπους της εταιρείας, όπως οι σχεδιαστές, οι προγραμματιστές και οι επιστήμονες δεδομένων, καθώς και η υποβολή αιτημάτων δημόσιων αρχείων και η εύρεση των ατόμων που επηρεάζονται, είναι τόσο σημαντική όσο ήταν πάντοτε. Δεν μπορώ να αναλύσω σε βάθος όλες αυτές τις μεθόδους σε αυτό το σύντομο κεφάλαιο, αλλά εδώ θέλω τουλάχιστον να αναφερθώ λίγο εκτενέστερα στον τρόπο με τον οποίο οι δημοσιογράφοι μπορούν να ερευνήσουν τους αλγορίθμους, χρησιμοποιώντας τον διαχειριστικό έλεγχο.

Οι τεχνικές διαχειριστικού ελέγχου χρησιμοποιούνται εδώ και δεκαετίες για τη μελέτη της κοινωνικής προκατάληψης σε συστήματα όπως οι αγορές κατοικιών και πρόσφατα προσαρμόστηκαν για τη μελέτη αλγορίθμων (Gaddis, 2017· Sandvig κ.α., 2014). Η βασική ιδέα είναι ότι, αν τα εισαγόμενα δεδομένα, που τροφοδοτούν τους αλγορίθμους, μεταβάλλονται με αρκετούς διαφορετικούς τρόπους και τα αποτελέσματα παρακολουθούνται, τότε τα εισαγόμενα δεδομένα και τα αποτελέσματα μπορούν να συσχετιστούν, για να δημιουργηθεί μια θεωρία για τον τρόπο λειτουργίας του αλγορίθμου (Diakopoulos, 2015). Εάν αναμένουμε κάποιο αποτέλεσμα το οποίο ο αλγόριθμος στρεβλώνει για κάποια εισαγόμενα δεδομένα, αυτό είναι κάτι

167. Για μια πιο ολοκληρωμένη προσέγγιση των μεθοδολογικών επιλογών, βλ.: Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Κέιμπριτζ, Μασαχουσέτη: Harvard University Press. Επίσης, βλ.: Diakopoulos, N. (2017) «Enabling accountability of algorithmic media: Transparency as a constructive and critical lens», στο Cerquitelli, T., Quercia, D. & Pasquale, F. (επιμ.) *Transparent data mining for big and small data*. Σαμ, Ελβετία: Springer International Publishing, σσ. 25-43. DOI: https://doi.org/10.1007/978-3-319-54024-5_2

που μπορεί να βοηθήσει στην πινακοποίηση των σφαλμάτων και για να διαπιστώσουμε αν τα σφάλματα είναι μεροληπτικά με συστηματικό τρόπο. Όταν οι αλγόριθμοι είναι προσβάσιμοι μέσω API[168] ή διαδικτυακών ιστοσελίδων, τα παραγόμενα δεδομένα, δηλαδή τα αποτελέσματα, μπορούν να συλλέγονται αυτόματα (Valentino-DeVries κ.α., 2012). Για τους εξατομικευμένους αλγορίθμους, οι τεχνικές διαχειριστικού ελέγχου έχουν, επίσης, συνδυαστεί με τον πληθοπορισμό («crowdsourcing»), προκειμένου να συγκεντρωθούν δεδομένα από μια σειρά ανθρώπων, καθένας εκ των οποίων μπορεί να έχει μια μοναδική «οπτική» σχετικά με τον αλγόριθμο. Το *AlgorithmWatch* στη Γερμανία χρησιμοποίησε αυτήν την τεχνική αποτελεσματικά, προκειμένου να μελετήσει την εξατομίκευση των αποτελεσμάτων αναζήτησης του Google, συλλέγοντας σχεδόν έξι εκατομμύρια αποτελέσματα αναζήτησης από περισσότερους από 4.000 χρήστες οι οποίοι μοιράστηκαν δεδομένα μέσω ενός πρόσθετου προγράμματος περιήγησης (όπως αναλύεται περαιτέρω από την Κριστίνα Έλμερ στο κεφάλαιό της[169] σε αυτόν τον τόμο)[170]. Το *Gizmodo* χρησιμοποίησε μια παραλλαγή αυτής της τεχνικής, για να βοηθήσει στην εξέταση του PYMK του Facebook. Οι χρήστες κατεβάζουν στον υπολογιστή τους ένα πρόγραμμα λογισμικού το οποίο παρακολουθεί περιοδικά τα αποτελέσματα του PYMK τοπικά στον υπολογιστή των χρηστών, διατηρώντας το απόρρητό τους. Οι δημοσιογράφοι μπορούν στη συνέχεια να ζητήσουν πληροφορίες από χρήστες που θεωρούν τα αποτελέσματά τους ανησυχητικά ή απρόσμενα (Hill & Mattu, 2018).

Ο διαχειριστικός έλεγχος αλγορίθμων δεν είναι για τους λιπόψυχους. Το έλλειμμα πληροφορίας περιορίζει την ικανότητα ενός ελεγκτή κάποιες φορές ακόμα και στο να γνωρίζει από πού να ξεκινήσει, τι να ζητήσει, πώς να ερμηνεύσει τα αποτελέσματα και πώς να εξηγήσει τα μοτίβα που βλέπει στη συμπεριφορά ενός αλγορίθμου. Υπάρχει, επίσης, η πρόκληση της επίγνωσης και του ορισμού του τι αναμένεται από έναν αλγόριθμο και πώς αυτές οι προσδοκίες μπορεί να διαφέρουν σε διάφορα πλαίσια και σύμφωνα με διαφορετικά παγκόσμια ηθικά, κοινωνικά, πολιτιστικά και νομικά πρότυπα και κανόνες. Για παράδειγμα, διαφορετικές προσδοκίες όσον αφορά στη δικαιοσύνη ισχύουν για έναν αλγόριθμο αξιολόγησης ποινικού κινδύνου σε σχέση με έναν αλγόριθμο που χρεώνει διαφορετικές τιμές για μια αεροπορική θέση. Για να εντοπίσετε ένα λάθος ή μια μεροληψία ειδησεογραφικού ενδιαφέροντος, πρέπει πρώτα να ορίσετε ποιο είναι το φυσιολογικό ή το αμερόληπτο. Ορισμένες φορές ο ορισμός αυτός προέρχεται από μια βάση δεδομένων, όπως στους διαχειριστικούς ελέγχους μας για τις δημοσιογραφικές πηγές στα αποτελέσματα αναζήτησης του Google κατά τη διάρκεια των αμερικανικών εκλογών του 2016 (Diakopoulos κ.α., 2018). Το ζήτημα της νομικής πρόσβασης στις πληροφορίες σχετικά με τους αλγορίθμους, επίσης, ανακύπτει και, φυσικά, εξαρτάται σε μεγάλο βαθμό από τη χώρα δικαιοδοσίας (Bhandari & Goodman, 2017). Στις Ηνωμένες Πολιτείες, οι νόμοι για την ελευθερία της πληροφόρησης προσδιορίζουν την πρόσβαση του κοινού σε κυβερνητικά έγγραφα, αλλά η ανταπόκριση από τις διάφορες υπηρεσίες για έγγραφα που αφορούν αλγορίθμους είναι στην καλύτερη περίπτωση ασυνεπής (Brauneis & Goodman, 2018· Diakopoulos, 2016· Fink, 2017). Ίσως πρέπει να γίνουν νομικές μεταρρυθμίσεις, ώστε να διευκολύνεται περισσότερο η πρόσβαση του κοινού σε πληροφορίες σχετικά με τους αλγορίθμους. Και αν τα ελλείμματα πληροφόρησης, οι δύσκολα διατυπώσιμες προσδοκίες και η αβέβαιη νομική πρόσβαση δεν αποτελούν αρκετή πρόκληση, θυμηθείτε ότι οι αλγόριθμοι μπορεί, επίσης, να είναι αρκετά ιδιότροποι. Η σημερινή έκδοση του αλγορίθμου μπορεί ήδη να διαφέρει από τη χθεσινή: για παράδειγμα, η Google κατά κανόνα αλλάζει τον αλγόριθμο αναζήτησης 500-600 φορές τον χρόνο. Ανάλογα με το διακύβευμα των πιθανών αλλαγών, οι αλγόριθμοι μπορεί να χρειαστεί να παρακολουθούνται με την πάροδο του χρόνου, προκειμένου να κατανοηθεί ο τρόπος με τον οποίο αλλάζουν και εξελίσσονται.

168. ΣτΜ: «Application Programming Interface» («διεπαφή προγραμματισμού εφαρμογών»).
169. ΣτΕ: Κεφάλαιο 6.5, «Ρίχνοντας Φως στους Αλγορίθμους: Συνεργατικές Έρευνες στο *Der Spiegel*»
170. https://algorithmwatch.org/filterblase-geplatzt-kaum-raum-fuer-personalisierung-bei-google-suchen-zur-bundestagswahl-2017/ (στα γερμανικά)

Προτάσεις για το μέλλον

Για να ξεκινήσετε και να αξιοποιήσετε στο έπακρο το ρεπορτάζ για την αλγοριθμική λογοδοσία, θα συνιστούσα τρία πράγματα. Πρώτον, έχουμε αναπτύξει ένα εργαλείο που ονομάζεται Algorithm Tips, το οποίο επιμελείται σχετικές μεθόδους, παραδείγματα και εκπαιδευτικές πηγές, και φιλοξενεί μια βάση δεδομένων αλγορίθμων για πιθανή έρευνα (αρχικά κάλυπτε αλγορίθμους στην ομοσπονδιακή κυβέρνηση των ΗΠΑ και, στη συνέχεια, επεκτάθηκε για να καλύψει περισσότερες περιοχές δικαιοδοσίας παγκοσμίως)[171]. Εάν αναζητάτε πόρους για να μάθετε περισσότερα και να βοηθήσετε να δρομολογηθεί ένα έργο, αυτό θα μπορούσε να είναι ένα σημείο εκκίνησης (Trielli κ.α., 2017). Δεύτερον, επικεντρωθείτε στα αποτελέσματα και στις επιπτώσεις που έχουν οι αλγόριθμοι αντί να προσπαθείτε να εξηγήσετε τον ακριβή μηχανισμό λήψης των αποφάσεών τους. Ο εντοπισμός των αλγοριθμικών διακρίσεων (δηλαδή ενός αποτελέσματος) έχει συχνά μεγαλύτερη αξία για την κοινωνία ως αρχικό βήμα από την εξήγηση του ακριβούς τρόπου με τον οποίο προέκυψαν αυτές οι διακρίσεις. Εστιάζοντας στα αποτελέσματα, οι δημοσιογράφοι μπορούν να παρέχουν μια πρώτης τάξεως διάγνωση και να σημάνουν συναγερμό, τον οποίο άλλοι ενδιαφερόμενοι μπορούν, στη συνέχεια, να εξετάσουν σε άλλα φόρουμ λογοδοσίας. Τέλος, μεγάλο μέρος των δημοσιευμένων ρεπορτάζ για την αλγοριθμική λογοδοσία που αναφέρω εδώ διεξάγεται ομαδικά και αυτό συμβαίνει για κάποιο λόγο. Το αποτελεσματικό ρεπορτάζ αλγοριθμικής λογοδοσίας απαιτεί όλες τις παραδοσιακές δεξιότητες που χρειάζονται οι δημοσιογράφοι για το ρεπορτάζ και τις συνεντεύξεις, τη γνώση ενός δημοσιογραφικού τομέα, τα αιτήματα δημόσιων αρχείων και την ανάλυση των εγγράφων που τους παρέχονται, καθώς και τη συγγραφή αποτελεσμάτων με σαφήνεια και πειστικότητα, ενώ συχνά βασίζεται επίσης σε πλήθος νέων δεξιοτήτων, όπως το scraping και ο καθαρισμός δεδομένων, ο σχεδιασμός μελετών ελέγχου και η χρήση προηγμένων στατιστικών τεχνικών. Η τεχνογνωσία σε αυτούς τους διαφορετικούς τομείς μπορεί να κατανεμηθεί σε μια ομάδα ή σε εξωτερικούς συνεργάτες, αρκεί να υπάρχει σαφής επικοινωνία, ενημέρωση και καθοδήγηση. Με αυτόν τον τρόπο, οι ειδικοί σε κάθε μέθοδο μπορούν να συνεργαστούν με διαφορετικούς ειδικούς σε διάφορους τομείς, για να κατανοήσουν την αλγοριθμική δύναμη σε μεγαλύτερη ποικιλία κοινωνικών τομέων.

Παραπομπές

Ananny, M. (2015) «Toward an ethics of algorithms», *Science, Technology & Human Values,* 41(1), σσ. 93-117.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016) «Machine bias», *ProPublica*, 23 Μαΐου. Διαθέσιμο σε: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Bhandari, E ., & Goodman, R. (2017) *Data journalism and the computer fraud and abuse act: Tips for moving forward in an uncertain landscape. Computation+Journalism Symposium,* 2017. Northwestern University, Έβανστον, Ιλινόις, 13-14 Οκτωβρίου. Διαθέσιμο σε: https://www.aclu.org/other/data-journalism-and-computer-fraud-and-abuse-act-tips-moving-forward-uncertain-landscape (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Brain, T. & Mattu, S. (άνευ ημερομηνίας). *Algorithmic disobedience*. Διαθέσιμο σε: https://samatt.github.io/algorithmic-disobedience/ (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Brauneis, R. & Goodman, E. P. (2018) «Algorithmic transparency for the smart city», *Yale Journal of Law & Technology*, 20, σσ. 103-176.

Bucher, T. (2018). *If...Then: Algorithmic power and politics.* Νέα Υόρκη: Oxford University Press.

171. http://algorithmtips.org/

Diakopoulos, N. (2013) «Algorithmic defamation: The case of the shameless autocomplete», *Tow Center for Journalism*, 6 Αυγούστου. Διαθέσιμο σε: https://towcenter.org/algorithmic-defamation-the-case-of-the-shameless-autocomplete (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Diakopoulos, N. (2015) «Algorithmic accountability: Journalistic investigation of computational power structures», *Digital Journalism*, 3(3), σσ. 398-415. DOI: https://doi.org/10.1080/21670811.2014.976411

Diakopoulos, N. (2016) «We need to know the algorithms the government uses to make important decisions about us», *The Conversation*, 24 Μαΐου. Διαθέσιμο σε: https://theconversation.com/we-need-to-know-the-algorithms-the-government-uses-to-make-important-decisions-about-us-57869 (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Diakopoulos, N. (2017) «Enabling Accountability of Algorithmic Media: Transparency as a Constructive and Critical Lens», στο T. Cerquitelli, D. Quercia & F. Pasquale (επιμ.), *Transparent data mining for Big and Small Data*. Σαμ, Ελβετία: Springer International Publishing, σσ. 25-43.

Diakopoulos, N. (2019). *Automating the News: How Algorithms are Rewriting the Media.* Κέιμπριτζ, Μασαχουσέτη: Harvard University Press.

Diakopoulos, N., Trielli, D., Stark, J. & Mussenden, S. (2018) «I vote for—How search informs our choice of candidate», στο M. Moore & D. Tambini (επιμ.), *Digital Dominance: The power of Google, Amazon, Facebook, and Apple.* Νέα Υόρκη: Oxford University Press, σσ. 320-341. Διαθέσιμο σε: https://www.academia.edu/37432634/I_Vote_For_How_Search_Informs_Our_Choice_of_Candidate (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor.* Νέα Υόρκη: St. Martin's Press.

Fink, K. (2017) «Opening the government's black boxes: Freedom of information and algorithmic accountability», *Digital Journalism*, 17(1). DOI: https://doi.org/10.1080/1369118X.2017.1330418

Gaddis, S. M. (2017) «An introduction to audit studies in the social sciences», στο M. Gaddis (επιμ.), *Audit studies: Behind the scenes with theory, method, and nuance*. Σαμ, Ελβετία: Springer International Publishing, σσ. 3-44.

Gillespie, T., & Seaver, N. (2015) «Critical algorithm studies: A reading list», *Social Media Collective*, 5 Νοεμβρίου. Διαθέσιμο σε: https://socialmediacollective.org/reading-lists/critical-algorithm-studies/ (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Hill, K. (2017a) «How Facebook outs sex workers», *Gizmodo*, 11 Οκτωβρίου. Διαθέσιμο σε: https://gizmodo.com/how-facebook-outs-sex-workers-1818861596 (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Hill, K. (2017b) «How Facebook figures out everyone you've ever met», *Gizmodo*, 7 Νοεμβρίου. Διαθέσιμο σε: https://gizmodo.com/how-facebook-figures-out-everyone-youve-ever-met-1819822691 (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Hill, K., & Mattu, S. (2018) «Keep track of who Facebook thinks you know with this nifty tool», *Gizmodo*, 10 Ιανουαρίου. Διαθέσιμο σε: https://gizmodo.com/keep-track-of-who-facebook-thinks-you-know-with-this-ni-1819422352 (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Larson, J. (2016) «Machine bias with Jeff Larson», *Data Stories* [Podcast], 20 Οκτωβρίου. Διαθέσιμο σε: https://datastori.es/85-machine-bias-with-jeff-larson/ (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Larson, J., Mattu, S., Kirchner, L., & Angwin, J. (2016) «How we analyzed the COMPAS recidivism algorithm», *ProPublica*, 23 Μαΐου. Διαθέσιμο σε: https://www.propublica.org/article/how-we-

analyzed-the-compas-recidivism-algorithm (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Lecher, C. (2018) «What happens when an algorithm cuts your health care», *The Verge*, 21 Μαρτίου. Διαθέσιμο σε: https://www.theverge.com/2018/3/21/17144260/healthcare-medicaid-algorithm-arkansas-cerebral-palsy (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Lepri, B., Oliver, N., Letouzé, E., Pentland, A. & Vinck, P. (2018) «Fair, transparent, and accountable algorithmic decision-making processes», *Philosophy & Technology,* 31(4), σσ. 611-627. DOI: https://doi.org/10.1007/s13347-017-0279-x

Lewis, S. C., Sanders, A. K. & Carmody, C. (2019) «Libel by algorithm? Automated journalism and the threat of legal liability», *Journalism and Mass Communication Quarterly,* 96(1), σσ. 60-81. DOI: https://doi.org/10.1177/1077699018755983

Maheshwari, S. (2017) «On YouTube Kids, startling videos slip past filters», *The New York Times,* 4 Νοεμβρίου. Διαθέσιμο σε: https://www.nytimes.com/2017/11/04/business/media/youtube-kids-paw-patrol.html?_r=0 (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Πόρτλαντ, Όρεγκον: Broadway Books.

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information.* Κέιμπριτζ, Μασαχουσέτη: Harvard University Press.

Sandvig, C., Hamilton, K., Karahalios, K. & Langbort, C. (2014). *Auditing algorithms: Research methods for detecting discrimination on Internet platforms, International Communication Association preconference on Data and Discrimination Converting Critical Concerns into Productive Inquiry.* Σιάτλ, Ουάσινγκτον, 22 Μαΐου.

Seaver, N. (2017) «Algorithms as culture: Some tactics for the ethnography of algorithmic systems», *Big Data & Society*, 4(2). DOI: https://doi.org/10.1177/2053951717738104

Stark, J., & Diakopoulos, N. (2016) «Uber seems to offer better service in areas with more White people. That raises some tough questions», *The Washington Post,* 10 Μαρτίου. Διαθέσιμο σε: https://www.washingtonpost.com/news/wonk/wp/2016/03/10/uber-seems-to-offer-better-service-in-areas-with-more-white-people-that-raises-some-tough-questions/ (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Trielli, D., & Diakopoulos, N. (2017) «How to report on algorithms even if you're not a data whiz», *Columbia Journalism Review*, 30 Μαίου. Διαθέσιμο σε: https://www.cjr.org/tow_center/algorithms-reporting-algorithmtips.php (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Trielli, D., Stark, J., & Diakopoulos, N. (2017). Algorithm tips: A resource for algorithmic accountability in Government, *Computation + Journalism Symposium*, 2017. Northwestern University, Έβανστον, Ιλινόις, 13-14 Οκτωβρίου.

Valentino-DeVries, J., Singer-Vine, J. & Soltani, A. (2012) «Websites vary prices, deals based on users' information», *The Wall Street Journal*, 24 Δεκεμβρίου. Διαθέσιμο σε: https:// www.wsj.com/articles/SB10001424127887323777204578189391813881534 (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Ο συγγραφέας

Ο Νίκολας Ντιακόπουλος είναι Επίκουρος Καθηγητής στο Τμήμα Επικοινωνίας του Πανεπιστημίου Northwestern, όπου διευθύνει το Εργαστήριο Υπολογιστικής Δημοσιογραφίας. Είναι ο συγγραφέας του βιβλίου *Automating the News: How Algorithms are Rewriting the Media* (Harvard University Press, 2019).

6.2
Αφήγηση Ιστοριών με τα Μέσα Κοινωνικής Δικτύωσης

Lam Thuy Vo

Περίληψη: Αυτό το κεφάλαιο αποτελεί μια εισαγωγή στη διερεύνηση των μέσων κοινωνικής δικτύωσης.

Λέξεις κλειδιά: μέσα κοινωνικής δικτύωσης, παραπληροφόρηση, δημοσιογραφία δεδομένων, ανάλυση δεδομένων, οπτικοποίηση δεδομένων, bot

Έχουμε γίνει οι μεγαλύτεροι παραγωγοί δεδομένων στην ανθρώπινη ιστορία[172]. Σχεδόν κάθε διαδικτυακό κλικ, κάθε κίνηση στο τάμπλετ μας και κάθε πάτημα στο κινητό μας τηλέφωνο παράγει ένα σημείο δεδομένων σε ένα εικονικό αποθετήριο. Το Facebook παράγει δεδομένα για τη ζωή περισσότερων από δύο δισεκατομμυρίων ανθρώπων. Το Twitter καταγράφει τις δραστηριότητες περισσότερων από 330 εκατομμυρίων χρηστών μηνιαίως. Μια μελέτη του MIT διαπίστωσε ότι ο μέσος Αμερικανός υπάλληλος γραφείου παράγει 5GB δεδομένων κάθε μέρα (Tucker, 2013). Αυτό γινόταν το 2013 και δεν έχουμε επιβραδύνει έκτοτε. Καθώς όλο και περισσότεροι άνθρωποι ζουν τη ζωή τους στο Διαδίκτυο και καθώς τα έξυπνα κινητά τηλέφωνα διεισδύουν σε περιοχές σε όλο τον κόσμο που ως χθες ήταν αποσυνδεδεμένες, αυτή η συλλογή ιστοριών θα μεγαλώνει συνεχώς.

Πολλοί ερευνητές τείνουν να αντιμετωπίζουν κάθε χρήστη των μέσων κοινωνικής δικτύωσης όπως θα αντιμετώπιζαν ένα ξεχωριστό υποκείμενο –ως μοναδικά και μεμονωμένα σημεία επαφής. Εάν το κάνουμε, όμως, αυτό με μια χούφτα χρηστών και τις μεμονωμένες αναρτήσεις τους, αγνοούμε τις δυνατότητες εκατοντάδων εκατομμυρίων άλλων χρηστών και τις αλληλεπιδράσεις μεταξύ τους. Υπάρχουν πολλές ιστορίες που θα μπορούσαν να ειπωθούν μέσα από τις τεράστιες ποσότητες δεδομένων που παράγονται από τους χρήστες και τις πλατφόρμες των μέσων κοινωνικής δικτύωσης, επειδή οι ερευνητές και οι δημοσιογράφοι μόλις τώρα αρχίζουν να αποκτούν την τεχνογνωσία επεξεργασίας δεδομένων μεγάλης κλίμακας και τις αναλυτικές τεχνικές που απαιτούνται για την αξιοποίησή τους.

Πρόσφατα γεγονότα έδειξαν, επίσης, ότι καθίσταται ζωτικής σημασίας για τους δημοσιογράφους να αποκτήσουν καλύτερη αντίληψη του κοινωνικού Ιστού. Η ρωσική ανάμειξη στις προεδρικές εκλογές του 2016 στις ΗΠΑ και στο Brexit, η επικίνδυνη εξάπλωση της αντιμουσουλμανικής ρητορικής μίσους στο Facebook σε χώρες της Ευρώπης και στη Μιανμάρ και η δυναμική χρήση του Twitter από παγκόσμιους ηγέτες –όλες αυτές οι εξελίξεις δείχνουν ότι υπάρχει μια διαρκώς αυξανόμενη ανάγκη να φτάσουμε σε ένα ικανό επίπεδο αλφαβητισμού γύρω από τη χρησιμότητα και τις παγίδες των δεδομένων των μέσων κοινωνικής δικτύωσης συνολικά.

172. Προηγούμενες εκδόσεις αυτού του κεφαλαίου: https://source.opennews.org/articles/what-buzzfeed-news-learned-after-year-mining-data-/, http://www.niemanlab.org/2016/12/the-primary-source-in-the-age-of-mechanical-multiplication/

Πώς μπορούν οι δημοσιογράφοι να χρησιμοποιήσουν τα δεδομένα των μέσων κοινωνικής δικτύωσης;

Ενώ υπάρχουν πολλοί διαφορετικοί τρόποι με τους οποίους τα μέσα κοινωνικής δικτύωσης μπορούν να βοηθήσουν στην πληροφόρηση, ενδέχεται να είναι χρήσιμο να εξετάσουμε τα δεδομένα που μπορούμε να συλλέξουμε από τις πλατφόρμες κοινωνικής δικτύωσης μέσα από δύο πρίσματα.

Πρώτον, τα μέσα κοινωνικής δικτύωσης μπορούν να χρησιμοποιηθούν ως μέσο για την καλύτερη κατανόηση των ατόμων και των ενεργειών τους. Είτε πρόκειται για δημόσιες διακηρύξεις είτε για ιδιωτικές επικοινωνίες μεταξύ ατόμων, πολλές από τις ενέργειες των ανθρώπων, όπως διαμεσολαβούνται και διαδίδονται μέσω της τεχνολογίας στις μέρες μας, αφήνουν ίχνη στο Διαδίκτυο, τα οποία μπορούν να αξιοποιηθούν για πληροφορίες. Αυτό είναι ιδιαίτερα χρήσιμο, όταν εξετάζουμε πολιτικούς και άλλες σημαντικές προσωπικότητες, των οποίων οι δημόσιες γνώμες μπορεί να είναι ενδεικτικές των στρατηγικών τους ή να έχουν συνέπειες στην πραγματική ζωή, όπως η πτώση των τιμών των μετοχών ή η απόλυση σημαντικών ατόμων.

Δεύτερον, ο Ιστός μπορεί να θεωρηθεί ως ένα αυτοτελές οικοσύστημα στο οποίο οι ιστορίες λαμβάνουν χώρα σε κοινωνικές πλατφόρμες (αν και εξακολουθούν να καθοδηγούνται από ανθρώπινες και αυτοματοποιημένες ενέργειες). Οι εκστρατείες παραπληροφόρησης, τα αλγοριθμικά διαστρεβλωμένα σύμπαντα πληροφοριών και οι επιθέσεις «τρολαρίσματος» είναι φαινόμενα μοναδικά στα μέσα κοινωνικής δικτύωσης.

Περιπτωσιολογικές μελέτες

Αντί να συζητάμε αυτού του είδους τις ιστορίες αφηρημένα, μπορεί να είναι πιο χρήσιμο να κατανοήσουμε τα δεδομένα των μέσων κοινωνικής δικτύωσης, στο πλαίσιο χρήσης τους για την αφήγηση συγκεκριμένων ιστοριών. Στις επόμενες ενότητες αναλύονται ορισμένα δημοσιογραφικά έργα που χρησιμοποίησαν δεδομένα από τα μέσα κοινωνικής δικτύωσης.

Κατανοώντας τα δημόσια πρόσωπα: δεδομένα από τα μέσα κοινωνικής δικτύωσης για ρεπορτάζ λογοδοσίας

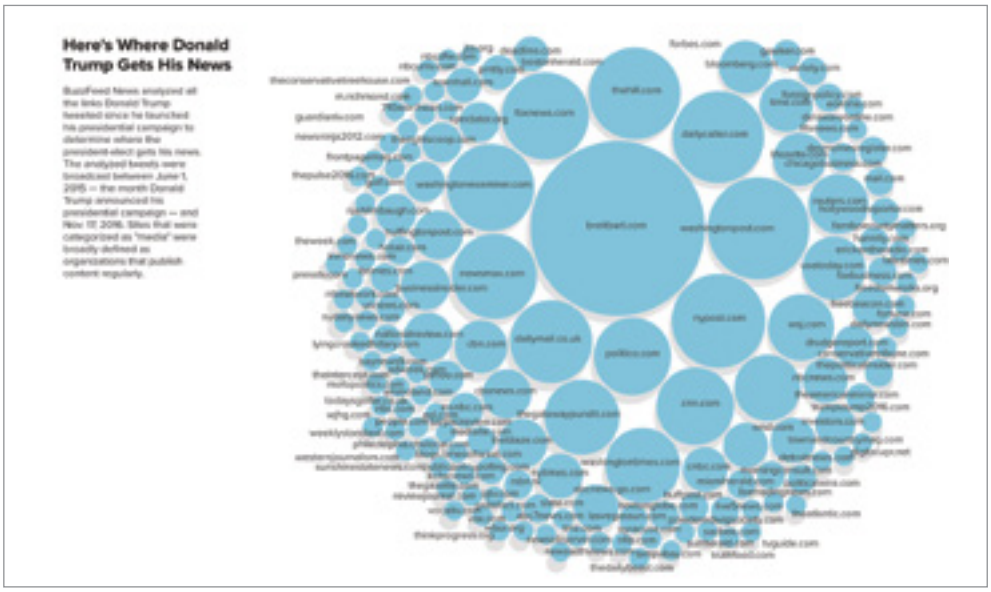
Τόσο για τα δημόσια πρόσωπα όσο και για τους καθημερινούς ανθρώπους, τα μέσα κοινωνικής δικτύωσης έχουν γίνει ένας τρόπος για να απευθύνονται στο κοινό με άμεσο τρόπο. Οι ενημερώσεις κατάστασης, τα tweet και οι αναρτήσεις μπορούν να χρησιμεύσουν ως τρόποι παράκαμψης παλαιότερων μηχανισμών προβολής, όπως οι συνεντεύξεις στα μέσα ενημέρωσης, τα δελτία Τύπου ή οι συνεντεύξεις Τύπου.

Για τους πολιτικούς, ωστόσο, αυτές οι δημόσιες ανακοινώσεις –αυτές οι προβολές του εαυτού τους– μπορεί να γίνουν δεσμευτικές δηλώσεις και, στην περίπτωση ισχυρών πολιτικών προσώπων, μπορεί να γίνουν προάγγελοι πολιτικών που δεν έχουν ακόμη τεθεί σε εφαρμογή.

Επειδή η δουλειά ενός πολιτικού είναι εν μέρει να αντιμετωπίζει το κοινό, η έρευνα στους λογαριασμούς ενός πολιτικού στα μέσα κοινωνικής δικτύωσης μπορεί να μας βοηθήσει να κατανοήσουμε καλύτερα την ιδεολογική του στάση. Για ένα θέμα, ο συνάδελφός μου Τσάρλι Γουόρζελ και εγώ συλλέξαμε και αναλύσαμε περισσότερα από 20.000 tweet του Ντόναλντ Τραμπ, προκειμένου να απαντήσουμε στο εξής ερώτημα: τι είδους πληροφορίες διαδίδει και πώς αυτές οι πληροφορίες μπορούν να χρησιμεύσουν ως δείκτης για το είδος των πληροφοριών που μπορεί να καταναλώνει;

Τα σημεία δεδομένων των μέσων κοινωνικής δικτύωσης δεν αποτελούν μια πλήρη εικόνα του ποιοι πραγματικά είμαστε, εν μέρει λόγω της επιτελεστικής τους φύσης και εν μέρει επειδή αυτά τα σύνολα δεδομένων είναι ελλιπή και, ως εκ τούτου, ανοιχτά σε ατομικές ερμηνείες. Μπορούν, ωστόσο, να βοηθήσουν συμπληρωματικά: η σχέση του προέδρου Τραμπ με το Breitbart

στο Διαδίκτυο, όπως φαίνεται στην Εικόνα 6.2.1, ήταν ένας πρώτος δείκτης για τους ισχυρούς δεσμούς του με τον Στιβ Μπάνον στην πραγματική ζωή. Η αναδημοσίευση μικρότερων συντηρητικών ιστολογίων, όπως το *Conservative Tree House* και το *News Ninja* 2012, ίσως να μαρτυρούσε τη δυσπιστία του απέναντι στα «κυρίαρχα μέσα ενημέρωσης».



Εικόνα 6.2.1: Στιγμιότυπο οθόνης από υπερσυνδέσμους προς μέσα ενημέρωσης σε tweet του Τραμπ κατά της διάρκειας της εκστρατείας για τις προεδρικές εκλογές του 2016.

Ανιχνεύοντας τις ανθρώπινες ενέργειες

Ενώ οι δημόσιες και ημι-δημόσιες επικοινωνίες, όπως τα tweet και οι ανοιχτές αναρτήσεις στο Facebook, μπορούν να δώσουν πληροφορίες για το πώς οι άνθρωποι παρουσιάζουν τον εαυτό τους στους άλλους, υπάρχουν, επίσης, δεδομένα που βρίσκονται στις κοινωνικές πλατφόρμες πίσω από κλειστούς τοίχους, όπως τα ιδιωτικά μηνύματα, οι αναζητήσεις στο Google ή τα δεδομένα γεωγραφικής θέσης.

Ο Κρίστιαν Ράντερ, συνιδρυτής του OKCupid και συγγραφέας του βιβλίου *Dataclysm*, διατυπώνει μια μάλλον εύστοχη περιγραφή αυτού του είδους δεδομένων: πρόκειται για στατιστικά στοιχεία της συμπεριφοράς μας που καταγράφονται, όταν «νομίζουμε ότι κανείς δεν μας παρακολουθεί» (Rudder, 2014).

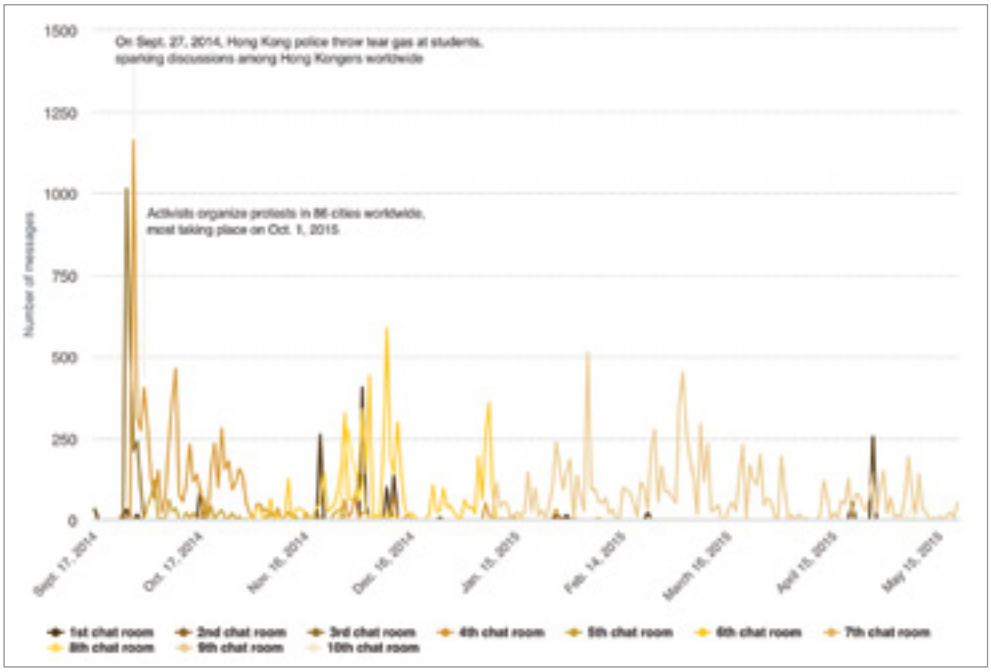
Λόγω της χρήσης μιας πλατφόρμας κοινωνικής δικτύωσης, ένα άτομο καταλήγει να παράγει διαχρονικά δεδομένα της δικής του συμπεριφοράς. Παρά το γεγονός ότι είναι δύσκολο να συμπεράνουμε πολλά από αυτά τα προσωπικά δεδομένα πέρα από το φυσικό αντικείμενο του προσώπου που τα παρήγαγε, αυτού του είδους τα δεδομένα μπορούν να είναι εξαιρετικά ισχυρά, όταν προσπαθούμε να αφηγηθούμε την ιστορία ενός ατόμου. Συχνά μου αρέσει να αναφέρομαι σε αυτό το είδος προσέγγισης ως «ποσοτικοποιημένη selfie», ένας όρος που επινόησε η Μορίν Ο'Κόνορ για λογαριασμό μου, περιγράφοντας κάποια από τα έργα μου.

Πάρτε, για παράδειγμα, την ιστορία του Τζέφρι Νο. Όταν άρχισαν οι διαδηλώσεις υπέρ της δημοκρατίας στη γενέτειρά του, το Χονγκ Κονγκ, στις αρχές Σεπτεμβρίου του 2014, ο Νο, φοιτητής του Πανεπιστημίου της Νέας Υόρκης με καταγωγή από το Χονγκ Κονγκ, αισθάνθηκε υποχρεωμένος να

δράσει. Ο Νο άρχισε να μιλάει με άλλους ομογενείς από το Χονγκ Κονγκ στη Νέα Υόρκη και στην Ουάσιγκτον. Κατέληξε να διοργανώνει διαμαρτυρίες σε 86 πόλεις σε ολόκληρο τον κόσμο και η ιστορία του είναι χαρακτηριστικό παράδειγμα πολλών κινημάτων που ξεκινούν από την παγκόσμια κατακραυγή για ένα ζήτημα. Για αυτό το θέμα του *Al Jazeera America*, ο Νο μάς επέτρεψε να διερευνήσουμε το προσωπικό του ιστορικό στο Facebook –ένα αρχείο που κάθε χρήστης του Facebook μπορεί να κατεβάσει από την πλατφόρμα (Vo, 2015). Ψάξαμε τα μηνύματα που αντάλλαξε με έναν άλλο κεντρικό διοργανωτή στο Χονγκ Κονγκ και βρήκαμε 10 διαφορετικά chat room στα οποία οι δυο τους και άλλοι διοργανωτές αντάλλασσαν σκέψεις σχετικά με τις πολιτικές τους δραστηριότητες.

Το παρακάτω διάγραμμα (Εικόνα 6.2.2) καταγράφει της διακυμάνσεις των επικοινωνιών τους. Πρώτα υπάρχει μια έξαρση της επικοινωνίας, όταν ένα ειδησεογραφικό γεγονός προκάλεσε τη δημόσια κατακραυγή –η αστυνομία του Χονγκ Κονγκ έριξε δακρυγόνα σε μια ειρηνική διαδήλωση. Στη συνέχεια, εμφανίζεται ένα chat room, εκείνο με το μπεζ χρώμα, το οποίο έγινε το chat room όπου οι βασικοί διοργανωτές σχεδίαζαν πολιτικές δραστηριότητες πολύ μετά τις αρχικές ειδήσεις.

Δεδομένου ότι το μεγαλύτερο μέρος του σχεδιασμού τους έλαβε χώρα μέσα σε αυτά τα chat room, μπορέσαμε, επίσης, να αφηγηθούμε τη στιγμή που ο Νο συνάντησε για πρώτη φορά τον συνδιοργανωτή του, Άντζελ Γιο. Ο ίδιος ο Νο δεν ήταν σε θέση να ανακαλέσει τις πρώτες τους συνομιλίες, αλλά χάρη στο αρχείο του Facebook καταφέραμε να ανασυνθέσουμε την πρώτη συζήτηση που είχε ο Νο με τον Γιο.



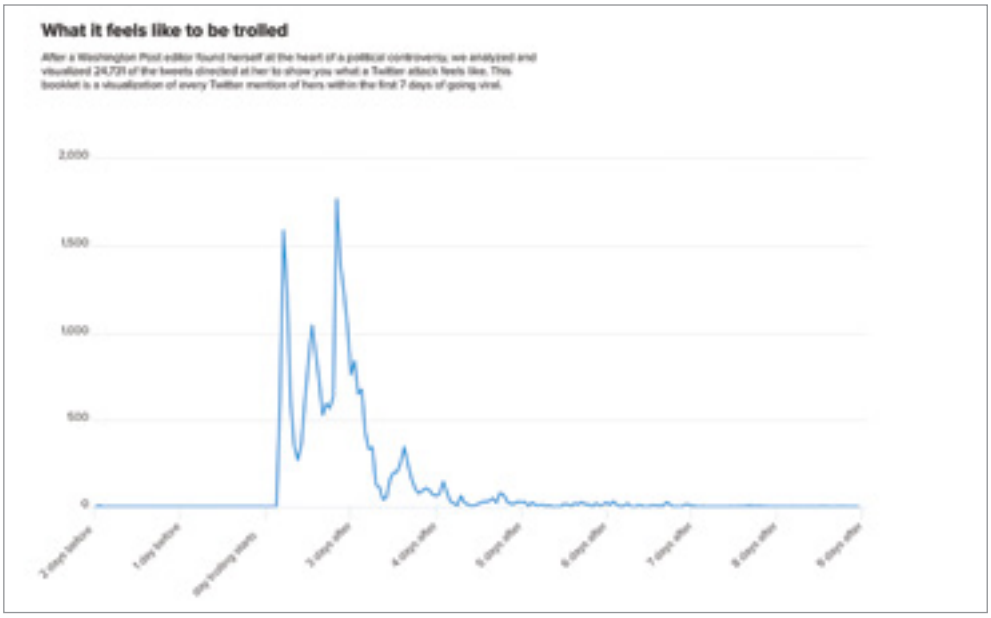
Εικόνα 6.2.2: Το Facebook group «United for Democracy: Global Solidarity with Hong Kong». Τα δεδομένα από το Facebook είναι ευγενική προσφορά του Τζέφρι Νο. Πηγή: *BuzzFeed News*.

Αν και είναι σαφές ότι η εξέλιξη του Νο ως πολιτικού διοργανωτή είναι αυτή ενός μεμονωμένου ατόμου και σε καμία περίπτωση αντιπροσωπευτική για κάθε άτομο που συμμετείχε στο κίνημά του, είναι, ωστόσο, ενδεικτική του είδους της πορείας που μπορεί να ακολουθήσει ένας πολιτικός διοργανωτής στην ψηφιακή εποχή.

Φαινόμενα που εμφανίζονται αποκλειστικά στα διαδικτυακά οικοσυστήματα

Πολλές από τις αλληλεπιδράσεις μας μετακινούνται αποκλειστικά σε διαδικτυακές πλατφόρμες. Αν και ένα μεγάλο μέρος της κοινωνικής μας συμπεριφοράς στο Διαδίκτυο είναι αλληλένδετη με τη συμπεριφορά μας στον φυσικό κόσμο, τα διαδικτυακά μας περιβάλλοντα παραμένουν αρκετά ιδιαίτερα, επειδή οι άνθρωποι στο Διαδίκτυο υποστηρίζονται από ισχυρά εργαλεία.

Καταρχάς, υπάρχει ο εκφοβισμός. Ο εκφοβισμός αναμφισβήτητα είναι τόσο παλιός όσο και η ανθρωπότητα. Τώρα οι εκφοβιστές υποστηρίζονται, όμως, από χιλιάδες άλλους εκφοβιστές, οι οποίοι μπορούν να κληθούν εν ριπή οφθαλμού. Οι εκφοβιστές έχουν πρόσβαση στις μηχανές αναζήτησης και στα ψηφιακά ίχνη της ζωής ενός ατόμου, που μερικές φορές φτάνουν μέχρι τις διαδικτυακές περσόνες αυτού του ατόμου. Και έχουν τον τρόπο να δυναμώνουν την ένταση – ένας εκφοβιστής που φωνάζει από την άλλη άκρη του διαδρόμου δεν είναι τόσο εκκωφαντικός όσο χιλιάδες από αυτούς που επιτίθενται ταυτόχρονα. Αυτή είναι η φύση του «τρολαρίσματος».



Εικόνα 6.2.3: Διάγραμμα που δείχνει πόσες φορές οι χρήστες του Twitter έκαναν mention στην Ντόρις Τρουόνγκ, ξεκινώντας από την ημέρα της επίθεσης. Πηγή: *BuzzFeed News*.

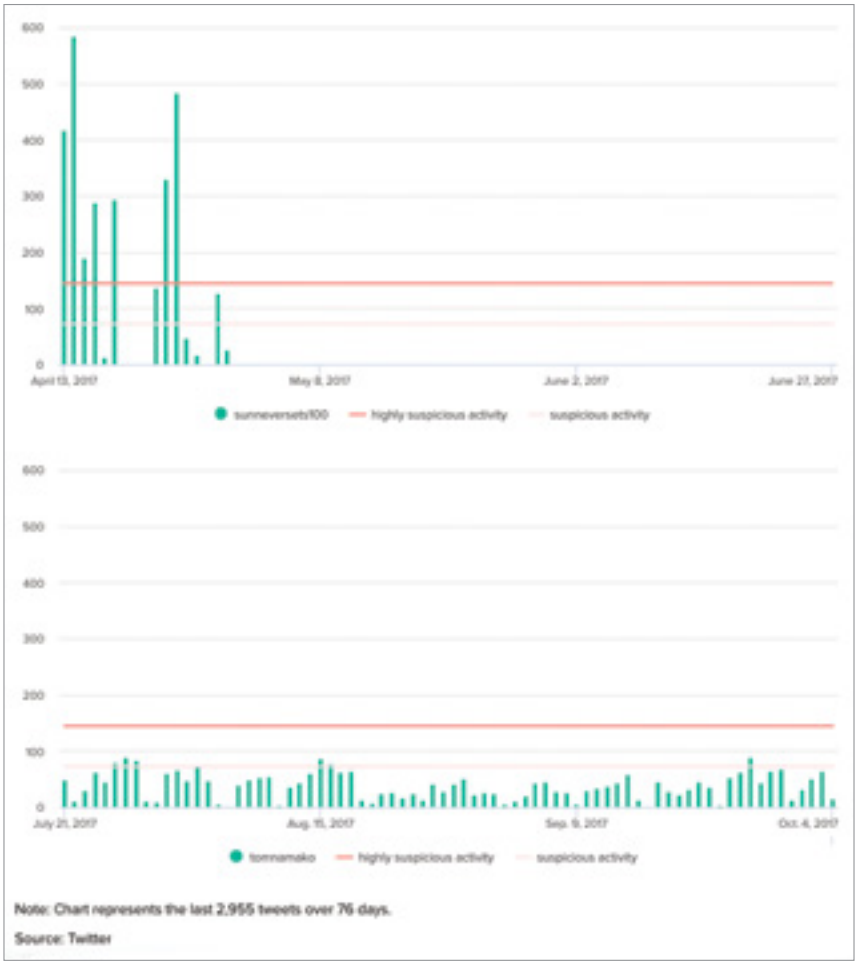
Η συντάκτρια της *Washington Post*, Ντόρις Τρουόνγκ, για παράδειγμα, βρέθηκε στο επίκεντρο μιας πολιτικής διαμάχης στο Διαδίκτυο. Για μερικές ημέρες, τα τρολ (και ένας ικανός αριθμός ανθρώπων που την υπερασπίστηκαν) της απηύθυναν 24.731 mention στο Twitter. Υπάρχει ένα όριο στο πόσο μπορεί να αντέξει κάποιος να του πετούν βιτριόλι στο Διαδίκτυο πριν να αρχίσει να τον καταβάλλει ένας συναισθηματικός αντίκτυπος.

Το «τρολάρισμα», όπως και πολλές άλλες διαδικτυακές επιθέσεις, έχει μετατραπεί σε πρόβλημα που μπορεί να πλήξει οποιοδήποτε άτομο –διάσημο ή μη. Από τις κριτικές του Yelp για επιχειρήσεις που γίνονται viral (όπως το ζαχαροπλαστείο που αρνήθηκε να ετοιμάσει γαμήλια τούρτα για ένα γκέι ζευγάρι) μέχρι τους τρόπους με τους οποίους η viral δημοσιότητα προκάλεσε την απόλυση και τη δημόσια διαπόμπευση της Τζαστίν Σάκο, μιας υπεύθυνης δημοσίων σχέσεων που έκανε ένα ατυχές αστείο για τον HIV και τους Νοτιοαφρικανούς λίγο πριν να απογειωθεί για

μια διππειρωτική πτήση– πολλές ιστορίες που επηρεάζουν την καθημερινότητά μας λαμβάνουν χώρα στο Διαδίκτυο στις μέρες μας.

Πόλεμοι της πληροφορίας

Η εμφάνιση και η πανταχού παρούσα χρήση των μέσων κοινωνικής δικτύωσης έχουν επιφέρει ένα νέο φαινόμενο στη ζωή μας που περιγράφεται με τον όρο virality[173].



Εικόνα 6.2.4: Το *BuzzFeed News* σύγκρινε τα δεδομένα ενός δικού του, ανθρώπου-συντάκτη στο Twitter, του @tomnamako, και τα δεδομένα διαφόρων λογαριασμών που εμφάνιζαν δραστηριότητα που έμοιαζε με bot, για να αναδείξει τις διαφορές τους στις περσόνες και στη συμπεριφορά. Το πρώτο διάγραμμα δείχνει ότι τα τελευταία 2.955 tweet του συντάκτη του *BuzzFeed News* κατανέμονται ομοιόμορφα σε αρκετούς μήνες. Ο καθημερινός αριθμός των tweet του σπάνια ξεπερνά τα 72 tweet την ημέρα, τα οποία το Digital Forensics Research Lab χαρακτήρισε ως ύποπτο επίπεδο δραστηριότητας. Το δεύτερο γράφημα δείχνει τα τελευταία 2.955 tweet του bot. Εξαπέλυε συστηματικά έναν ύποπτο αριθμό tweet, φτάνοντας τα 584 σε μία ημέρα, και σταματώντας, κατόπιν, απότομα. Πηγή: *BuzzFeed News*.

173. ΣτΜ: Η τάση μιας πληροφορίας, σε μορφή εικόνας, βίντεο ή κειμένου, να κυκλοφορήσει ταχύτατα και ευρέως από τον έναν χρήστη του Ίντερνετ στον άλλον.

Η κοινοποίηση στα μέσα κοινωνικής δικτύωσης επιτρέπει σε οποιοδήποτε είδος περιεχομένου να το δουν όχι μόνο μερικές εκατοντάδες αλλά εκατομμύρια άνθρωποι, χωρίς δαπανηρές εκστρατείες μάρκετινγκ ή την αγορά τηλεοπτικού χρόνου.

Αυτό, όμως, σημαίνει ότι πολλοί άνθρωποι έχουν, επίσης, βρει τρόπους να εξαπατούν τους αλγορίθμους με ψεύτικους ή αγορασμένους follower, καθώς και με (ημι)αυτοματοποιημένους λογαριασμούς όπως τα bot και τα cyborg (Vo, 2017a).

Τα bot δεν είναι εξαρχής κακά: υπάρχουν πολλά bot που μπορούν να μας ευχαριστήσουν με ιδιόρρυθμα χαϊκού ή με συμβουλές αυτοφροντίδας. Ωστόσο, όπως μου είπε ο Μπεν Νίμο, υπότροφος του Atlantic Council, ο οποίος ερευνά εδώ και χρόνια τους στρατούς των bot, για ένα θέμα του *BuzzFeed*: «[Τα bot] έχουν τη δυνατότητα να διαστρεβλώσουν σοβαρά οποιαδήποτε συζήτηση [...]. Μπορούν να κάνουν μια ομάδα έξι ατόμων να φαίνεται σαν 46.000 άτομα». Οι ίδιες οι πλατφόρμες κοινωνικής δικτύωσης βρίσκονται σε ένα κομβικό σημείο της ύπαρξής τους, όπου πρέπει να αναγνωρίσουν την ευθύνη τους για τον καθορισμό και τον περιορισμό αυτού που μπορεί να θεωρηθεί «προβληματικό bot». Εν τω μεταξύ, οι δημοσιογράφοι θα πρέπει να αναγνωρίσουν τη συνεχώς αυξανόμενη παρουσία των μη-ανθρώπων και τη δύναμή τους στο Διαδίκτυο.

Για ένα επεξηγηματικό κομμάτι σχετικά με τους αυτοματοποιημένους λογαριασμούς θέλαμε να συγκρίνουμε τα tweet ενός ανθρώπου με εκείνα ενός bot (Vo, 2017b). Αν και δεν υπάρχει ασφαλής τρόπος για να προσδιοριστεί πραγματικά αν ένας λογαριασμός λειτουργεί μέσω ενός προγράμματος και συνεπώς δεν είναι άνθρωπος, υπάρχουν τρόποι να εξετάσετε διάφορα χαρακτηριστικά ενός χρήστη, για να δείτε αν η συμπεριφορά του μπορεί να είναι ύποπτη. Ένα από τα χαρακτηριστικά που αποφασίσαμε να εξετάσουμε είναι η δραστηριότητα ενός λογαριασμού.

Για τον σκοπό αυτό, συγκρίναμε τη δραστηριότητα ενός πραγματικού ατόμου με αυτή ενός bot. Κατά τη διάρκεια της πιο δραστήριας ώρας της πιο δραστήριας ημέρας του, το bot που εξετάσαμε έγραψε στο Twitter πάνω από 200 φορές. Ο αντίστοιχος άνθρωπος έγραψε στο Twitter μόνο 21 φορές.

Πώς να συλλέξετε δεδομένα από τα μέσα κοινωνικής δικτύωσης

Υπάρχουν, σε γενικές γραμμές, τρεις διαφορετικοί τρόποι συλλογής δεδομένων από τα μέσα κοινωνικής δικτύωσης: API, προσωπικά αρχεία και scraping.

Το είδος των δεδομένων που παρέχονται μέσα από επίσημες οδούς, όπως οι ροές δεδομένων API, είναι πολύ περιορισμένο. Παρά το γεγονός ότι έχουν στην κατοχή τους ολόκληρες αποθήκες δεδομένων σχετικά με τη συμπεριφορά των καταναλωτών, οι εταιρείες μέσων κοινωνικής δικτύωσης διαθέτουν μόνο ένα ελάχιστο μέρος τους μέσω API. Για το Facebook, οι ερευνητές μπορούσαν κάποτε να λαμβάνουν δεδομένα για δημόσιες σελίδες και ομάδες, αλλά αυτό δεν είναι πλέον εφικτό, αφού η εταιρεία εφάρμοσε περιορισμούς στη διάθεση αυτών των δεδομένων ως απάντηση στο σκάνδαλο της Cambridge Analytica. Για το Twitter, η πρόσβαση αυτή συχνά περιορίζεται σε έναν καθορισμένο αριθμό tweet από το χρονολόγιο ενός χρήστη ή σε ένα καθορισμένο χρονικό πλαίσιο για αναζήτηση.

Στη συνέχεια, υπάρχουν περιορισμοί όσον αφορά στο είδος των δεδομένων που μπορούν να ζητήσουν οι χρήστες σχετικά με τη δική τους διαδικτυακή περσόνα και συμπεριφορά. Ορισμένες υπηρεσίες, όπως το Facebook ή το Twitter, επιτρέπουν στους χρήστες να κατεβάζουν ένα ιστορικό των δεδομένων που συνθέτουν τον διαδικτυακό τους εαυτό –τις αναρτήσεις τους, τα μηνύματά τους ή τις φωτογραφίες του προφίλ τους–, αλλά αυτό το αρχείο δεδομένων δεν περιλαμβάνει πάντα όλα όσα έχει στη διάθεσή της κάθε εταιρεία κοινωνικής δικτύωσης.

Για παράδειγμα, οι χρήστες μπορούν να δουν μόνο τις διαφημίσεις στις οποίες έχουν κάνει κλικ κατά τη διάρκεια των τελευταίων τριών μηνών, κάτι που καθιστά πραγματικά δύσκολο το να δουν αν έχουν κάνει ή όχι κλικ σε μια ανάρτηση που χρηματοδοτείται από τη Ρωσία.

Τέλος, η εξαγωγή δεδομένων των μέσων κοινωνικής δικτύωσης από τις πλατφόρμες με scraping συχνά έρχεται σε αντίθεση με τους όρους χρήσης των υπηρεσιών. Το scraping σε μια πλατφόρμα κοινωνικής δικτύωσης μπορεί να οδηγήσει σε αποκλεισμό των χρηστών από μια υπηρεσία και ενδεχομένως ακόμη και σε μια αγωγή (Facebook, Inc. v. Power Ventures, Inc., 2016).

Για τις πλατφόρμες κοινωνικής δικτύωσης, η δικαστική δίωξη όσων επιδίδονται σε scraping μπορεί να έχει οικονομικό νόημα. Πολλές από τις πληροφορίες που συλλέγουν οι πλατφόρμες κοινωνικής δικτύωσης για τους χρήστες τους είναι προς πώληση –όχι άμεσα, αλλά οι εταιρείες και οι διαφημιστές μπορούν να επωφεληθούν από αυτές μέσω διαφημίσεων και μάρκετινγκ. Οι ανταγωνιστές θα μπορούσαν να αντλήσουν πληροφορίες από το Facebook για να δημιουργήσουν μια παρόμοια πλατφόρμα, για παράδειγμα. Όμως, οι αγωγές μπορεί ακούσια να αποθαρρύνουν όχι μόνο όσους συγκεντρώνουν δεδομένα με οικονομικά κίνητρα, αλλά και ακαδημαϊκούς και δημοσιογράφους που θέλουν να συλλέξουν πληροφορίες από πλατφόρμες μέσων κοινωνικής δικτύωσης για ερευνητικούς σκοπούς.

Αυτό σημαίνει ότι οι δημοσιογράφοι ίσως χρειαστεί να γίνουν πιο δημιουργικοί στον τρόπο με τον οποίο κάνουν ρεπορτάζ και αφηγούνται αυτές τις ιστορίες. Οι δημοσιογράφοι ίσως πρέπει να αγοράσουν bot, για να κατανοήσουν καλύτερα τον τρόπο με τον οποίο αυτά δρουν στο Διαδίκτυο, ή οι ρεπόρτερ μπορεί να πρέπει να αγοράσουν διαφημίσεις στο Facebook, για να κατανοήσουν καλύτερα τον τρόπο με τον οποίο αυτό λειτουργεί (Angwin κ.α., 2017).

Όποιο και αν είναι το μέσο, η λειτουργία εντός και εκτός των ορίων που θέτουν οι εταιρείες μέσων κοινωνικής δικτύωσης θα αποτελέσει σημαντική πρόκληση για τους δημοσιογράφους, καθώς αυτοί κινούνται στο διαρκώς μεταβαλλόμενο κυβερνοπεριβάλλον.

Για τι *δεν* είναι καλά τα δεδομένα των μέσων κοινωνικής δικτύωσης

Φαίνεται επιτακτική ανάγκη να κατανοήσουμε καλύτερα το σύμπαν των δεδομένων από τα μέσα κοινωνικής δικτύωσης –και από τη σκοπιά των περιορισμών του.

Κατανοώντας ποιος *χρησιμοποιεί* τα μέσα κοινωνικής δικτύωσης και ποιος *όχι*

Ένα από τα μεγαλύτερα προβλήματα με τα δεδομένα των μέσων κοινωνικής δικτύωσης είναι ότι δεν μπορούμε να υποθέσουμε ότι οι άνθρωποι που παρακολουθούμε στο Twitter ή στο Facebook είναι αντιπροσωπευτικά δείγματα των ευρύτερων πληθυσμών στον φυσικό κόσμο.

Αν και υπάρχει μεγάλος αριθμός ανθρώπων που διαθέτουν λογαριασμό στο Facebook ή στο Twitter, οι δημοσιογράφοι θα πρέπει να είναι επιφυλακτικοί απέναντι στη θέση ότι οι απόψεις που εκφράζονται στο Διαδίκτυο είναι αυτές του γενικού πληθυσμού. Όπως δείχνει μια έρευνα του Pew από το 2018, η χρήση των μέσων κοινωνικής δικτύωσης ποικίλλει από πλατφόρμα σε πλατφόρμα (Smith & Anderson, 2018). Ενώ περισσότερα από τα δύο τρίτα των ενήλικων Αμερικανών που βρίσκονται στο Διαδίκτυο χρησιμοποιούν το YouTube και το Facebook, λιγότερο από το ένα τέταρτο χρησιμοποιεί το Twitter. Αυτό το είδος δεδομένων μπορεί να είναι πολύ πιο αποτελεσματικό για μια συγκροτημένη και συγκεκριμένη ιστορία, είτε πρόκειται για την εξέταση της ρητορικής μίσους που εξαπλώνεται από συγκεκριμένους πολιτικούς στη Μιανμάρ είτε για την εξέταση του είδους της κάλυψης που δημοσιεύεται από τη συνωμοσιολογική έκδοση Infowars σε βάθος χρόνου.

Δεν αντιστοιχεί ο κάθε χρήστης σε ένα πραγματικό ανθρώπινο ον

Επιπλέον, κάθε χρήστης δεν αντιστοιχεί απαραίτητα σε ένα άτομο. Υπάρχουν αυτοματοποιημένοι λογαριασμοί (bot) και λογαριασμοί που είναι μερικώς αυτοματοποιημένοι και μερικώς ελεγχόμενοι από άνθρωπο (cyborg). Και υπάρχουν, επίσης, χρήστες που διαχειρίζονται πολλαπλούς λογαριασμούς.

Και πάλι, το να κατανοήσουμε το γεγονός ότι υπάρχει ένα πλήθος φορέων που χειραγωγούν τη ροή των πληροφοριών για οικονομικό ή πολιτικό όφελος είναι μια σημαντική πτυχή που πρέπει να έχουμε κατά νου, όταν εξετάζουμε τα δεδομένα των μέσων κοινωνικής δικτύωσης σε μαζική μορφή (αν και αυτό το θέμα από μόνο του –η χειραγώγηση των μέσων ενημέρωσης και των πληροφοριών– έχει εξελιχθεί σε μια καθαυτό σημαντική ιστορία, την οποία οι δημοσιογράφοι προσπαθούν να αφηγηθούν με όλο και πιο σύνθετους τρόπους).

Η τυραννία του πιο φωνακλά

Τέλος, είναι σημαντικό να αναγνωρίσουμε ότι δεν αξιολογούνται όλες οι συμπεριφορές. Πολλοί άνθρωποι συχνά επιλέγουν να παραμείνουν σιωπηλοί. Και καθώς οι πιο μετριοπαθείς φωνές καταγράφονται λιγότερο, είναι μόνο οι ακραίες αντιδράσεις που καταγράφονται και ανατροφοδοτούνται μέσω αλγορίθμων που ενισχύουν δυσανάλογα την ήδη υπάρχουσα εξέχουσα θέση των πιο φωνακλάδων.

Αυτό σημαίνει ότι το περιεχόμενο που το Facebook, το Twitter και άλλες πλατφόρμες προβάλλουν αλγοριθμικά στα χρονολόγιά μας συχνά βασίζεται στα like, στα retweet και στα σχόλια όσων επέλεξαν να συμμετάσχουν. Εκείνοι που δεν έχουν μιλήσει, έχουν παραγκωνιστεί δυσανάλογα σε αυτή τη διαδικασία. Ως εκ τούτου, πρέπει να προσέχουμε τόσο τι δεν αξιολογείται όσο και τι αξιολογείται αλλά και τον τρόπο με τον οποίο οι πληροφορίες κατατάσσονται και αναδύονται ως αποτέλεσμα αυτών των αξιολογούμενων και μη αξιολογούμενων δεδομένων.

Παραπομπές

Angwin, J., Varner, M. & Tobin, A. (2017) «Facebook enabled advertisers to reach "Jew haters"», *ProPublica*, 14 Σεπτεμβρίου. Διαθέσιμο σε: https://www.propublica.org/article/facebook-enabled-advertisers-to-reach-jew-haters (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Facebook, Inc. v. Power Ventures, Inc., No. 13-17102 (2016). United States Ninth Circuit, 12 Ιουλίου. Διαθέσιμο σε: https://caselaw.findlaw.com/summary/opinion/us-9th-circuit/2016/07/12/276979.html (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Rudder, C. (2014). *Dataclysm: Who we are* (When we think no one's looking). London: Fourth Estate.

Smith, A., & Anderson, M. (2018) «Social media use 2018: Demographics and statistics», *Pew Research Center*, 1 Μαρτίου. Διαθέσιμο σε: https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018/ (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Tucker, P. (2013) «Has big data made anonymity impossible?», *MIT Technology Review,* 7 Μαΐου. Διαθέσιμο σε: https://www.technologyreview.com/2013/05/07/178542/has-big-data-made-anonymity-impossible/ (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Vo, L. T. (2015) «The umbrella network», *Al Jazeera America*, 3 Ιουνίου. Διαθέσιμο σε: http://projects.aljazeera.com/2015/04/loving-long-distance/hong-kong-umbrella-protest.html (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Vo, L. T. (2016) «The primary source in the age of mechanical multiplication», *Nieman Lab.* Διαθέσιμο σε: https://www.niemanlab.org/2016/12/the-primary-source-in-the-age-of-mechanical-multiplication/ (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Vo, L. T. (2017a) «Twitter bots are trying to influence you. These six charts show you how to spot one», *BuzzFeed* News, 11 Οκτωβρίου. Διαθέσιο σε: https://www.buzzfeednews.com/article/lamvo/twitter-bots-v-human (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Vo, L. T. (2017b) «Here's what we learned from staring at social media data for a year», *BuzzFeed*, 20 Οκτωβρίου. Διαθέσιμο σε: https://www.buzzfeed.com/lamvo/heres-what-we-learned-from-staring-at-social-media-data-for (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Vo, L. T. (2017c) «What we learned from staring at social media data for a year», *Source*, 20 Οκτωβρίου. Διαθέσιμο σε: https://source.opennews.org/articles/what-buzzfeed-news-learned-after-year-mining-data-/ (Προσπελάστηκε: 8 Δεκεμβρίου 2021).

Η συγγραφέας

Η Λαμ Θούι Βο είναι επικεφαλής δημοσιογράφος δεδομένων στο *BuzzFeed News* και συγγραφέας του βιβλίου *Mining Social Media: Finding Stories in Internet Data* (No Starch Press, 2019), ένα Python book.

6.3

Ψηφιακή Εγκληματολογία: Εκ νέου χρήση των ID[174] στα Google Analytics

Richard Rogers

Περίληψη: Αυτό το κεφάλαιο περιγράφει μια τεχνική εντοπισμού δικτύου με βάση τους ιστοτόπους που μοιράζονται τα ίδια αναγνωριστικά Google Analytics ή/και AdSense.

Λέξεις-κλειδιά: ψηφιακές μέθοδοι, ψηφιακή εγκληματολογία, ανώνυμες πηγές, χαρτογράφηση δικτύου, Google Analytics, δημοσιογραφία δεδομένων

Όταν ένας ερευνητής δημοσιογράφος αποκάλυψε τον Ιούλιο του 2015 ένα κρυφό δίκτυο ρωσικών ιστοτόπων που δημοσίευαν παραπληροφόρηση για την Ουκρανία, η αποκάλυψη αυτή όχι μόνο προοιώνισε την κρατικά χρηματοδοτούμενη εκστρατεία επιρροής πριν από τις προεδρικές εκλογές του 2016 στις ΗΠΑ[175], αλλά διέδωσε, επίσης, μια τεχνική ανακάλυψης δικτύων για δημοσιογράφους δεδομένων και κοινωνικούς ερευνητές (Alexander, 2015). Ποιοι ιστότοποι μοιράζονται το ίδιο ID στα Google Analytics (βλ. Εικόνα 6.3.1); Εάν οι ιστότοποι μοιράζονται το ίδιο ID, προκύπτει ότι λειτουργούν από τον ίδιο εγγεγραμμένο χρήστη, είτε πρόκειται για άτομο, είτε για οργανισμό, είτε για όμιλο μέσων ενημέρωσης. Ο δημοσιογράφος, Λόρενς Αλεξάντερ, οδηγήθηκε στο έργο του από την έλλειψη μιας πηγής πίσω από το emaidan.com.ua, έναν ιστότοπο που εμφανιζόταν να παρέχει πληροφορίες σχετικά με τις διαδηλώσεις της Euromaidan το 2013-2014 στην Ουκρανία, οι οποίες τελικά ανέτρεψαν τον φιλορώσο Ουκρανό πρόεδρο υπέρ ενός φιλοδυτικού. Αναζητώντας την πηγή και «παρακινημένος από την ανωνυμία της», ο Αλεξάντερ σκάλισε τον κώδικα της ιστοσελίδας (Alexander, 2015).

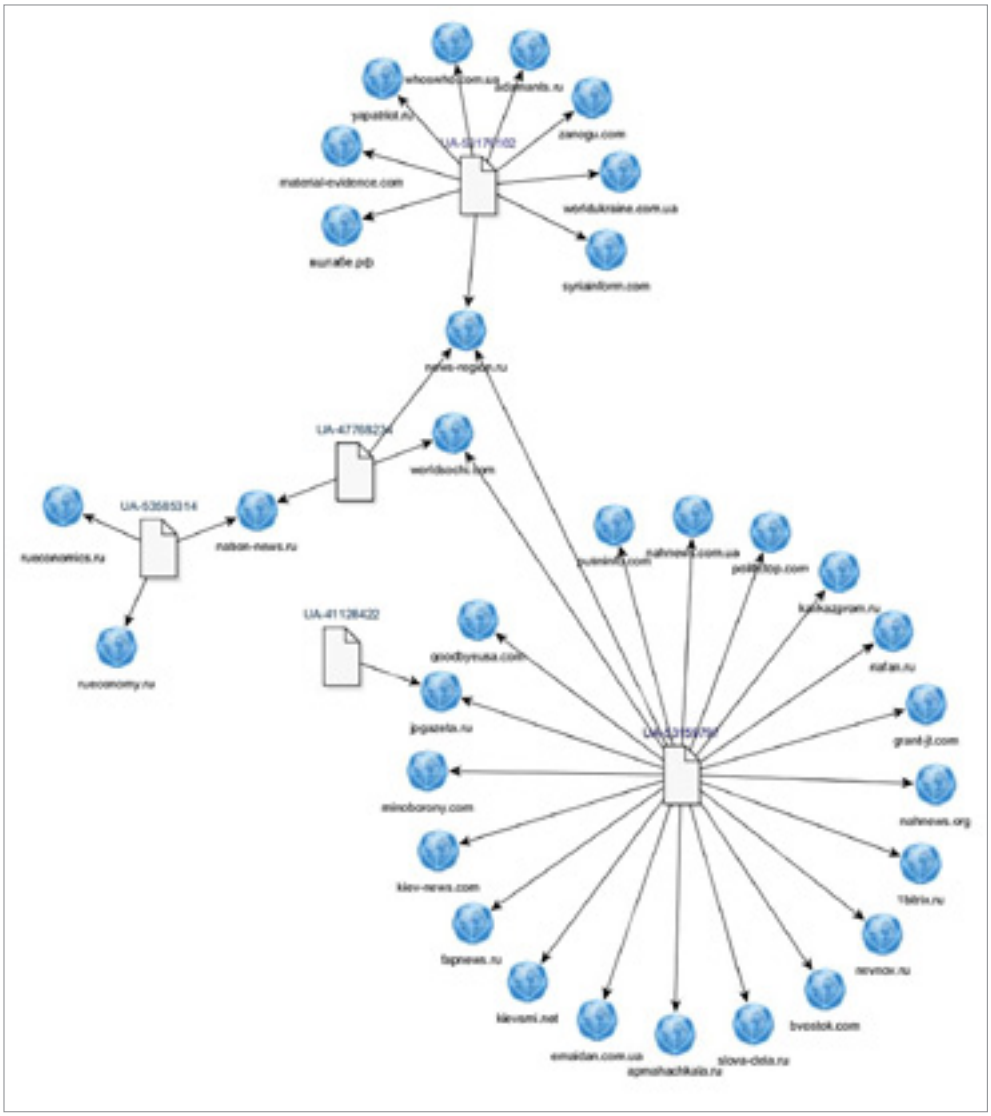
Βλέποντας τον πηγαίο κώδικα της ιστοσελίδας, βρήκε ένα Google Analytics ID, το οποίο εισήγαγε σε ένα λογισμικό αντίστροφης αναζήτησης που παρέχει έναν κατάλογο άλλων ιστοσελίδων οι οποίες χρησιμοποιούν το ίδιο ID[176]. Βρήκε ένα (αστεροειδές) δίκτυο ενός Google Analytics ID που συνδέεται με οκτώ άλλους ιστοτόπους (στην Εικόνα 6.3.1 στην κορυφή του διαγράμματος), οι οποίοι μοιράζονται μια παρόμοια αντι-ουκρανική αφήγηση. Ένας από αυτούς τους ιστοστόπους χρησιμοποιούσε, επίσης, ένα πρόσθετο ID στα Google Analytics, το οποίο οδήγησε σε ένα άλλο σύμπλεγμα συναφών ιστοτόπων (στην εικόνα 6.3.1, κάτω δεξιά),

174. ΣτΜ: Συντομογραφία της λέξης «identifier»: εδώ, σημαίνει το μοναδικό αναγνωριστικό του λογαριασμού Google Analytics που μπορεί να χρησιμοποιείται από μία ή περισσότερες ιστοσελίδες.

175. Μια εκτενέστερη έκδοση αυτού του κεφαλαίου είναι διαθέσιμη στο Rogers, R. (2019). *Doing digital methods*. Λονδίνο: Sage. Ο συγγραφέας επιθυμεί να αναγνωρίσει το έργο αναφοράς του Μίσα Σζπιρτ. Για περισσότερα σχετικά με αυτήν την προσέγγιση, βλ. Rogers, R. (2019). *Doing digital methods*, Λονδίνο: Sage (Κεφάλαιο 11) και Bounegru, L., Gray, J., Venturini, T. & Mauri, M. (2017). *A field guide to "fake news": A collection of recipes for those who love to cook with digital methods*. Public Data Lab (Κεφάλαιο 3).

176. Η αναζήτηση μπορεί, επίσης, να δώσει τη διεύθυνση IP κάθε ιστοτόπου, το αναγνωριστικό Google AdSense, τις εγγραφές τομέα WHOIS και άλλες πληροφορίες.

με παρόμοια πολιτική τοποθέτηση. Εξετάζοντας τα αρχεία WHOIS αρκετών από αυτά τα domain name, βρήκε μια συσχετιζόμενη διεύθυνση ηλεκτρονικού ταχυδρομείου και, στη συνέχεια, το προφίλ και τη φωτογραφία ενός ατόμου στο VKontakte, τον ρωσικό ιστότοπο κοινωνικής δικτύωσης. Το όνομα αυτού του ατόμου βρέθηκε, στη συνέχεια, σε έναν κατάλογο υπαλλήλων που διέρρευσε από την Υπηρεσία Έρευνας Διαδικτύου στην Αγία Πετρούπολη, γνωστή ως ο τόπος εργασίας του χρηματοδοτούμενου από τη ρωσική κυβέρνηση «στρατού των τρολ» (Chen, 2015· Toler, 2015). Συνδέοντας τα στοιχεία των δεδομένων μεταξύ τους, ο Αλεξάντερ έδωσε όνομα και πρόσωπο στο λεγόμενο ρωσικό τρολ. Επίσης, εξανθρώπισε κάπως το τρολ, δείχνοντας τη σελίδα με το χόμπι του στο Pinterest, όπου είναι αναρτημένη μια φωτογραφία με τα ρωσικά διαστημικά επιτεύγματα. Το τρολ είναι, επίσης, θαυμαστής των κοσμοναυτών του διαστήματος.



Εικόνα 6.3.1: Δίκτυο ιστοτόπων που ανακαλύφθηκε μέσω (κοινών) Google Analytics ID. Πηγή: Alexander (2015).

Χρησιμοποιώντας τα λεγόμενα εργαλεία «πληροφοριών ανοιχτού κώδικα» («open-source intelligence tools», OSINT) ως τεχνικές ανακάλυψης (αλλά και ως ψηφιακές μεθόδους, με την έννοια της εκ νέου χρήσης των Google Analytics και του λογισμικού αντίστροφης αναζήτησης), ο Αλεξάντερ και άλλοι δημοσιογράφοι δημιουργούν και ακολουθούν συνδέσμους σε κώδικες, δημόσια αρχεία, βάσεις δεδομένων και διαρροές, συνθέτοντάς τα όλα μαζί για να παρουσιάσουν «ποιος κρύβεται πίσω» από συγκεκριμένες δραστηριότητες (Bazzell, 2016). Η «ανακάλυψη» είναι μια ερευνητική ή ακόμη και ψηφιακή εγκληματολογική προσέγγιση για τη δημοσιογραφική αναζήτηση και αποκάλυψη, όπου κάποιος θα εντοπίσει και, στη συνέχεια, θα προσπαθήσει να έρθει σε επαφή με το άτομο, τον οργανισμό ή την ομάδα μέσων ενημέρωσης, για να του πάρει συνέντευξη και να του δώσει την ευκαιρία να λογοδοτήσει για τη δουλειά του[177]. Η διπλή εξιστόρηση –η ανακάλυψη του δημοσιογράφου και η εξήγηση του ανθρώπου που ανακαλύφθηκε– αποτελεί την ιστορία που πρέπει να ειπωθεί. Ο σκοπός είναι να δοθεί δημοσιότητα, να αποκαλυφθεί μέσα από τον δυσπρόσιτο κώδικα των ιστοσελίδων το συγκεκαλυμμένο πολιτικό έργο που εκπονείται και να αναγνωριστεί αυτή η συγκεκριμένη απόδειξη (Latour, 2005).

Εικόνα 6.3.2: Ενσωματωμένα ψηφιακά αντικείμενα σε ιστοτόπους, που απεικονίζονται ως διάγραμμα δικτύου. Πηγή: Alexander (2015).

Η ερευνητική δουλειά πάνω στα Google Analytics ID έχει παράδοση στην πρακτική της αποκάλυψης ανώνυμων διαδικτυακών φορέων μέσω εκμετάλλευσης ή σημείων εισόδου σε προσωπικά αναγνωρίσιμα δεδομένα που δεν έχουν προβλεφθεί από τους δημιουργούς του. Η εξόρυξη Google Analytics ID για την ανακάλυψη και τη χαρτογράφηση δικτύων είναι, επίσης, μια άσκηση εκ νέου χρήσης, που χρησιμοποιεί το λογισμικό με μη προβλεπόμενο τρόπο για κοινωνική έρευνα. Ο εμπνευστής της τεχνικής αυτής, ο Άντι Μπάιο, δημοσιογράφος του περιοδικού *Wired*, αφηγείται την ιστορία ενός ανώνυμου μπλόγκερ που αναρτούσε εξαιρετικά προσβλητικό υλικό και ο οποίος είχε καλύψει τα ίχνη του με τους «συνήθεις τρόπους»: «κρύβοντας προσωπικές πληροφορίες στις οποίες το domain name θα μπορούσε να οδηγήσει,

177. Η ψηφιακή εγκληματολογία έχει τις ρίζες της στη διερεύνηση της εταιρικής απάτης μέσω τεχνικών όπως το «data carving» («χάραξη δεδομένων»), οι οποίες επιτρέπουν την ανάκτηση διαγραμμένων αρχείων.

χρησιμοποιώντας μια διαφορετική διεύθυνση IP από τους άλλους ιστοτόπους του και διαγράφοντας τυχόν κοινόχρηστους πόρους από την εγκατάσταση του WordPress» (Baio, 2011). Ο Μπάιο τον ταυτοποίησε, επειδή ο μπλόγκερ μοιραζόταν ένα Google Analytics ID με άλλους ιστοτόπους που διαχειριζόταν σε κοινή θέα. Η διδακτική ιστορία σχετικά με αυτήν την ανακάλυψη και τεχνική αποκάλυψης καταλήγει με τον Μπάιο να προσφέρει έναν οδηγό ασφαλείας για άλλους, δικαίως, ανώνυμους μπλόγκερ, όπως αυτοί που παρακολουθούν τα μεξικανικά καρτέλ ναρκωτικών, των οποίων η αποκάλυψη θα μπορούσε να οδηγήσει σε κίνδυνο, ή ακόμη και σε απώλεια, ζωής. Εδώ, θα μπορούσε κανείς να δοκιμάσει, επίσης, την ανθεκτικότητα της ανωνυμίας και να ενημερώσει τους δημοσιογράφους που εργάζονται μυστικά στο Διαδίκτυο για τυχόν τρωτά σημεία ή πιθανές αδυναμίες.

Εν κατακλείδι, προσφέρω ένα ερευνητικό πρωτόκολλο για την ανακάλυψη δικτύων με χρήση των Google Analytics ID, το οποίο συνοψίζεται στην παρακάτω λίστα:

- Συγκεντρώστε έναν κατάλογο ιστοτόπων που δεν παραθέτουν τις πηγές τους. Εντοπίστε τα ID τους στο Google Analytics και το AdSense.
- Εισάγετε τη λίστα URL σε ένα λογισμικό αντίστροφης αναζήτησης, όπως το dnslytics.com. Αναζητήστε ιστοτόπους που μοιράζονται τα ίδια ID.
- Ομαδοποιήστε και προσδιορίστε θεματικά τους ιστοτόπους που μοιράζονται τα ID. Μελετήστε την οπτικοποίηση του δικτύου με τη χρήση του Gephi.

Παραπομπές

Alexander, L. (2015) «Open-source information reveals pro-Kremlin web campaign», *Global Voices*, 13 Ιουλίου. Διαθέσιμο σε: https://globalvoices.org/2015/07/13/open-source-information-reveals-pro-kremlin-web-campaign/ (Προσπελάστηκε: 10 Δεκεμβρίου 2021).

Baio, A. (2011) «Think you can hide, anonymous blogger? Two words: Google analytics», *Wired*, 15 Νοεμβρίου. Διαθέσιμο σε: https://www.wired.com/2011/11/goog-analytics-anony-bloggers/ (Προσπελάστηκε: 10 Δεκεμβρίου 2021).

Bazzell, M. (2016). *Open source intelligence techniques: Resources for searching and analyzing online information* (5th ed.). CreateSpace Independent Publishing Platform.

Bounegru, L., Gray, J., Venturini, T. & Mauri, M. (2017). *A field guide to "fake news": A collection of recipes for those who love to cook with digital methods*. Public Data Lab. Διαθέσιμο σε: https://www.researchgate.net/profile/Liliana-Bounegru/publication/320237117_A_Field_Guide_to_Fake_News_A_Collection_of_Recipes_for_Those_Who_Love_to_Cook_with_Digital_Methods/links/59d666f4aca27213df9e7b2d/A-Field-Guide-to-Fake-News-A-Collection-of-Recipes-for-Those-Who-Love-to-Cook-with-Digital-Methods.pdf (Προσπελάστηκε: 10 Δεκεμβρίου 2021).

Chen, A. (2015) «The agency», *The New York Times Magazine*, 2 Ιουνίου. Διαθέσιμο σε: https://www.nytimes.com/2015/06/07/magazine/the-agency.html (Προσπελάστηκε: 10 Δεκεμβρίου 2021).

Latour, B. (2005) «From realpolitik to dingpolitik—An introduction to making things public», στο B. Latour & P. Weibel (επιμ.), *Making things public: Atmospheres of democracy,* Κέιμπριτζ, Μασαχουσέτη: MIT Press, σσ. 14-41. Διαθέσιμο σε: http://www.bruno-latour.fr/node/208.html (Προσπελάστηκε: 10 Δεκεμβρίου 2021).

Rogers, R. (2019). *Doing digital methods*. Λονδίνο: Sage.

Toler, A. (2015) «Inside the Kremlin troll army machine: Templates, guidelines, and paid posts»,

Global Voices, 14 Μαρτίου. Διαθέσιμο σε: https://globalvoices.org/2015/03/14/russia-kremlin-troll-army-examples/ (Προσπελάστηκε: 10 Δεκεμβρίου 2021).

Ο συγγραφέας

Ο Ρίτσαρντ Ρότζερς είναι Καθηγητής Νέων Μέσων και Ψηφιακής Κουλτούρας στο Τμήμα Σπουδών Μέσων στο Πανεπιστήμιο του Άμστερνταμ και διευθυντής της Πρωτοβουλίας Ψηφιακών Μεθόδων (Digital Methods Initiative), καθώς και της Ολλανδικής Σχολής Έρευνας για τις Σπουδές Μέσων (Netherlands Research School for Media Studies).

6.4
Εφαρμογές και οι Δυνατότητές τους για Έρευνες Δεδομένων

Esther Weltevrede

Περίληψη: Διερεύνηση των σχέσεων εφαρμογής-πλατφόρμας για έρευνες δεδομένων.

Λέξεις-κλειδιά: εφαρμογές, πλατφόρμες μέσων κοινωνικής δικτύωσης, ψηφιακές μέθοδοι, υποδομές δεδομένων, δημοσιογραφία δεδομένων, έρευνες δεδομένων

Πρόσφατα, το Netvizz, ένα εργαλείο για την εξαγωγή δεδομένων από το Facebook, έχασε την πρόσβαση στη λειτουργία «Πρόσβασης στο Δημόσιο Περιεχόμενο Σελίδων» του Facebook («Facebook Page Public Content Access). Αυτό φαίνεται πως τερμάτισε την επισφαλή σχέση που διατηρούσε ο δημιουργός του, ο ερευνητής ψηφιακών μεθόδων Μπέρνχαρντ Ρίντερ, με το API[178] του Facebook τα τελευταία εννέα χρόνια[179]. Το τέλος του Netvizz είναι χαρακτηριστικό μιας ευρύτερης αλλαγής στην ψηφιακή έρευνα, όπου οι πλατφόρμες περιορίζουν περαιτέρω τη συλλογή δεδομένων μέσω API και τις πολιτικές τους για προγραμματιστές. Παρόλο που η πραγματική αποτελεσματικότητα των μεθόδων της Cambridge Analytica αμφισβητείται (Lomas, 2018· Smout & Busvine, 2018), το σκάνδαλο προκάλεσε συζήτηση για την προστασία της ιδιωτικής ζωής και των δεδομένων στα μέσα κοινωνικής δικτύωσης και, με τη σειρά του, το Facebook απάντησε περιορίζοντας περαιτέρω την πρόσβαση σε δεδομένα από τις πλατφόρμες του.

Από την αρχική ανακοίνωση τον Μάρτιο του 2018[180], η κλιμακωτή εφαρμογή των περιορισμών πρόσβασης στα δεδομένα από το Facebook, στο πλαίσιο της ευρύτερης οικογένειας εφαρμογών του, έχει καταστήσει ορατό το τεράστιο δίκτυο ενδιαφερόμενων τρίτων μερών που βασίζονται στην πλατφόρμα για ποικίλους σκοπούς. Οι εφαρμογές έπαψαν να λειτουργούν, οι διαφημιστικοί στόχοι περιορίστηκαν, αλλά το πιο σοβαρό πλήγμα φαίνεται να το έχουν υποστεί οι ψηφιακοί ερευνητές. Αυτό οφείλεται στο γεγονός ότι οι εφαρμογές που έχουν ως πρωταρχικό σκοπό τη συλλογή δεδομένων δεν επιτρέπονται πλέον. Οι ψηφιακοί ερευνητές αντιτάχθηκαν στις αλλαγές (Bruns, 2018), υποστηρίζοντας ότι αυτές θα επιβάρυναν το κόστος της έρευνας που διεξάγεται προς όφελος του δημόσιου συμφέροντος. Ο κατάλογος των παραπομπών στο άρθρο του Netvizz (Rieder, 2013) περιλαμβάνει περισσότερες από 450 δημοσιεύσεις, οι οποίες στην πραγματικότητα υπερβαίνουν κατά πολύ αυτόν τον αριθμό –σκεφτείτε μόνο τις πολλές ερευνητικές εργασίες φοιτητών που χρησιμοποιούν το εργαλείο. Ομοίως, μια ad hoc καταγραφή από τον Μπέχμαν των μελετών που «δεν θα μπορούσαν να έχουν υπάρξει χωρίς πρόσβαση σε δεδομένα API»[181] περιλαμβάνει έναν εντυπωσιακό κατάλογο δημοσιογραφικών, κοινωνικών επιστημονικών και άλλων ψηφιακών ερευνητικών δημοσιεύσεων. Αναλογιζόμενοι τον αντίκτυπο που έχουν στην ψηφιακή έρευνα οι περιορισμοί πρόσβασης στα δεδομένα, οι

178. ΣτΜ: «Application Programming Interface» (διεπαφή προγραμματισμού εφαρμογών).
179. http://thepoliticsofsystems.net/?s=netvizz
180. https://about.fb.com/news/2018/03/cracking-down-on-platform-abuse/
181. https://docs.google.com/document/d/15YKeZFSUc1j03b4lW9YXxGmhYEnFx3TSy68qCrX9BEI/edit

συγγραφείς έχουν εντάξει αυτές τις εξελίξεις σε ένα πλαίσιο και έχουν χαρακτηρίσει την τελευταία δεκαετία ως «έρευνα που βασίζεται σε API» (Venturini & Rogers, 2019) ή «έρευνα που σχετίζεται με API» (Perriam κ.α., 2020). Ως τέτοιου είδους έρευνες ορίζονται οι προσεγγίσεις της ψηφιακής έρευνας οι οποίες βασίζονται στην εξαγωγή δεδομένων που διατίθενται από διαδικτυακές πλατφόρμες μέσω των API τους. Σίγουρα, τα API –με τα δεδομένα τους ήδη έτοιμα για κοινωνική έρευνα– έχουν διευκολύνει ιδιαίτερα την έρευνα με δεδομένα από τα μέσα κοινωνικής δικτύωσης, χώρια που επέτρεψαν σε μια ολόκληρη γενιά φοιτητών να πειραματιστούν με την ψηφιακή έρευνα. Ο λόγος είναι ότι δεν απαιτούν τεχνικές δεξιότητες και, για τα πρότυπα διαδικτυακών δεδομένων, τα δεδομένα είναι σχετικά καθαρά. Η έρευνα που βασίζεται σε API έχει, επίσης, δεχθεί κριτική από την αρχή, κυρίως λόγω των ερευνητικών πλεονεκτημάτων των API που καθοδηγούνται από την ευκολία, επηρεάζοντας την ικανότητα των ερευνητών να αναπτύσσουν σχετικά ερευνητικά ερωτήματα (Marres, 2017).

Το κεφάλαιο αυτό ακολουθεί τις πρόσφατες προτροπές των Βεντουρίνι και Ρότζερς (Venturini & Rogers, 2019) και της Πρωτοβουλίας Ψηφιακών Μεθόδων («Digital Methods Initiative») για «έρευνα μετά τα API», και επικεντρώνεται στις ευκαιρίες που προκύπτουν ως απάντηση στις πρόσφατες εξελίξεις στα οικοσυστήματα των μέσων κοινωνικής δικτύωσης[182]. Η ψηφιακή έρευνα, με την έννοια που χρησιμοποιείται σε αυτό το κεφάλαιο, ορίζεται από τη μεθοδολογική αρχή της «παρακολούθησης του μέσου», της ανταπόκρισης και της διασύνδεσης των μεθόδων με τις εξελίξεις στο ψηφιακό περιβάλλον. Στη συνέχεια, προσεγγίζω τους πρόσφατους περιορισμούς των API, υποστηρίζοντας την ανανεωμένη ανάγκη και τη δυνατότητα δημιουργικής και εφευρετικής εξερεύνησης διαφορετικών τύπων κοινωνικοτεχνικών δεδομένων που αποτελούν το κλειδί για τη διαμόρφωση των σημερινών περιβαλλόντων των πλατφορμών. Συνεχίζω λαμβάνοντας υπόψη τις ευκαιρίες που έχουν επισημανθεί από τους ψηφιακούς ερευνητές και προσθέτοντας σε αυτές μια μεθοδολογική προοπτική για τη μελέτη των σχέσεων μεταξύ εφαρμογών και πλατφορμών. Με αυτόν τον τρόπο, ελπίζω να προσφέρω στους δημοσιογράφους δεδομένων που ενδιαφέρονται για τις δυνατότητες αφήγησης τις οποίες προσφέρουν τα δεδομένα από τα μέσα κοινωνικής δικτύωσης (βλ., π.χ., το κεφάλαιο της Λαμ Θούι Βο[183] σε αυτόν τον τόμο), κάποια σημεία εκκίνησης για την προσέγγιση των ερευνών *με* και *για* τις πλατφόρμες, και τα δεδομένα τους στην τρέχουσα μετα-API εποχή.

Το κοινό χαρακτηριστικό των ψηφιακών μεθόδων είναι ότι χρησιμοποιούν μια σειρά από τεχνικές συλλογής και ανάλυσης δεδομένων που βελτιστοποιούν τη χρήση εγγενών ψηφιακών μορφών δεδομένων. Αυτά προκύπτουν με την εισαγωγή των ψηφιακών μέσων στην κοινωνική ζωή. Οι ερευνητές ψηφιακών μεθόδων αναπτύσσουν εργαλεία εμπνευσμένα από τα ψηφιακά μέσα, ώστε να είναι σε θέση να χειρίζονται αυτές τις μορφές δεδομένων με μεθοδολογικά καινοτόμους τρόπους. Επομένως, η ιστορία των ψηφιακών μεθόδων μπορεί, επίσης, να διαβαστεί ως αφήγηση μιας ιστορίας των βασικών φορμάτ και δομών των δεδομένων του Διαδικτύου, που προσαρμόζονται στις αλλαγές των μέσων και τις περιλαμβάνουν στην ανάλυση.

Παρακάτω, θα ήθελα να συνεισφέρω στις μετα-API ερευνητικές προσεγγίσεις, προτείνοντας μια προοπτική για τη μελέτη των πλατφορμών ως υποδομών δεδομένων από την οπτική γωνία της εφαρμογής-πλατφόρμας. Ο αντίκτυπος που έχουν οι περιορισμοί πρόσβασης στα δεδομένα σε ευρύτερα οικοσυστήματα των μέσων ενημέρωσης αποδεικνύει ότι απαιτείται επειγόντως προηγμένη, εξειδικευμένη γνώση των υποδομών των πλατφορμών και της αλληλεπίδρασής τους με τις εφαρμογές τρίτων και καταδεικνύει την ανάγκη για διευρυμένο αλφαβητισμό στις υποδομές δεδομένων (Gray κ.α., 2018), πέραν της γνώσης σχετικά με τον τρόπο λειτουργίας τρίτων εταιρειών και εφαρμογών σε περιβάλλοντα μέσων κοινωνικής δικτύωσης.

182. https://wiki.digitalmethods.net/Dmi/WinterSchool2020

183. ΣτΕ: Κεφάλαιο 6.2, «Αφήγηση Ιστοριών με τα Μέσα Κοινωνικής Δικτύωσης»

Εφαρμογές και πλατφόρμες ως υποδομή

Οι περιορισμοί στα δεδομένα των πλατφορμών αποτελούν αναπόσπαστο μέρος της εξέλιξης των μέσων κοινωνικής δικτύωσης σε πλατφόρμες ως υποδομές. Οι εξελίξεις αυτές αναδεικνύουν την εξέλιξη της εστίασης των ψηφιακών οικοσυστημάτων στις εταιρικές συμπράξεις (Helmond κ.α., 2019). Μετά από έναν χρόνο αρνητικής κάλυψης σε συνέχεια του ρόλου που έπαιξε το Facebook στις εκλογές, ο Ζούκερμπεργκ δημοσίευσε ένα σημείωμα στο οποίο σκιαγραφούσε την αλλαγή οπτικής της πλατφόρμας από τη «σύνδεση ανθρώπων» στη δημιουργία μιας «κοινωνικής υποδομής» (Helmond κ.α., 2019· Hoffmann κ.α., 2018)[184]. Η έννοια της κοινωνικής υποδομής αναδεικνύει, αφενός, τις κοινωνικές δραστηριότητες ως το βασικό προϊόν της πλατφόρμας που συνδέει και δημιουργεί αξία για τις πολλαπλές πλευρές της αγοράς και, αφετέρου, τη μετατόπιση της εταιρείας από μέσο κοινωνικής δικτύωσης σε υποδομή δεδομένων, επεκτείνοντας την πλατφόρμα, ώστε να συμπεριλάβει άλλους ιστοτόπους και την ευρύτερη οικογένεια 70 εφαρμογών (Nieborg & Helmond, 2019)[185]. Αυτή η στροφή στην υποδομή σηματοδοτεί ένα επόμενο βήμα στην ικανότητα της πλατφόρμας να επεκτείνει την υποδομή δεδομένων της σε εφαρμογές, πλατφόρμες και ιστοτόπους τρίτων, καθώς και να διευκολύνει τις εσωτερικές ενσωματώσεις.

Παρόλο που οι πλατφόρμες ως υποδομή λαμβάνουν όλο και μεγαλύτερη προσοχή (Plantin κ.α., 2018), όπως και οι μεμονωμένες εφαρμογές, ο τρόπος με τον οποίο οι εφαρμογές λειτουργούν σε υποδομές δεδομένων, και μεταξύ αυτών, δεν έχει μελετηθεί επαρκώς και συχνά δεν λαμβάνεται υπόψη. Ωστόσο, οι εφαρμογές μεταμορφώνουν και αξιοποιούν συνεχώς τις καθημερινές πρακτικές μέσα σε περιβάλλοντα πλατφόρμας. Χρησιμοποιώ έναν σχεσιακό ορισμό των εφαρμογών, εστιάζοντας στις εφαρμογές τρίτων, οι οποίες ορίζονται ως εφαρμογές που έχουν κατασκευαστεί σε μια πλατφόρμα από εξωτερικούς προγραμματιστές, χωρίς να ανήκουν στην πλατφόρμα ή να λειτουργούν μέσω αυτής. Όταν μια εφαρμογή συνδέεται σε μια πλατφόρμα, παρέχεται πρόσβαση σε λειτουργίες και δεδομένα της πλατφόρμας, ανάλογα με την παροχή συγκατάθεσης. Οι εφαρμογές δίνουν, επίσης, τη δυνατότητα στους ενδιαφερόμενους –για παράδειγμα, σε app store, διαφημιστές ή χρήστες– να τις ενσωματώσουν και να τις αξιοποιήσουν με πολλαπλούς, ταυτόχρονους τρόπους. Με άλλα λόγια, οι εφαρμογές έχουν ενσωματωμένη την τάση να σχετίζονται και να συσχετίζονται με διαφορετικές υποδομές δεδομένων λειτουργίας. Αυτή η συγκεκριμένη τοποθέτηση των εφαρμογών τρίτων τις καθιστά ιδιαίτερα κατάλληλες για έρευνες στα περιβάλλοντα της πλατφόρμας ως υποδομής.

Οι πλατφόρμες των μέσων κοινωνικής δικτύωσης θέτουν μεθοδολογικές προκλήσεις, διότι, όπως αναφέρθηκε, η πρόσβαση στα δεδομένα που δημιουργούνται από τους χρήστες περιορίζεται ολοένα και περισσότερο, γεγονός που προκαλεί τους ερευνητές να εξετάσουν εκ νέου τι είναι τα «κοινωνικά δεδομένα» («social data») και να δημιουργήσουν εναλλακτικές προοπτικές. Σε αντίθεση με τον τρόπο με τον οποίο οι πλατφόρμες κοινωνικής δικτύωσης προσφέρουν πρόσβαση σε δεδομένα που δημιουργούνται από χρήστες για ψηφιακή έρευνα, δομημένα μέσω API, οι πηγές δεδομένων εφαρμογών χαρακτηρίζονται όλο και περισσότερο από τον κλειστό ή ιδιόκτητο χαρακτήρα τους. Παρόλο που η συσκότιση είναι μια ευρέως χρησιμοποιούμενη τεχνική στη μηχανική λογισμικού (Matviyenko κ.α., 2015), οι ενέργειες που καθιστούν τον κώδικα και τα δεδομένα δυσανάγνωστα ή απρόσιτα έχουν σημαντικό αντίκτυπο στην ψηφιακή έρευνα. Αυτές οι αυξημένες προκλήσεις που θέτουν οι πλατφόρμες και τα περιβάλλοντα εφαρμογών για την παράκαμψη ή την εκτροπή της εμπειρικής έρευνας είναι αυτό που οι συνάδελφοι και εγώ έχουμε ονομάσει «αντίσταση υποδομής» («infrastructural resistance», Dieter κ.α., 2019). Αντίθετα, τα δεδομένα που είναι διαθέσιμα για την ψηφιακή έρευνα σήμερα χαρακτηρίζονται από ετερογενή φορμάτ δεδομένων τα οποία ποικίλλουν, από δεδομένα που βασίζονται σε συσκευές (π.χ.

184. https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10154544292806634/
185. https://www.appannie.com/en/

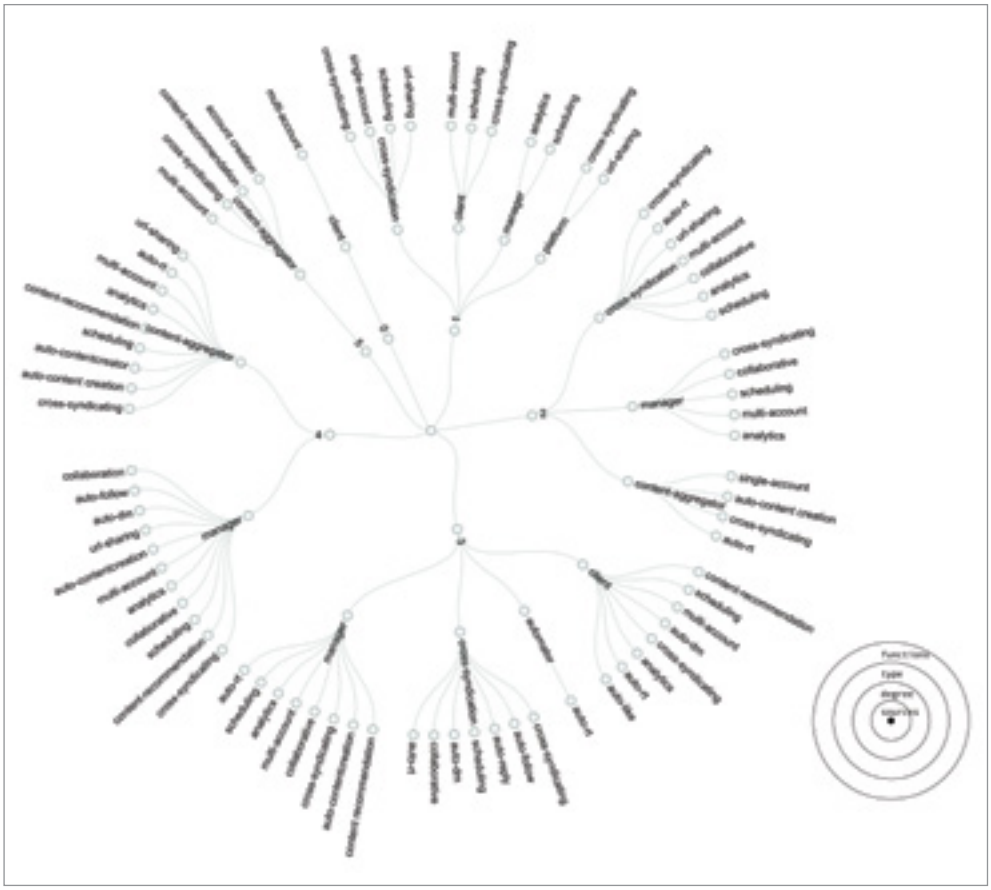
GPS) ως βιβλιοθήκες λογισμικού (π.χ. κιτ ανάπτυξης λογισμικού, SDK) και συνδέσεις δικτύου (π.χ. δίκτυα διαφημίσεων). Οι εφαρμογές μπορούν να συλλέγουν δεδομένα που δημιουργούνται από τους χρήστες, αλλά ως επί το πλείστον δεν προσφέρουν πρόσβαση μέσω ανοικτών API. Επομένως, δεν υπάρχουν έτοιμα δεδομένα για έρευνες.

Ακολούθως, παρουσιάζω τρεις διαφορετικές προσεγγίσεις έρευνας δεδομένων από τα κάτω, μέσω των οποίων οι ψηφιακοί ερευνητές και οι δημοσιογράφοι μπορούν να αξιοποιήσουν ενεργά διαφορετικές «ερευνητικές δυνατότητες» (Weltevrede, 2016) και να τις χρησιμοποιήσουν για να προωθήσουν ή να ξεκινήσουν μια έρευνα. Οι ερευνητικές δυνατότητες συντονίζονται με τις δυνατότητες δράσης εντός του λογισμικού από την οπτική γωνία και σε συμφωνία με τα ενδιαφέροντα του ερευνητή. Η προσέγγιση αυτή επιτρέπει την ανάπτυξη εφευρετικών ψηφιακών μεθόδων (Lury & Wakeford, 2012· Rogers, 2013). Αυτά απαιτούν την επανεξέταση των τεχνικών μορφών και τύπων των σχέσεων μεταξύ εφαρμογών και πλατφορμών, με την εξερεύνηση των δυνατοτήτων τους για ανάλυση. Οι μελέτες βασίζονται σε πρόσφατες έρευνες που πραγματοποιήσαμε εγώ και οι συνάδελφοί μου, αντλώντας έμπνευση αλλά και σημειώνοντας προκλήσεις και κάνοντας προτάσεις για το είδος των ερευνών που επιτρέπουν αυτές οι πηγές δεδομένων, για να αυξήσουμε την κατανόησή μας για το περιβάλλον της πλατφόρμας ως υποδομής.

Ψεύτικες κοινωνικές υποδομές

Η πρώτη έρευνα εξετάζει τους ψεύτικους ακολούθους και τη σχέση τους με την κοινωνική υποδομή του Facebook. Όλο και μεγαλύτερη προσοχή δίνεται στην έκταση των ψεύτικων ακολούθων στα περιβάλλοντα των μέσων κοινωνικής δικτύωσης, τόσο από τις πλατφόρμες όσο και από την ψηφιακή έρευνα. Από τη σκοπιά των πλατφορμών, η αγορά των ψεύτικων ακολούθων συχνά αποκλείεται από τις συζητήσεις για τις πλατφόρμες ως πολύπλευρες αγορές. Η αγορά των ψεύτικων ακολουθών δεν θεωρείται «πλευρά» και σίγουρα δεν αποτελεί μέρος της «οικογένειας». Οι ψεύτικοι ακόλουθοι δημιουργούν μια ανεπίσημη υποδομή σχέσεων, οι οποίες αναγνωρίζονται από τις πλατφόρμες ως ανεπιθύμητες καταχρήσεις. Οι πλατφόρμες δεν τις επιδιώκουν, αλλά εκείνες λειτουργούν σε συνδυασμό με τους μηχανισμούς των πλατφορμών και χάρη σε αυτούς. Επιπλέον, οι πρακτικές αυτές μειώνουν την αξία του βασικού προϊόντος, δηλαδή της κοινωνικής δραστηριότητας.

Συνάδελφοι και εγώ διερευνήσαμε την πορεία προς το δημοψήφισμα για το Brexit στο Twitter, εστιάζοντας στις πιο συχνά χρησιμοποιούμενες εφαρμογές σε αυτό το σύνολο δεδομένων (Gerlitz & Weltevrede, 2019) –βλ. Εικόνα 6.4.1. Η συστηματική ανάλυση αυτών των εφαρμογών και των λειτουργιών τους παρέχει πληροφορίες σχετικά με τους μηχανισμούς αυτοματοποιημένων και ψεύτικων αλληλεπιδράσεων στο πλαίσιο της δομής διακυβέρνησης της πλατφόρμας. Σε ένα έργο εν εξελίξει με τον Γιόχαν Λίντκβιστ, διερευνούμε ένα σύνολο από πάνω από 1.200 πλατφόρμες μεταπώλησης που επιτρέπουν την αγορά και την πώληση ψεύτικων αλληλεπιδράσεων σε ένα ευρύ φάσμα πλατφορμών. Αυτές οι αρχικές έρευνες δείχνουν πώς οι ψεύτικοι ακόλουθοι σχετίζονται τεχνικά με πλατφόρμες –τόσο με επίσημες εφαρμογές τρίτων που συνδέονται μέσω API, όσο και με μια υποδομή πλατφορμών που επίσης συνδέονται ανεπίσημα με πλατφόρμες μέσων κοινωνικής δικτύωσης. Αυτό που έδειξε η εν λόγω αρχική διερεύνηση είναι ότι η έρευνα θα πρέπει να φιλοξενήσει ποικιλία δεδομένων αυτοματοποιημένης και ψεύτικης προέλευσης. Οι αυτοματοποιημένοι και ψεύτικοι λογαριασμοί δεν μπορούν να αντιμετωπίζονται (μόνο) ως τύπος ή ως φορέας, αλλά ως πρακτική, η οποία τοποθετείται και αναδύεται σε σχέση με τις δυνατότητες του μέσου. Όπως φαίνεται στην περίπτωση του Twitter, ένας λογαριασμός δεν αντιπροσωπεύει απαραίτητα έναν άνθρωπο-χρήστη, καθώς επιτυγχάνεται με κατανεμημένους και εγκαθιδρυμένους τρόπους, όπως ακριβώς ένα tweet δεν είναι ένα tweet, που συνήθως νοείται ως μια μοναδικά δακτυλογραφημένη ανάρτηση (Gerlitz & Weltevrede, 2019).



Εικόνα 6.4.1: Λειτουργίες αυτοματισμού. Το δενδρόγραμμα απεικονίζει την ιεραρχία των πηγών, τους βαθμούς αυτοματοποίησης, τους τύπους των πηγών και τις λειτουργίες τους στο σύνολο δεδομένων σχετικά με το Brexit, 17-23 Ιουνίου 2016. Πηγή: Gerlitz & Weltevrede (2019).

Σχέσεις εφαρμογών - πλατφόρμας

Η δεύτερη έρευνα εξετάζει τα app store ως υποδομές δεδομένων για εφαρμογές. Σήμερα, το κύριο σημείο εισόδου στις εφαρμογές –τόσο για τους προγραμματιστές όσο και για τους χρήστες– είναι μέσω των app store, όπου οι χρήστες μπορούν να αναζητήσουν μεμονωμένες εφαρμογές ή να ταξινομήσουν συλλογές ή είδη εφαρμογών. Αξιοποιώντας μεθόδους από μελέτες αλγορίθμων (Rogers, 2013· Sandvig κ.α., 2014), μπορεί κανείς να ασχοληθεί με την τεχνικότητα των «πολιτισμών κατάταξης» (Rieder κ.α., 2018) –για παράδειγμα, στο Google Play και στο App Store. Ένα τέτοιο εγχείρημα αφορά τόσο την αλγοριθμική και οικονομική ισχύ όσο και τις κοινωνικές τους συνέπειες. Μπορεί να χρησιμοποιηθεί για την απόκτηση γνώσεων σχετικά με τους μηχανισμούς κατάταξής τους και την κατανόηση των λόγων για τους οποίους αυτό έχει σημασία για την κυκλοφορία του πολιτισμικού περιεχομένου.

Τα app store μπορούν, επίσης, να χρησιμοποιηθούν για την ταξινόμηση συλλογών ή ειδών εφαρμογών για τη μελέτη των σχέσεων εφαρμογών-πλατφόρμας από την πλευρά των εφαρμογών. Στο *Regramming the Platform* (Gerlitz κ.α., 2019), οι συνάδελφοί μου και εγώ ερευνήσαμε πάνω από 18.500 εφαρμογές και τους διαφορετικούς τρόπους με τους οποίους οι εφαρμογές σχετίζονται με τα χαρακτηριστικά και τις λειτουργίες της πλατφόρμας. Ένα από τα βασικά

ευρήματα της παρούσας μελέτης είναι ότι οι προγραμματιστές εφαρμογών βρίσκουν δημιουργικές λύσεις, για να παρακάμψουν τα επίσημα API των πλατφορμών, παρακάμπτοντας έτσι και τα επίσημα συστήματα διακυβέρνησης των πλατφορμών (Πίνακας 6.4.1). Η «εφαρμογοκεντρική» προσέγγιση των πλατφορμών ως υποδομών παρέχει πληροφορίες σχετικά με τις εφαρμογές τρίτων που αναπτύσσονται περιφερειακά των πλατφορμών των μέσων κοινωνικής δικτύωσης, τις πρακτικές και τα χαρακτηριστικά που υποστηρίζονται και αναπτύσσονται από τις εν λόγω εφαρμογές, καθώς και τις ακατάστατες και συγκυριακές σχέσεις μεταξύ των εφαρμογών και των πλατφορμών των μέσων κοινωνικής δικτύωσης (Gerlitz κ.α., 2019).

Σχέση	**Facebook**	**Instagram**	**Snapchat**	**Twitter**
Εμπορικό σήμα (mention)	1.449 (34,96%)	2.945 (80,03%)	614 (12,17%)	1.107 (21,80%)
Νομικά (mention)	302 (7,29%)	318 (8,65%)	268 (5,31%)	305 (6,01%)
Τεχνικά (mention)	61 (1,47%)	62 (1,68%)	70 (1,39%)	114 (2,24%)
Τεχνικά (βιβλιοθήκες, SDK)*	83 (33,20%)	0 (0%)	0 (0%)	40 (16,13%)
Τεχνικά (αιτήματα HTTP)	156 (62.40%)	89 (35.60%)	12 (4.80%)	102 (41.13%)

**Μόνο για αποτελέσματα αναζήτησης Google Play (N = 998)*
Σημείωση: Οι κατηγορίες Τεχνικά (βιβλιοθήκες, SDK) και Τεχνικά (αιτήματα HTTP) υποδεικνύουν εφαρμογές που έχουν μια «επίσημη» σχέση με το μέσο κοινωνικής δικτύωσης μέσω των API τους. Πηγή: Gerlitz κ.α. (2019).

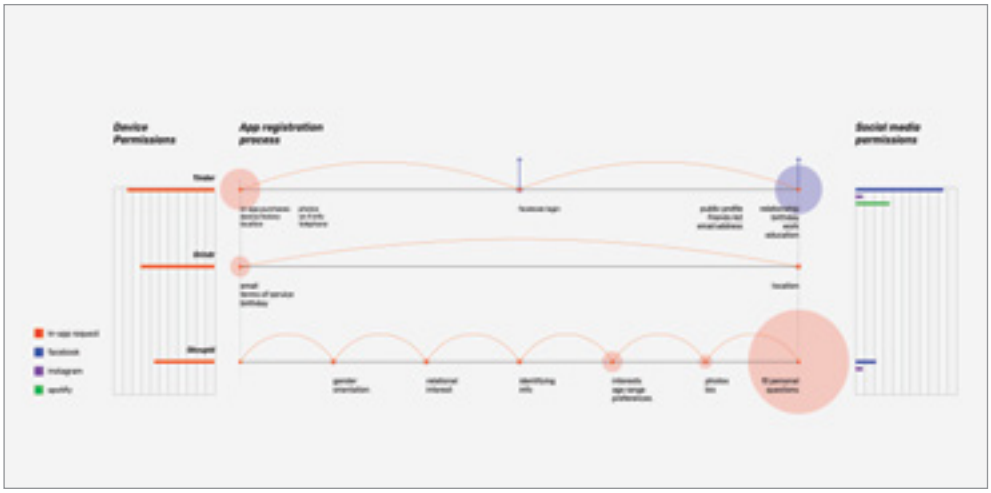
Πίνακας 6.4.1: Ανιχνευμένες σχέσεις εφαρμογής-πλατφόρμας ανά σύνολο πηγών

Συνδέσεις δεδομένων τρίτων

Η τρίτη μελέτη εξετάζει το λογισμικό εφαρμογών και τον τρόπο με τον οποίο αυτό σχετίζεται με τις υποδομές δεδομένων των εξωτερικών ενδιαφερομένων μερών. Με αυτού του είδους την έρευνα είναι δυνατόν να χαρτογραφηθεί ο τρόπος με τον οποίο η εφαρμογή ως αντικείμενο λογισμικού ενσωματώνει εξωτερικές υποδομές δεδομένων, καθώς και οι δυναμικές ροές δεδομένων μέσα και έξω από τις εφαρμογές (Weltevrede & Jansen, 2019). Οι εφαρμογές εμφανίζονται σε εμάς ως διακριτά και οριοθετημένα αντικείμενα, ενώ είναι εξ ορισμού αντικείμενα υποδομής δεδομένων, που σχετίζονται με πλατφόρμες, για να επεκταθούν και να ενσωματωθούν στην υποδομή δεδομένων. Προκειμένου να ενεργοποιήσουμε και να διερευνήσουμε τις εισερχόμενες και εξερχόμενες ροές δεδομένων, χρησιμοποιήσαμε μια παραλλαγή της «μεθόδου περιδιαβάσματος» («walkthrough method», Light κ.α., 2016). Εστιάζοντας στις συνδέσεις δεδομένων, η οπτικοποίηση που προκύπτει δείχνει ποια δεδομένα διοχετεύονται στις εφαρμογές από τις πλατφόρμες κοινωνικής δικτύωσης και την πλατφόρμα κινητής τηλεφωνίας (Εικόνα 6.4.2).

Σε ένα δεύτερο βήμα, χαρτογραφήσαμε τα διαφημιστικά δίκτυα, τις υπηρεσίες cloud, τα δίκτυα ανάλυσης και άλλα δίκτυα τρίτων μερών στα οποία συνδέονται οι εφαρμογές, μεταξύ άλλων, για τη χρηματική εκμετάλλευση των δεδομένων της εφαρμογής, τη βελτίωση της λειτουργικότητας ή τη διανομή της παροχής φιλοξενίας σε εξωτερικά μέρη (Εικόνα 6.4.3).

Η χαρτογράφηση των ροών δεδομένων, μέσα και έξω από τις εφαρμογές, παρέχει μια κρίσιμη εικόνα της πολιτικής οικονομίας της κυκλοφορίας και του ανασυνδυασμού των δεδομένων: οι συνδέσεις δεδομένων που δημιουργούνται, ο τρόπος με τον οποίο ενεργοποιούνται και ποιοι τύποι δεδομένων μεταφέρονται σε ποια μέρη.



Εικόνα 6.4.2: Διεπαφή περιδιαβάσματος στη ροή δεδομένων κατά τη διαδικασία εγγραφής. Πηγή: Weltevrede & Jansen (2019).

Εικόνα 6.4.3: Συνδέσεις δικτύου που δημιουργήθηκαν μεταξύ των εφαρμογών γνωριμιών Tinder, Grindr και OKCupid και των τρίτων μερών τους. Πηγή: Weltevrede & Jansen (2019).

Συμπέρασμα

Οι πλατφόρμες και οι εφαρμογές είναι τόσο θεμελιωδώς συνυφασμένες με την καθημερινή ζωή που συχνά περνούν απαρατήρητες, χωρίς κανέναν προβληματισμό. Αυτή η τάση, να περνούν σε δεύτερη μοίρα, είναι ακριβώς ο λόγος για τον οποίο οι ψηφιακοί ερευνητές, οι δημοσιογράφοι δεδομένων και οι ακτιβιστές θα πρέπει να διερευνήσουν τον τρόπο λειτουργίας τους και τις συνθήκες που διέπουν τη δημιουργία και τη χρήση τους. Είναι σημαντικό να βελτιωθεί ο αλφαβητισμός στις υποδομές δεδομένων, ώστε να γίνει κατανοητή η σύνδεσή τους με διάφορες πλατφόρμες και δίκτυα, η μεταξύ τους αλληλεπίδραση και η συχνή εμπλοκή άγνωστων ενδιαφερομένων μερών. Στον απόηχο του σκανδάλου της Cambridge Analytica, τα δεδομένα που είναι έτοιμα για κοινωνικές έρευνες και είναι προσβάσιμα μέσω δομημένων API περιορίζονται ολοένα και περισσότερο από τις πλατφόρμες ως απάντηση στη δημόσια πίεση. Σε αυτό το κεφάλαιο, πρότεινα ότι, ως ανταπάντηση, οι ερευνητές, οι δημοσιογράφοι και οι ομάδες της κοινωνίας των πολιτών θα πρέπει να είναι δημιουργικοί και εφευρετικοί στη διερεύνηση νέων τύπων δεδομένων όσον αφορά στην ερευνητική δυναμική τους. Έχω διερευνήσει τρεις τύπους δεδομένων για τη μελέτη των εφαρμογών. Υπάρχουν, άλλωστε, πολλαπλές ευκαιρίες για περαιτέρω ανάπτυξη. Θα πρέπει να τονιστεί ότι ασχολήθηκα κυρίως με τις εφαρμογές, αλλά αυτό μπορεί να προσφέρει έμπνευση για έρευνες σε διαφορετικά περιβάλλοντα, πλούσια σε δεδομένα, συμπεριλαμβανομένων των έξυπνων πόλεων και του «Διαδικτύου των πραγμάτων» («Internet of things»). Μια πιο διαφοροποιημένη αντίληψη των υποδομών δεδομένων που διαμορφώνουν ολοένα και περισσότερο τις πρακτικές της καθημερινής ζωής παραμένει ένα έργο εν εξελίξει.

Παραπομπές

Bruns, A. (2018) «Facebook shuts the gate after the horse has bolted, and hurts real research in the process», *Internet Policy Review,* 25 Απριλίου. Διαθέσιμο σε: https://policyreview.info/articles/news/facebook-shuts-gate-after-horse-has-bolted-and-hurts-real-research-process/786 (Προσπελάστηκε: 10 Δεκεμβρίου 2021).

Dieter, M., Gerlitz, C., Helmond, A., Tkacz, N., Van der Vlist, F. N., & Weltevrede, E. (2019) «Multi-situated app studies: Methods and propositions», *Social Media + Society.* DOI: https://doi.org/10.1177/2056305119846486

Gerlitz, C., Helmond, A., Van der Vlist, F. N., & Weltevrede, E. (2019). «Regramming the platform: Infrastructural relations between apps and social media», *Computational Culture,* 7. Διαθέσιμο σε: http://computationalculture.net/regramming-the-platform/ (Προσπελάστηκε: 10 Δεκεμβρίου 2021).

Gerlitz, C., & Weltevrede, E. (2019) «What happens to ANT, and its emphasis on the socio-material grounding of the social, in digital sociology?», στο A. Blok, I. Farias, & C. Roberts (επιμ.), *Companion to actor–network theory,* Λονδίνο: Routledge, σσ. 345–356. DOI: https://doi.org/10.4324/9781315111667-38

Gray, J., Gerlitz, C., & Bounegru, L. (2018) «Data infrastructure literacy», *Big Data & Society,* 5(2), σσ. 1–13. DOI: https://doi.org/10.1177/2053951718786316

Helmond, A., Nieborg, D. B., & Van der Vlist, F. N. (2019) «Facebook's evolution: Development of a platform-as-infrastructure», *Internet Histories*, 3(2), σσ. 123–146. DOI: https://doi.org/10.1080/24701475.2019.1593667

Hoffmann, A. L., Proferes, N., & Zimmer, M. (2018) «"Making the world more open and connected": Mark Zuckerberg and the discursive construction of Facebook and its users», *New Media & Society*, 20(1), σσ. 199–218. DOI: https://doi.org/10.1177/1461444816660784

Light, B., Burgess, J., & Duguay, S. (2016) «The walkthrough method: An approach to the study of apps», *New Media & Society,* 20(3), σσ. 881–900. DOI: https://doi.org/10.1177/1461444816675438

Lomas, N. (2018) «Kogan: "I don't think Facebook has a developer policy that is valid», *TechCrunch*, 24 Απριλίου. Διαθέσιμο σε: https://social.techcrunch.com/2018/04/24/kogan-i-dont-think-facebook-has-a-developer-policy-that-is-valid/ (Προσπελάστηκε: 11 Δεκεμβρίου 2021).

Lury, C., & Wakeford, N. (2012). *Inventive methods: The happening of the social.* Λονδίνο, Νέα Υόρκη: Routledge.

Marres, N. (2017). *Digital sociology: The reinvention of social research.* Κέιμπριτζ: Polity Press.

Matviyenko, S., Ticineto Clough, P., & Galloway, A. R. (2015) «On governance, blackboxing, measure, body, affect and apps: A conversation with Patricia Ticineto Clough and Alexander R. Galloway», *The Fibreculture Journal,* 25, σσ. 10–29. Διαθέσιμο σε: http://twentyfive.fibreculturejournal.org/fcj-179-on-governance-blackboxing-measure-body-affect-and-apps-a-conversation-with-patricia-ticineto-clough-and-alexander-r-galloway/ (Προσπελάστηκε: 11 Δεκεμβρίου 2021).

Nieborg, D. B., & Helmond, A. (2019) «The political economy of Facebook's platformization in the mobile ecosystem: Facebook Messenger as a platform instance», *Media, Culture & Society,* 41(2), σσ. 196–218. DOI: https://doi.org/10.1177/0163443718818384

Perriam, J., Birkbak, A., & Freeman, A. (2020) «Digital methods in a post-API environment», *International Journal of Social Research Methodology,* 23(3), σσ. 277–290. DOI: https://doi.org/10.1080/13645579.2019.1682840

Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018) «Infrastructure studies meet platform studies in the age of Google and Facebook», *New Media & Society,* 20(1), σσ. 293–310. DOI: https://doi.org/10.1177/1461444816661553

Rieder, B. (2013) «Studying Facebook via data extraction: The Netvizz application», στο *Proceedings of the 5th Annual ACM Web Science Conference,* Παρίσι, 2-4 Μαΐου 2013, σσ. 346–355. DOI: https://doi.org/10.1145/2464464.2464475

Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018) «From ranking algorithms to "ranking cultures": Investigating the modulation of visibility in YouTube search results», *Convergence*, 24(1), σσ. 50–68. DOI: https://doi.org/10.1177/1354856517736982

Rogers, R. (2013). *Digital methods.* Κέιμπριτζ, Μασαχουσέτη: MIT Press.

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014) «*Auditing algorithms: Research methods for detecting discrimination on Internet platforms», 64th Annual Meeting of the International Communication Association preconference on Data and Discrimination Converting Critical Concerns into Productive Inquiry,* Σιάτλ, Ουάσινγκτον, 22 Μαΐου 2014. Διαθέσιμο σε: http://www-personal.umich.edu/~csandvig/research/Auditing%20Algorithms%20--%20Sandvig%20--%20ICA%202014%20Data%20and%20Discrimination%20Preconference.pdf (Προσπελάστηκε: 11 Δεκεμβρίου 2021).

Smout, A., & Busvine, D. (2018) «Researcher in Facebook scandal says: My work was worthless to Cambridge Analytica», *Reuters*, 24 Απριλίου. Διαθέσιμο σε: https://www.reuters.com/article/us-facebook-privacy-cambridge-analytica-idUSKBN1HV17M (Προσπελάστηκε: 11 Δεκεμβρίου 2021).

Venturini, T., & Rogers, R. (2019) «"API-based research" or how can digital sociology and journalism studies learn from the Cambridge Analytica data breach», *Digital Journalism,* 7(4), σσ. 532–540. DOI: https://doi.org/10.1080/21670811.2019.1591927

Weltevrede, E. (2016). *Repurposing digital methods: The research affordances of platforms and*

engines [Διδακτορική διατριβή], Πανεπιστήμιο του Άμστερνταμ.

Weltevrede, E., & Jansen, F. (2019). «Infrastructures of intimate data: Mapping the inbound and outbound data flows of dating apps», *Computational Culture*, 7, 21 Οκτωβρίου. Διαθέσιμο σε: http://computationalculture.net/infrastructures-of-intimate-data-mapping-the-inbound-and-outbound-data-flows-of-dating-apps/ (Προσπελάστηκε: 10 Δεκεμβρίου 2021).

Η συγγραφέας

Η Έσθερ Βέλτεβρεντ είναι Επίκουρη Καθηγήτρια Νέων Μέσων και Ψηφιακού Πολιτισμού και συντονίστρια της Πρωτοβουλίας Ψηφιακών Μεθόδων (Digital Methods Initiative) στο Πανεπιστήμιο του Άμστερνταμ, όπου διερευνά τις διάφορες ερευνητικές δυνατότητες των ψηφιακών μέσων, με ιδιαίτερο ενδιαφέρον για τις μεθοδολογικές καινοτομίες στη μελέτη των υποδομών των πλατφορμών των μέσων κοινωνικής δικτύωσης και των τρίτων μερών.

6.5

Ρίχνοντας Φως στους Αλγορίθμους: Συνεργατικές Έρευνες στο *Der Spiegel*

Christina Elmer

Περίληψη: Χρησιμοποιώντας διάφορα παραδείγματα από τα γερμανικά μέσα ενημέρωσης, το κεφάλαιο αυτό παρέχει πληροφορίες για συνεργατικά έργα που διερευνούν αλγορίθμους και συνοψίζει τα σημαντικότερα συμπεράσματα.

Λέξεις-κλειδιά: αλγόριθμοι, αλγοριθμική λογοδοσία, δημοσιογραφία δεδομένων, crowdsourcing, διαφάνεια, συνεργασία

Το αίτημα για διαφάνεια σχετικά με τους αλγορίθμους δεν είναι καινούργιο στη Γερμανία. Το 2012, ο αρθρογράφος του *Der Spiegel*, Σάσα Λόμπο, ζήτησε να γνωστοποιηθούν οι μηχανισμοί του αλγορίθμου αναζήτησης της Google (Lobo, 2012), ακόμη και αν αυτό θα έβλαπτε την εταιρεία. Ο λόγος ήταν ότι η Google μπορεί να διαμορφώσει τον τρόπο με τον οποίο αντιλαμβανόμαστε τον κόσμο, για παράδειγμα μέσω της λειτουργίας αυτόματης συμπλήρωσης, όπως έδειξε μια χαρακτηριστική περίπτωση στη Γερμανία. Στην περίπτωση αυτή, η σύζυγος του πρώην ομοσπονδιακού προέδρου είχε στραφεί νομικά κατά της Google, επειδή στην αυτόματη συμπλήρωση προτείνονταν προβληματικοί όροι, όταν γινόταν αναζήτηση του ονόματός της. Δύο χρόνια αργότερα, ο Γερμανός Υπουργός Δικαιοσύνης επανέλαβε αυτό το αίτημα, για το οποίο υπερθεμάτισε η ομοσπονδιακή καγκελάριος το 2016: η Άνγκελα Μέρκελ απαίτησε να γίνουν οι αλγόριθμοι πιο διαφανείς (Kartell, 2014· Reinbold, 2016).

Τα τελευταία χρόνια, το θέμα της αλγοριθμικής λογοδοσίας συζητείται διαρκώς στο *Der Spiegel*, αλλά αρχικά μόνο ως αφορμή για ρεπορτάζ και όχι με τη μορφή δικού μας ερευνητικού ή αναλυτικού έργου. Μπορεί να υπάρχουν δύο βασικοί λόγοι για τους οποίους τα γερμανικά μέσα ενημέρωσης άρχισαν να πειραματίζονται σε αυτόν τον τομέα αργότερα από τους συναδέλφους τους στις Ηνωμένες Πολιτείες. Πρώτον, οι δημοσιογράφοι στη Γερμανία δεν έχουν στη διάθεσή τους τόσο ισχυρά δικαιώματα και μέσα σχετικά με την ελευθερία της πληροφόρησης («freedom of information», FOI). Δεύτερον, η δημοσιογραφία δεδομένων δεν έχει τόσο μακρά παράδοση στη χώρα, όσο στις Ηνωμένες Πολιτείες. Το *Der Spiegel* απέκτησε το δικό του τμήμα δημοσιογραφίας δεδομένων μόλις το 2016 και επεκτείνεται αργά αλλά σταθερά σε αυτόν τον τομέα. Είναι, φυσικά, εφικτό να δραστηριοποιηθούν σε αυτόν τον τομέα και εφημερίδες με μικρότερους οικονομικούς πόρους –για παράδειγμα, μέσω της συνεργασίας με οργανισμούς ή με ελεύθερους επαγγελματίες. Και στη δική μας περίπτωση, όλα τα προηγούμενα έργα στον τομέα του ρεπορτάζ αλγοριθμικής λογοδοσίας προέκυψαν με αυτόν τον τρόπο. Ως εκ τούτου, το παρόν κεφάλαιο θα επικεντρωθεί στις συνεργασίες και στα διδάγματα που αντλήσαμε από αυτές.

Google, Facebook και Schufa: τρία έργα με μια ματιά

Η συντακτική μας ομάδα βασίζεται κυρίως στη συνεργασία, όταν πρόκειται για τη διερεύνηση αλγορίθμων. Ενόψει των ομοσπονδιακών εκλογών του 2017, ενώσαμε τις δυνάμεις μας

με τη ΜΚΟ AlgorithmWatch, για να αποκτήσουμε πληροφορίες σχετικά με την εξατομίκευση των αποτελεσμάτων αναζήτησης της Google[186]. Οι χρήστες κλήθηκαν να εγκαταστήσουν μια πρόσθετη εφαρμογή («plug-in») που εκτελούσε τακτικά προκαθορισμένες αναζητήσεις στον υπολογιστή τους. Συνολικά, περίπου 4.400 συμμετέχοντες παραχώρησαν σχεδόν έξι εκατομμύρια αποτελέσματα αναζήτησης και παρείχαν, έτσι, τα δεδομένα για μια ανάλυση που θα αμφισβητούσε τη θέση περί «φίλτρου-φούσκας» –τουλάχιστον όσον αφορά στην Google και τον υπό εξέταση τομέα.

Για το έργο αυτό, οι συνεργάτες μας από το AlgorithmWatch προσέγγισαν το *Der Spiegel*, καθώς ζητούσαν να συνεργαστούν με ένα μέσο με μεγάλη απήχηση για τη συλλογή των απαιτούμενων δεδομένων με crowdsourcing[187]. Ενώ το περιεχόμενο του ρεπορτάζ ήταν εξ ολοκλήρου ευθύνη του τμήματός μας, που καλύπτει θέματα σχετικά με το Διαδίκτυο και την τεχνολογία, το τμήμα δημοσιογραφίας δεδομένων υποστήριξε τον σχεδιασμό και τη μεθοδολογική αξιολόγηση του εγχειρήματος. Επιπλέον, η υποστήριξη από το νομικό μας τμήμα ήταν απαραίτητη για την υλοποίηση του έργου με τρόπο που να είναι νομικά ακλόνητος. Για παράδειγμα, τα θέματα προστασίας των δεδομένων έπρεπε να αποσαφηνιστούν στο πλαίσιο του ρεπορτάζ και να είναι πλήρως κατανοητά για όλους τους συμμετέχοντες στο έργο.

Σχεδόν την ίδια περίοδο, το *Der Spiegel* συνεργάστηκε με την *ProPublica* για την ανάπτυξη του AdCollector στη Γερμανία τους μήνες πριν από τις εκλογές (Angwin & Larson, 2017). Το έργο είχε ως στόχο να εξασφαλίσει διαφάνεια στον τρόπο με τον οποίο τα γερμανικά κόμματα απευθύνονται στους χρήστες του Facebook με διαφημίσεις. Ως εκ τούτου, με ένα plug-in συλλέχθηκαν οι πολιτικές διαφημίσεις που βλέπουν οι χρήστες στη ροή ενημερώσεων και αποκαλύφθηκαν οι διαφημίσεις που δεν τους εμφανίζονται. Για το έργο αυτό, το *Der Spiegel* ένωσε τις δυνάμεις του με άλλα γερμανικά μέσα ενημέρωσης, όπως η *Süddeutsche Zeitung* και το *Tagesschau* –ένας ασυνήθιστος συνδυασμός οργανισμών που συνήθως ανταγωνίζονται μεταξύ τους. Στην περίπτωση αυτή, ήταν απαραίτητο να επιτευχθεί η προσέγγιση όσο το δυνατόν περισσότερων ανθρώπων για την εξυπηρέτηση του δημόσιου συμφέροντος. Τα αποτελέσματα θα μπορούσαν, επίσης, να δημοσιευθούν ως δημοσιογραφικά θέματα, αλλά η πρωταρχική μας εστίαση ήταν η διαφάνεια. Μετά από δύο εβδομάδες, είχαν συγκεντρωθεί και διατεθεί στο κοινό περίπου 600 πολιτικές διαφημίσεις.

Οι Τζούλια Άνγκουιν και Τζεφ Λάρσον από την *ProPublica* παρουσίασαν την ιδέα της συνεργασίας στο ετήσιο συνέδριο της γερμανικής ένωσης ερευνητών δημοσιογράφων, Netzwerk Recherche, στο Αμβούργο, όπου έκαναν μια συζήτηση σχετικά με το ρεπορτάζ για την αλγοριθμική λογοδοσία. Η ιδέα αναπτύχθηκε από την αρχή σε συνεργασία με ειδικούς σε θέματα τεχνικής και μεθοδολογίας από διάφορα τμήματα της αίθουσας σύνταξης του *Der Spiegel*. Η ανταλλαγή απόψεων με τον προηγούμενο εταίρο μας, τον μη κερδοσκοπικό οργανισμό AlgorithmWatch, ήταν επίσης πολύ πολύτιμη για εμάς, προκειμένου να ρίξουμε φως στο νομικό υπόβαθρο και να το συμπεριλάβουμε στην έρευνά μας. Μετά το συνέδριο, αναπτύξαμε περαιτέρω την ιδέα μέσω τακτικών τηλεφωνικών διασκέψεων. Οι συνεργάτες μας από τα άλλα γερμανικά μέσα ενημέρωσης συμμετείχαν, επίσης, σε μεταγενέστερα στάδια.

Το 2018, το *Der Spiegel* συμμετείχε σε ένα μεγάλο έργο για τη διερεύνηση ενός εξαιρετικά ισχυρού αλγορίθμου στη Γερμανία –της έκθεσης πίστωσης Schufa. Η έκθεση χρησιμοποιείται για την αξιολόγηση της πιστοληπτικής ικανότητας ιδιωτών και οφείλει να αναφέρει την πιθανότητα να μπορεί κάποιος να εξοφλήσει τους λογαριασμούς του, να πληρώσει το ενοίκιο ή να εξυπη-

186. https://algorithmwatch.org/filterblase-geplatzt-kaum-raum-fuer-personalisierung-bei-google-suchen-zur-bundestagswahl-2017/ (στα γερμανικά)

187. ΣτΕ: Δηλαδή, με τη μέθοδο του «πληθοπορισμού», όπως λέγεται στα ελληνικά, υποδηλώνοντας τις προσπάθειες και τα έργα που γίνονται με τη συμμετοχή του πλήθους –εν προκειμένω, οι χρήστες συνέβαλαν στη συλλογή των δεδομένων με τη συνεισφορά των δικών τους δεδομένων αναζήτησης.

ρετήσει ένα δάνειο. Ως εκ τούτου, μπορεί να έχει εκτεταμένες επιπτώσεις στην ιδιωτική ζωή ενός ατόμου και αρνητικό αντίκτυπο στην κοινωνία στο σύνολό της. Για παράδειγμα, είναι πιθανό η βαθμολογία να αυξήσει τις κοινωνικές διακρίσεις και την άνιση μεταχείριση των ατόμων, ανάλογα με τον όγκο των δεδομένων που είναι διαθέσιμα για εκείνα. Λανθασμένα δεδομένα ή κάποιο μπέρδεμα θα μπορούσαν να αποβούν μοιραία για μεμονωμένα άτομα. Η υποκείμενη βαθμολόγηση του αλγορίθμου δεν είναι διαφανής. Δεν είναι γνωστό ποια δεδομένα λαμβάνονται υπόψη και με ποια στάθμιση. Και όσοι επηρεάζονται συχνά δεν έχουν καμία γνώση της διαδικασίας. Αυτό καθιστά τη Schufa έναν αμφιλεγόμενο φορέα στη Γερμανία –και έργα όπως το OpenSCHUFA είναι απολύτως ζωτικής σημασίας για τη δημόσια συζήτηση σχετικά με την αλγοριθμική λογοδοσία, κατά τη γνώμη μας[188].

Το έργο προωθήθηκε κυρίως από τις ΜΚΟ Open Knowledge Foundation (OKFN) και AlgorithmWatch. Το *Der Spiegel* ήταν ένας από τους δύο συνεργαζόμενους εταίρους, μαζί με την Bayerischer Rundfunk (Βαυαρική Ραδιοτηλεόραση). Η ιδέα για αυτό το έργο προέκυψε λίγο πολύ ταυτόχρονα, με τη συμμετοχή πολλών μερών. Μετά από ορισμένα επιτυχημένα έργα με τις ΜΚΟ AlgorithmWatch και OKFN, καθώς και με την ομάδα δημοσιογραφίας δεδομένων της Bayerischer Rundfunk, το *Der Spiegel* συμπεριλήφθηκε στις αρχικές συζητήσεις.

Αυτός ο συνασπισμός δημιουργούσε ιδιαίτερες προκλήσεις. Για τις δύο δημοσιογραφικές ομάδες, ήταν σημαντικό να εργαστούν χωριστά από τις ΜΚΟ, κυρίως προκειμένου να διασφαλιστεί η ανεξαρτησία τους από τη διαδικασία της χρηματοδότησης από το πλήθος («crowdfunding»). Ως εκ τούτου, αν και υπήρξαν, φυσικά, συζητήσεις μεταξύ των εμπλεκόμενων φορέων, δεν κατέστη δυνατή ούτε μία επίσημη εταιρική συνεργασία, ούτε μία κοινή αξιολόγηση των δεδομένων. Το παράδειγμα αυτό υπογραμμίζει πόσο σημαντικό είναι για τους δημοσιογράφους να προβληματιστούν σχετικά με την αυτονομία τους, ειδικά σε θέματα τόσο μεγάλης δημοσιότητας.

Η προβολή του OpenSCHUFA ήταν ένας από τους κεντρικούς παράγοντες επιτυχίας αυτού του έργου. Το πρώτο βήμα ήταν η χρήση crowdfunding για τη δημιουργία της απαραίτητης υποδομής για τη συλλογή των δεδομένων, τα οποία αποκτήθηκαν μέσω crowdsourcing. Τα αποτελέσματα αξιολογήθηκαν από κοινού από τους εταίρους, κατά τη διάρκεια του έτους, σε ανωνυμοποιημένη μορφή. Το κεντρικό ερώτημα που κρύβεται πίσω από αυτό είναι αν ο αλγόριθμος της Schufa κάνει διακρίσεις εις βάρος ορισμένων πληθυσμιακών ομάδων και αν αυξάνει την ανισότητα στην κοινωνία. Σύμφωνα με τα αποτελέσματα, αυτό ακριβώς συμβαίνει. Διαπιστώσαμε ότι η βαθμολογία ευνοούσε τους ηλικιωμένους και τις γυναίκες, καθώς και εκείνους που αλλάζουν λιγότερο συχνά κατοικία. Και ανακαλύψαμε ότι διαφορετικές εκδόσεις αλγορίθμων εντός της κλίμακας βαθμολόγησης παρήγαγαν διαφορετικά αποτελέσματα για άτομα με τα ίδια χαρακτηριστικά, ένα είδος διάκρισης που δεν ήταν προηγουμένως γνωστό σχετικά με αυτή τη βαθμολογία.

Η εξαγωγή αυτών των αποτελεσμάτων δεν θα ήταν δυνατή χωρίς τη συμμετοχή πολλών εθελοντών και υποστηρικτών. Η εκστρατεία crowdfunding ήταν σε μεγάλο βαθμό επιτυχής και η χρηματοδότηση του λογισμικού εξασφαλίστηκε βάσει πλάνου[189]. Και στη διαδικασία του crowdsourcing που ακολούθησε, περίπου 2.800 άτομα έστειλαν τις προσωπικές τους εκθέσεις πίστωσης. Το δείγμα αυτό δεν ήταν, φυσικά, αντιπροσωπευτικό, αλλά διασφάλιζε επαρκώς, τη διαφοροποίηση του δείγματος για να αποκαλύψει τα ευρήματα που περιγράφονται.

Δείκτες επιπτώσεων και επιτυχίας

Τόσο η έρευνα για το Facebook όσο και η έρευνα για τη Google ήταν μάλλον μη θεαματικές από άποψη ευρημάτων και επιβεβαίωσαν τις υποθέσεις μας. Φαίνεται ότι τα πολιτικά κόμματα

188. http://www.openschufa.de/ (στα γερμανικά)
189. https://www.startnext.com/openschufa (στα γερμανικά)

δεν χρησιμοποίησαν σχεδόν καθόλου τις επιλογές στόχευσης του Facebook και το περιβόητο «φίλτρο-φούσκα» της Google δεν εντοπίστηκε στο πείραμα crowdsourcing που κάναμε στη Γερμανία. Ωστόσο, για εμάς, η αξία αυτών των έργων έγκειται στην αύξηση του αλφαβητισμού των αναγνωστών μας γύρω από τις λειτουργίες και τους κινδύνους που σχετίζονται με τους αλγορίθμους για την κοινωνία.

Η απήχηση των άρθρων μας ήταν ένας δείκτης ότι είχαμε καταφέρει να κάνουμε το θέμα ευρύτερα γνωστό. Το εισαγωγικό άρθρο στην αρχή του έργου για τη Schufa είχε απήχηση σε μεγάλο κοινό (περίπου 335.000 αναγνώστες)[190]. Ο χρόνος ανάγνωσης ήταν, επίσης, σαφώς μεγαλύτερος από τον συνηθισμένο –με μέσο όρο σχεδόν τα τρία λεπτά. Επιπλέον, το θέμα συζητήθηκε ευρέως σε δημόσιους χώρους και καλύφθηκε από πολλά μέσα ενημέρωσης και συνέδρια.

Το θέμα έχει συζητηθεί, επίσης, και σε πολιτικούς κύκλους. Μετά τη δημοσίευση των ερευνών για τη Schufa, η Γερμανίδα Υπουργός Προστασίας των Καταναλωτών ζήτησε μεγαλύτερη διαφάνεια στον τομέα της βαθμολόγησης της πιστοληπτικής ικανότητας. Η ίδια απαίτησε να γνωρίζει ο κάθε πολίτης ποια βασικά χαρακτηριστικά έχουν συμπεριληφθεί στον υπολογισμό της πιστοληπτικής του ικανότητας και πώς αυτά σταθμίζονται.

Και ο αντίκτυπος στην πραγματικότητα της καθημερινής ζωής; Ως ένα πρώτο βήμα, ήταν σημαντικό για εμάς να συμβάλουμε στην εδραίωση του θέματος στη συνείδηση του κοινού. Μέχρι στιγμής, δεν έχουμε δει οι πολιτικοί φορείς να αντιμετωπίζουν τους αλγόριθμους που έχουν ευρύτερες κοινωνικές συνέπειες με ριζικά διαφορετικό τρόπο.

Ωστόσο, το θέμα του ρεπορτάζ αλγοριθμικής λογοδοσίας είναι πολύ σημαντικό για εμάς. Αυτό οφείλεται στο γεγονός ότι στην Ευρώπη ακόμη έχουμε την ευκαιρία να συζητήσουμε το θέμα των αλγορίθμων στο πλαίσιο της κοινωνίας και να διαμορφώσουμε τον τρόπο με τον οποίο θέλουμε να το χειριστούμε. Αποτελεί μέρος της λειτουργίας μας ως δημοσιογράφων, να παρέχουμε τις απαραίτητες γνώσεις, ώστε οι πολίτες να μπορούν να κατανοήσουν και να διαμορφώσουν το μέλλον των αλγορίθμων στην κοινωνία. Στο μέτρο του δυνατού, αναλαμβάνουμε, επίσης, τον ρόλο του παρατηρητή-ελεγκτή, προσπαθώντας να καταστήσουμε τους αλγορίθμους και τα αποτελέσματά τους διαφανή, να εντοπίσουμε τους κινδύνους και να έρθουμε αντιμέτωποι με τους υπευθύνους. Για να το επιτύχουμε αυτό, πρέπει να καθιερώσουμε συνεργασίες που διαφορετικά θα θεωρούνταν ασυνήθιστες, με ανταγωνιστές και φορείς από άλλους τομείς. Ελπίζουμε ότι τέτοιες συμμαχίες, τελικά, θα αυξήσουν την πίεση για νομοθεσία και πρότυπα διαφάνειας στον τομέα αυτό.

Περισσότερες προσπάθειες και πόροι πρέπει να αφιερωθούν στις έρευνες για την αλγοριθμική λογοδοσία –το Markup έχει δημοσιεύσει μερικές πολύ συναρπαστικές έρευνες στον τομέα αυτό. Απαιτείται περαιτέρω πειραματισμός, εν μέρει επειδή ακόμη υπάρχουν περιθώρια για δράση όσον αφορά στη νομοθετική ρύθμιση για τους αλγορίθμους. Ο τομέας του ρεπορτάζ για την αλγοριθμική λογοδοσία έχει αρχίσει να αναπτύσσεται μόλις τα τελευταία χρόνια. Και θα πρέπει να αναπτυχθεί ταχύτατα, για να ανταποκριθεί στις προκλήσεις ενός ολοένα και περισσότερο ψηφιοποιημένου κόσμου.

Οργάνωση συνεργατικών ερευνών

Η διεξαγωγή συνεργατικών ερευνών απαιτεί ένα σύνολο νέων ή λιγότερο χρησιμοποιούμενων δεξιοτήτων στην αίθουσα σύνταξης. Αυτό περιλαμβάνει την ανάλυση μεγάλων συνόλων δε-

190. Η πλειονότητα αυτών, ωστόσο, κατέληξε στο άρθρο μέσω εσωτερικών καναλιών, όπως η αρχική μας σελίδα. Αυτό δεν συνέβη στην περίπτωση ενός άλλου άρθρου, μιας επιτόπιας έρευνας με την προσωπική ιστορία του συγγραφέα της, το οποίο διαβάστηκε από περίπου 220.000 άτομα. Το ένα πέμπτο από αυτά έφτασε στο άρθρο μέσω των μέσων κοινωνικής δικτύωσης, ποσοστό πολύ υψηλότερο από τον μέσο όρο. Φαίνεται, λοιπόν, ότι καταφέραμε να προσεγγίσουμε νέες ομάδες-στόχους με αυτή τη θεματολογία.

δομένων και τον προγραμματισμό συγκεκριμένων διαδικασιών αλλά και τη διαχείριση έργων. Το τελευταίο είναι πολύ εύκολο να το παραβλέψουμε και θα το περιγράψουμε λεπτομερέστερα εδώ, με συγκεκριμένα παραδείγματα από προηγούμενα έργα μας.

Η συνεργασία σε ποικιλόμορφους συνδυασμούς όχι μόνο διευκολύνει την ανταλλαγή δεξιοτήτων και πόρων, αλλά επιτρέπει, επίσης, τον σαφή καθορισμό των ρόλων. Ως μέσο-εταίρος, το *Der Spiegel* κινήθηκε σε αυτές τις συνεργασίες περισσότερο ως ουδέτερος σχολιαστής, χωρίς να εμπλέκεται ιδιαίτερα στο ίδιο το έργο. Αυτό επέτρεψε στους συντάκτες να παραμείνουν ανεξάρτητοι και, έτσι, να κερδίσουν την εμπιστοσύνη των αναγνωστών τους: συνέχισαν να εφαρμόζουν τα ποιοτικά τους κριτήρια για το ρεπορτάζ, στο πλαίσιο του έργου –για παράδειγμα, δίνοντας πάντα την ευκαιρία σε κάθε πρόσωπο που εμπλεκόταν στο ρεπορτάζ τους να σχολιάσει τις κατηγορίες. Σε σύγκριση με τις εμπλεκόμενες ΜΚΟ, οι μηχανισμοί αυτοί μπορεί να ρίξουν τους ρυθμούς των μέσων-εταίρων περισσότερο από όσο τους αρέσει, αλλά ταυτόχρονα διασφαλίζουν ότι οι αναγνώστες ενημερώνονται πλήρως από τα δημοσιεύματά τους –και ότι αυτά θα εμπλουτίσουν μακροπρόθεσμα τον δημόσιο διάλογο.

Η επίτευξη συμφωνίας σχετικά με τους ρόλους αυτούς εκ των προτέρων έχει αποδειχθεί σημαντικό κριτήριο επιτυχίας για τις συνεργασίες στον τομέα της αλγοριθμικής λογοδοσίας. Θα πρέπει, επίσης, να αναπτυχθεί ένα κοινό χρονοδιάγραμμα σε πρώιμο στάδιο και να καθοριστούν γλωσσικοί κανόνες για την παρουσίαση του έργου στους διάφορους διαύλους επικοινωνίας. Διότι, σε τελική ανάλυση, μια σαφής κατανομή ρόλων μπορεί να λειτουργήσει μόνο αν εφαρμόζεται με συνέπεια. Αυτό περιλαμβάνει, για παράδειγμα, μια σαφή ορολογία σχετικά με τους ρόλους των διαφόρων εταίρων στο έργο και τον συντονισμό των δηλώσεων αποποίησης ευθύνης σε περίπτωση σύγκρουσης συμφερόντων.

Στο παρασκήνιο, οι μέθοδοι διαχείρισης έργου θα πρέπει να χρησιμοποιούνται με σύνεση, οι στόχοι του έργου θα πρέπει να τίθενται με σαφήνεια και οι διαθέσιμοι πόροι θα πρέπει να συζητούνται. Οι συντονιστές θα πρέπει να βοηθούν στη συνολική επικοινωνία και να δίνουν, έτσι, στους συμμετέχοντες συντάκτες τον χώρο που χρειάζονται για τις έρευνές τους. Προκειμένου να είναι όλοι ενήμεροι, οι δίαυλοι πληροφόρησης θα πρέπει να είναι όσο το δυνατόν πιο απλοί, ιδίως κατά την έναρξη σημαντικών σταδίων του έργου.

Όσον αφορά τον εκδοτικό σχεδιασμό, οι τρεις θεματικές ενότητες αποτέλεσαν πρόκληση. Αν και γενικά η συνάφεια και η αξία της είδησης δεν αμφισβητήθηκαν ποτέ, χρειάζονταν ξεχωριστές ιστορίες, για να προσεγγίσουμε ένα ευρύ αναγνωστικό κοινό. Συχνά, οι ιστορίες αυτές επικεντρώνονταν στις προσωπικές επιπτώσεις των υπό εξέταση αλγορίθμων. Για παράδειγμα, τα λανθασμένα δεδομένα της Schufa δυσκόλεψαν έναν συνάδελφο της συντακτικής ομάδας του *Der Spiegel* να συνάψει σύμβαση με έναν πάροχο Διαδικτύου. Η αναφορά στην εμπειρία του έδειξε με εντυπωσιακό τρόπο τις επιπτώσεις που μπορεί να έχει ο αλγόριθμος της Schufa σε προσωπικό επίπεδο και, έτσι, συνδέθηκε με την πραγματικότητα της ζωής του κοινού μας (Seibt, 2018).

Έτσι, προσαρμόσαμε το εύρος του ρεπορτάζ μας στα ενδιαφέροντα του κοινού μας, στο μέτρο του δυνατού. Φυσικά, οι εμπλεκόμενοι δημοσιογράφοι δεδομένων ενδιαφέρονταν, επίσης, πολύ για τη λειτουργία των υπό διερεύνηση αλγορίθμων –ένα ενδιαφέρον που είναι εξαιρετικά χρήσιμο για ερευνητικούς σκοπούς. Ωστόσο, μόνο αν αυτές οι λεπτομέρειες έχουν σχετική επιρροή στα αποτελέσματα των αλγορίθμων μπορούν να αποτελέσουν αντικείμενο δημοσιογραφικής κάλυψης –και μόνο αν η αφήγησή τους γίνεται με τρόπο προσιτό στους αναγνώστες μας.

Εσωτερικά στη διεύθυνση σύνταξης, η υποστήριξη και για τα τρία έργα ήταν πολύ μεγάλη. Ωστόσο, δεν ήταν εύκολο να απελευθερωθούν πόροι για καθημερινό ρεπορτάζ στην καθημερινή ρουτίνα μιας συντακτικής ομάδας με γνώμονα τις ειδήσεις –ειδικά όταν τα αποτελέσματα των ερευνών μας δεν ήταν πάντα θεαματικά.

Τι μάθαμε

Εν κατακλείδι, συνοψίζω τα όσα μάθαμε από αυτά τα έργα.

Συνεργαστείτε όπου είναι δυνατόν. Οι καλές έρευνες αλγοριθμικής λογοδοσίας είναι εφικτές μόνο με τη συνένωση δυνάμεων με άλλους και τη δημιουργία ομάδων με διαφορετικές δεξιότητες. Αυτό είναι, επίσης, σημαντικό βάσει της έλλειψης πόρων και των νομικών περιορισμών που αντιμετωπίζουν οι περισσότεροι δημοσιογράφοι. Δεδομένου, όμως, ότι τα έργα αυτά συγκεντρώνουν φορείς από διαφορετικούς τομείς, είναι ζωτικής σημασίας να συζητηθούν εκ των προτέρων τα θεμελιώδη κριτήρια συνάφειας, οι απαιτήσεις και οι δυνατότητες.

Καθορίστε τους στόχους σας συστηματικά. Η ευαισθητοποίηση σχετικά με τις αρχές λειτουργίας των αλγορίθμων μπορεί να είναι ένας πρώτος ισχυρός στόχος σε τέτοια έργα. Φυσικά, τα έργα θα πρέπει, επίσης, να προσπαθούν να επιτύχουν όσο το δυνατόν μεγαλύτερη διαφάνεια. Στην καλύτερη περίπτωση, θα μπορούσαμε να ελέγξουμε αν οι αλγόριθμοι κάνουν διακρίσεις –αλλά οι εταίροι του έργου θα πρέπει να έχουν κατά νου ότι αυτός είναι ένας πιο δύσκολος στόχος που απαιτεί εκτεταμένα σύνολα δεδομένων και πόρους.

Να είστε προσεκτικοί κατά την υλοποίηση του έργου. Ανάλογα με τον φόρτο εργασίας και την καθημερινή πίεση των εμπλεκόμενων δημοσιογράφων, μπορεί να χρειαστείτε ακόμη και έναν project manager. Λάβετε υπόψη ότι το χρονοδιάγραμμα του έργου μπορεί μερικές φορές να έρχεται σε σύγκρουση με τις απαιτήσεις του ρεπορτάζ. Να το λαμβάνετε υπόψη σας κατά την επικοινωνία με άλλους εταίρους και, αν είναι δυνατόν, να προετοιμάζετε εναλλακτικές λύσεις για τέτοιες περιπτώσεις.

Επενδύστε στον σχεδιασμό της έρευνας. Για να δημιουργήσετε ένα ουσιαστικό σχέδιο που παράγει χρήσιμα δεδομένα, ίσως να χρειαστείτε εξειδικευμένους συνεργάτες. Οι στενές συμμαχίες με επιστήμονες από τον χώρο της πληροφορικής, των μαθηματικών και συναφών κλάδων είναι ιδιαίτερα χρήσιμες για τη διερεύνηση ορισμένων από τις πιο τεχνικές πτυχές των αλγορίθμων. Επιπλέον, μπορεί να είναι επίσης χρήσιμη η συνεργασία με κοινωνικούς και πολιτισμικούς ερευνητές, για να αποκτήσουμε μια βαθύτερη κατανόηση των ταξινομήσεων και των κανόνων που εφαρμόζονται σε αυτές.

Προστατέψτε τα δεδομένα των χρηστών. Οι δωρεές δεδομένων από χρήστες μπορεί να είναι χρήσιμες για τη διερεύνηση αλγορίθμων. Σε τέτοιου είδους έργα crowdsourcing, είναι απαραίτητη η νομική υποστήριξη, προκειμένου να διασφαλιστεί η προστασία των δεδομένων και να ληφθούν υπόψη οι απαιτήσεις της εθνικής νομοθεσίας και των σχετικών κανονισμών. Εάν η εταιρεία σας διαθέτει υπεύθυνο προστασίας δεδομένων, εμπλέξτε τον από νωρίς στο έργο.

Παραπομπές

Angwin, J., & Larson, J. (2017) «Help us monitor political ads online», *Pro-Publica*, 7 Σεπτεμβρίου. Διαθέσιμο σε: https://www.propublica.org/article/help-us-monitor-political-ads-online (Προσπελάστηκε: 11 Δεκεμβρίου 2021).

Kartell, S. vor. (2014) «Maas hätte gerne, dass Google geheime Suchformel offenlegt», *Der Spiegel*, 16 Σεπτεμβρίου. Διαθέσιμο σε: https://www.spiegel.de/wirtschaft/unternehmen/google-heiko-maas-fordert-offenlegung-von-algorithmus-a-991799.html (Προσπελάστηκε: 11 Δεκεμβρίου 2021).

Lobo, S. (2012) «Was Bettina Wulff mit Mettigeln verbindet», *Der Spiegel,* 11 Σεπτεμβρίου. Διαθέσιμο σε: https://www.spiegel.de/netzwelt/netzpolitik/google-suchvorschlaege-was-bettina-wulff-mit-mettigeln-verbindet-a-855097.html (Προσπελάστηκε: 11 Δεκεμβρίου 2021).

Reinbold, F. (2016) «Warum Merkel an die Algorithmen will», *Der Spiegel*, 26 Οκτωβρίου. Δια-

θέσιμο σε: https://www.spiegel.de/netzwelt/netzpolitik/angela-merkel-warum-die-kanzlerin-an-die-algorithmen-von-facebook-will-a-1118365.html (Προσπελάστηκε: 11 Δεκεμβρίου 2021).

Seibt, P. (2018) «Wie ich bei der Schufa zum "deutlich erhöhten Risiko" wurde», *Der Spiegel,* 9 Μαρτίου. Διαθέσιμο σε: https://www.spiegel.de/wirtschaft/service/schufa-wie-ich-zum-deutlich-erhoehten-risiko-wurde-a-1193506.html (Προσπελάστηκε: 11 Δεκεμβρίου 2021).

Η συγγραφέας

Η Κριστίνα Έλμερ δημιούργησε το τμήμα δημοσιογραφίας δεδομένων και προωθεί έργα καινοτομίας στη συντακτική ομάδα έρευνας και ανάπτυξης (Research & Development, R&D) του *Der Spiegel*, ενώ είναι μέλος του διοικητικού συμβουλίου του Netzwerk Recherche.

7.

Οργάνωση της Δημοσιογραφίας Δεδομένων

7.1
Το Hashtag #ddj στο Twitter

Eunice Au, Marc Smith

Περίληψη: Πώς χρησιμοποιήσαμε το πακέτο NodeXL για την ανάλυση και οπτικοποίηση της κοινωνικής δικτύωσης, ώστε να μελετήσουμε τι τουιτάρει η παγκόσμια κοινότητα της δημοσιογραφίας δεδομένων.

Λέξεις-κλειδιά: #ddj, Twitter, ανάλυση κοινωνικής δικτύωσης, δημοσιογραφία δεδομένων, μέσα κοινωνικής δικτύωσης, ανάλυση δεδομένων

Το να επιλέξει κανείς έναν και μόνο όρο για την ιχνηλάτηση του πεδίου της δημοσιογραφίας δεδομένων δεν είναι εύκολο. Οι δημοσιογράφοι δεδομένων χρησιμοποιούν πλήθος hashtag που σχετίζονται με τη δουλειά τους, όπως #datajournalism, #ddj, #dataviz, #infographics και #data. Όταν το Παγκόσμιο Δίκτυο Ερευνητικής Δημοσιογραφίας (Global Investigative Journalism Network, GIJN) –μια διεθνής ένωση οργανισμών ερευνητικής δημοσιογραφίας που υποστηρίζει την εκπαίδευση και την ανταλλαγή πληροφοριών μεταξύ ερευνητών και δημοσιογράφων δεδομένων– άρχισε για πρώτη φορά να καταγράφει τις συζητήσεις γύρω από τη δημοσιογραφία δεδομένων στο Twitter, πριν από έξι χρόνια, το πιο δημοφιλές hashtag ήταν το #ddj («data-driven journalism»)[191].

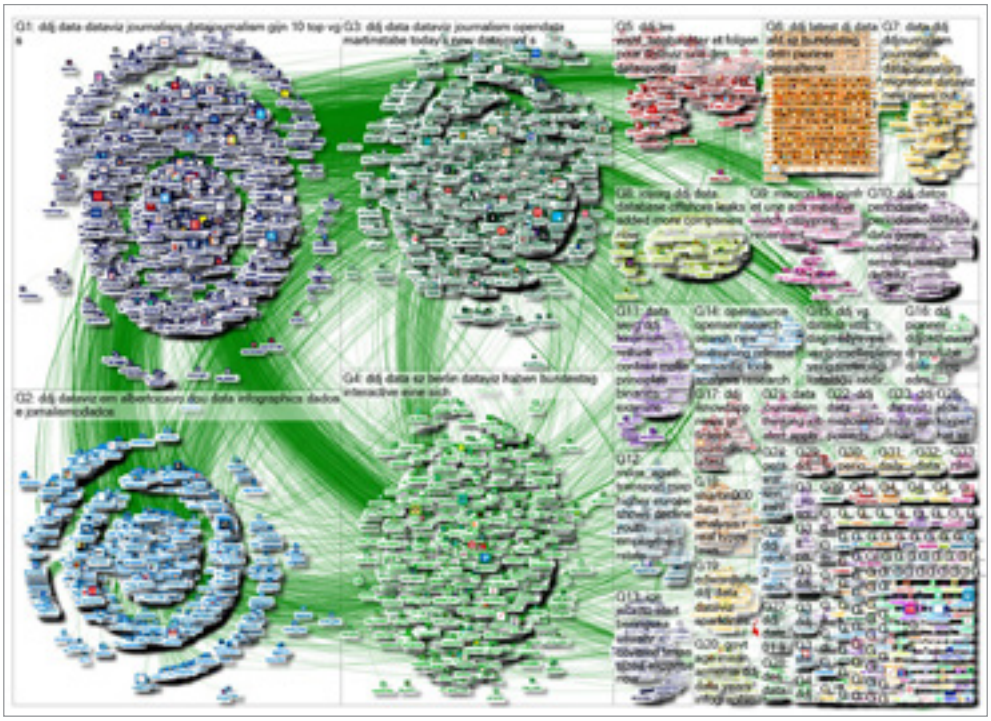
Ο ίδιος ο όρος, «δημοσιογραφία με γνώμονα τα δεδομένα», είναι αμφιλεγόμενος, καθώς κανείς θα μπορούσε να υποστηρίξει ότι η δημοσιογραφία δεν καθορίζεται από τα δεδομένα· τα δεδομένα έχουν απλώς ενημερωτικό χαρακτήρα ή είναι ένα εργαλείο που χρησιμοποιείται για τη δημοσιογραφία. Τα δεδομένα αποτελούνται από δομημένα γεγονότα και στατιστικά στοιχεία που απαιτούν από τους δημοσιογράφους να φιλτράρουν, να αναλύουν και να ανακαλύπτουν μοτίβα, προκειμένου να παράγουν ιστορίες. Ακριβώς όπως κανείς δεν θα αποκαλούσε ένα προφίλ «δημοσιογραφία με γνώμονα τη συνέντευξη» ή ένα άρθρο που βασίζεται σε δημόσια έγγραφα «δημοσιογραφία με γνώμονα τα έγγραφα», οι μεγάλες ιστορίες δημοσιογραφίας δεδομένων χρησιμοποιούν τα δεδομένα ως μόνο ένα από τα πολλά συστατικά τους.

Ο ρόλος του #ddj

Πέρα από αυτές τις εκτιμήσεις, η ευρεία χρήση του hashtag #ddj ανάμεσα στις κοινότητες της δημοσιογραφίας δεδομένων το έχει καταστήσει μια εξέχουσα πηγή για την ανταλλαγή θεμάτων και δραστηριοτήτων σε ολόκληρο τον κόσμο. Οι δημοσιογράφοι δεδομένων χρησιμοποιούν το hashtag, για να προωθήσουν το έργο τους και να το προβάλουν σε ευρύτερα διεθνή ακροατήρια.

Το hashtag βοηθά, επίσης, στη διευκόλυνση των συζητήσεων στα μέσα κοινωνικής δικτύωσης, όπου τα μέλη της κοινότητας της δημοσιογραφίας δεδομένων μπορούν να αναζητούν, να ανακαλύπτουν και να μοιράζονται περιεχόμενο με τη χρήση του hashtag. Οι συζητήσεις που περιλαμβάνουν το hashtag #ddj αφορούν ένα εύρος θεμάτων –από εκλογικές προβλέψεις και παρερμηνείες των διαγραμμάτων πιθανοτήτων μέχρι την ηθική των δεδομένων και τη λογοδοσία της τεχνητής νοημοσύνης.

191. https://gijn.org/



Εικόνα 7.1.1: Χαρτογράφηση του hashtag #ddj στο Twitter, από την 1η Ιανουαρίου 2018 έως τις 13 Αυγούστου 2018. Πηγή: NodeXL.

Η γέννηση του top 10 του #ddj

Η εβδομαδιαία σειρά Top 10 #ddj του GIJN ξεκίνησε τον Ιανουάριο του 2014, όταν ένας από εμάς δημοσίευσε για πρώτη φορά στο Twitter ένα διάγραμμα δικτύου για το #ddj (Smith, 2014). Το διάγραμμα, το οποίο απεικόνιζε τα tweet που ανέφεραν το hashtag #ddj, συμπεριλαμβανομένων των απαντήσεων σε αυτά τα tweet, δημιουργήθηκε με τη χρήση του NodeXL, ενός πακέτου ανάλυσης και οπτικοποίησης κοινωνικών δικτύων που βασίζεται στο λογισμικό υπολογιστικών φύλλων Excel. Τέτοια διαγράμματα δικτύου αποκαλύπτουν τα μοτίβα διασύνδεσης που προκύπτουν από δραστηριότητες όπως η απάντηση, το mention και το retweet. Τα μοτίβα αυτά αναδεικνύουν τα βασικά πρόσωπα, ομάδες και θέματα που συζητούνται.

Ως διεθνής οργανισμός ερευνητικής δημοσιογραφίας, το GIJN αναζητά πάντα τρόπους για να ευαισθητοποιήσει το κοινό σχετικά με το τι συμβαίνει στους τομείς της ερευνητικής δημοσιογραφίας και της δημοσιογραφίας δεδομένων. Όταν ο διευθυντής του GIJN, Ντέιβιντ Κάπλαν, είδε το διάγραμμα δικτύου του Σμιθ, πρότεινε να χρησιμοποιηθεί ο «χάρτης» για την παραγωγή ενός εβδομαδιαίου Top 10 #ddj, προκειμένου να παρουσιάζονται δημοφιλή και ενδιαφέροντα παραδείγματα δημοσιογραφίας δεδομένων. (Ο ίδιος και ο Σμιθ επιχείρησαν, επίσης, μια εβδομαδιαία ανασκόπηση της ερευνητικής δημοσιογραφίας, αλλά κανένα άλλο hashtag δεν έφτανε στο επίπεδο του #ddj για τη δημοσιογραφία δεδομένων.) Παρόλο που το GIJN ακολουθεί πιστά τα προτεινόμενα ευρήματα του διαγράμματος δικτύου, απαιτείται κάποια ανθρώπινη επιμέλεια για να εξαλειφθούν τα διπλότυπα και να αναδειχθούν τα πιο ενδιαφέροντα στοιχεία.

Από την έναρξη της σειράς, έχουμε συγκεντρώσει περισσότερα από 250 στιγμιότυπα από τις συζητήσεις της κοινότητας της δημοσιογραφίας δεδομένων με το hashtag #ddj τα τελευταία

έξι χρόνια (GIJN, άνευ ημερομηνίας). Η σειρά χρησιμεύει τώρα ως μια καλή, γρήγορη σύνοψη για τους ενδιαφερόμενους που δεν μπορούν να παρακολουθήσουν κάθε tweet του #ddj. Η χρήση του όρου «στιγμιότυπο» δεν είναι απλώς μια μεταφορά. Αυτή η ανάλυση μας δίνει μια εικόνα της κοινότητας της δημοσιογραφίας δεδομένων στο Twitter, με τον ίδιο τρόπο που η φωτοειδησεογραφία απεικονίζει τα πραγματικά πλήθη στα πρωτοσέλιδα των μεγάλων ειδησεογραφικών οργανισμών.

Η εξέλιξη του #ddj στη ροή του Twitter

Για να καταλάβουμε πώς έχει εξελιχθεί η ροή του Twitter με τη χρήση του #ddj, κάναμε μια πολύ βασική και πρόχειρη ανάλυση των δεδομένων με το #ddj που συλλέξαμε από το 2014 έως το 2019. Επιλέξαμε ένα μικρό δείγμα οκτώ εβδομάδων, τον Φεβρουάριο και τον Μάρτιο, από καθένα από τα έξι έτη, ήτοι 48 εβδομάδες. Υπήρχε μια ποικιλία περιεχομένου που κοινοποιήθηκε και απασχόλησε, και τα πιο δημοφιλή στοιχεία περιλάμβαναν αναλύσεις και σκέψεις, βραβεία, επιχορηγήσεις, εκδηλώσεις, μαθήματα, θέσεις εργασίας, εργαλεία, πόρους και έρευνες. Το είδος του περιεχομένου που κοινοποιείται παρέμεινε σταθερό με την πάροδο των ετών.

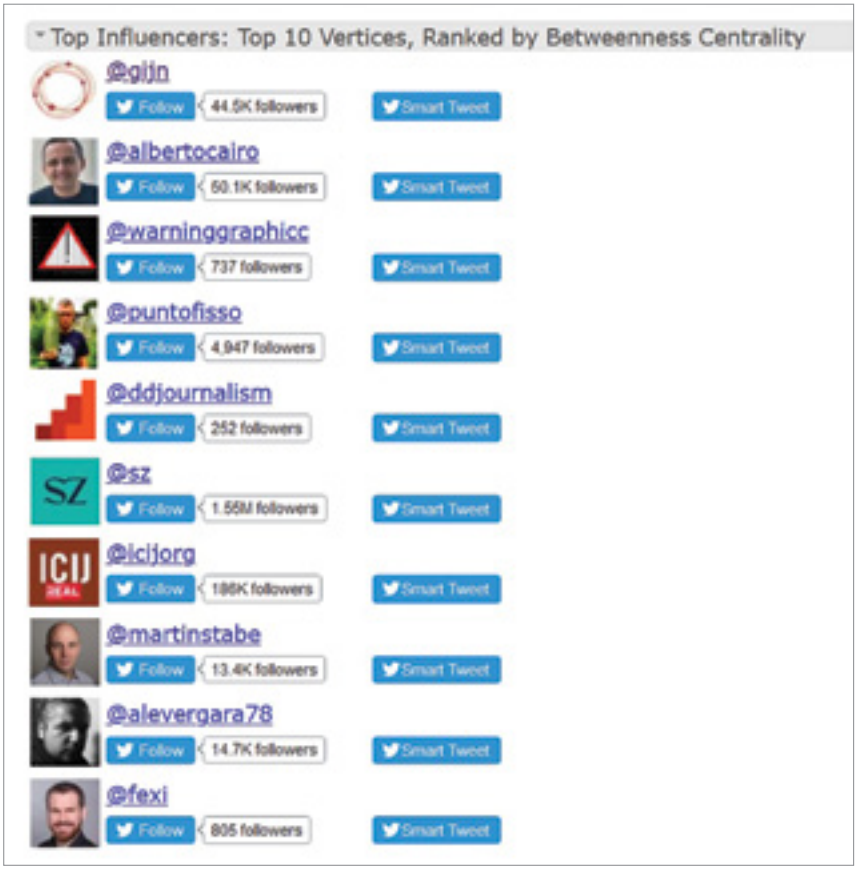
Το 2014, είδαμε άρθρα που μιλούσαν για έναν εκκολαπτόμενο τομέα της δημοσιογραφίας δεδομένων. Αυτό περιλάμβανε άρθρα που υποστήριζαν ότι η δημοσιογραφία δεδομένων είναι απαραίτητη, επειδή ενισχύει τη λογοδοσία και την κατανόηση (Howard, 2014) και προέβλεπαν ότι η ανάλυση δεδομένων είναι το μέλλον για τους δημοσιογράφους (Arthur, 2010). Τα επόμενα χρόνια, παρατηρήσαμε να συζητούνται νέα θέματα, όπως η τεχνητή νοημοσύνη, οι μαζικές διαρροές δεδομένων και οι συνεργατικές έρευνες δεδομένων. Υπήρχαν, επίσης, εμπεριστατωμένα άρθρα με οδηγίες, όπου οι δημοσιογράφοι δεδομένων άρχισαν να προσφέρουν πληροφορίες σχετικά με τις διαδικασίες δημοσιογραφίας δεδομένων που ακολουθούσαν (Grossenbacher, 2019) και να μοιράζονται τον καλύτερο τρόπο αξιοποίησης των βάσεων δεδομένων (Gallego, 2018), αντί να συζητούν το κατά πόσο ο κλάδος των μέσων ενημέρωσης θα πρέπει να ενσωματώσει τη δημοσιογραφία δεδομένων στις αίθουσες σύνταξης. Παρατηρήσαμε, επίσης, ότι ανάμεσα στις έρευνες που κοινοποιήθηκαν υπήρχαν συχνά αναλύσεις για τις εκλογές, τη μετανάστευση, τη ρύπανση, το κλίμα και το ποδόσφαιρο.

Η εβδομαδιαία ανασκόπηση του GIJN για το #ddj όχι μόνο αναδεικνύει τα πιο δημοφιλή tweet και διευθύνσεις URL, αλλά και απαριθμεί τους κεντρικούς συμμετέχοντες στη συζήτηση για το #ddj. Ορισμένοι από τους συνήθεις ύποπτους στο επίκεντρο των συζητήσεων του #ddj περιλαμβάνουν τους ειδικούς στη δημοσιογραφία δεδομένων Έντουαρντ Τάφτ, Αλμπέρτο Κάιρο, Μάρτιν Στέιμπ, Νέιτ Σίλβερ και Νέιθαν Γιο, μαζί με ομάδες δεδομένων από την Ευρώπη και τη Βόρεια Αμερική, συμπεριλαμβανομένων εκείνων των *Telegramme*, *Tages-Anzeiger*, *Berliner Morgenpost*, *FiveThirtyEight*, *Financial Times*, και του Upshot των *New York Times*. Το έργο τους κατά καιρούς μπορεί να είναι επιμορφωτικό και εμπνευσμένο, αλλά και να πυροδοτήσει περαιτέρω συζητήσεις. Η κοινότητα της δημοσιογραφίας δεδομένων μπορεί, επίσης, να επωφεληθεί και να δικτυωθεί με αυτούς τους influencer.

Μια σειρά από άλλα hashtag συνοδεύουν συχνά το #ddj, όπως αποκαλύπτει η χαρτογράφηση του Connected Action, επιτρέποντας στα μέλη της κοινότητας να αναζητούν παρόμοιες ιστορίες.

Μακράν, τα πιο συνηθισμένα hashtag που εμφανίζονται μαζί με το #ddj είναι τα #dataviz, #visualization, #datajournalism, #opendata, #data και #infographics. Αυτό μας υποδεικνύει ότι όσοι ασχολούνται με αυτόν τον τομέα ενδιαφέρονται ιδιαίτερα όχι μόνο για τη διαθεσιμότητα των δημόσιων δεδομένων, αλλά και για τον τρόπο με τον οποίο τα δεδομένα παρουσιάζονται δημιουργικά και οπτικοποιούνται για τους αναγνώστες.

Ωστόσο, η χαρτογράφηση του NodeXL για το #ddj δεν είναι σε καμία περίπτωση αντιπροσωπευτική ολόκληρου του πεδίου, καθώς αναλύει μόνο άτομα που τουιτάρουν. Επιπλέον, εκείνοι που έχουν γενικά περισσότερους ακολούθους στο Twitter και συγκεντρώνουν περισσότερα retweet τείνουν να εμφανίζονται σε πιο περίοπτη θέση στην ανασκόπηση.



Εινόνα 7.1.2: Παράδειγμα κορυφαίων influencer (από την 1η Ιανουαρίου 2018 έως τις 13 Αυγούστου 2018). Πηγή: NodeXL.

Παρατηρήσαμε, επίσης, ότι η πλειονότητα των κορυφαίων tweet προέρχεται συνήθως από την Ευρώπη και την Αμερική, ιδίως τη Γερμανία και τις Ηνωμένες Πολιτείες, ενώ υπάρχουν και μερικά tweet από την Ασία και την Αφρική. Αυτό θα μπορούσε να οφείλεται στην ασύμμετρη βάση των χρηστών του Twitter, στο γεγονός ότι άλλες περιοχές έχουν σχετικά λιγότερο ισχυρές κοινότητες δημοσιογράφων δεδομένων ή στο γεγονός ότι οι κοινότητες δημοσιογράφων δεδομένων σε άλλες περιοχές δεν οργανώνονται μέσω των ίδιων hashtag στο Twitter ή δεν οργανώνονται καθόλου στο Twitter.

Κατά τη διάρκεια του περασμένου έτους, παρατηρήσαμε ότι ορισμένες εργασίες από διακεκριμένους οργανισμούς δημοσιογραφίας δεδομένων που κοινοποιήθηκαν ευρέως στο Twitter δεν εμφανίζονταν στο διάγραμμα δικτύου μας. Αυτό θα μπορούσε ενδεχομένως να οφείλεται στο γεγονός ότι οι χρήστες δεν χρησιμοποίησαν το hashtag #ddj, όταν ανέβαζαν την ιστορία τους στο Twitter, ή χρησιμοποίησαν άλλα ή καθόλου hashtag. Υποψιαζόμαστε ότι η επέκταση του αριθμού χαρακτήρων στο Twitter από 140 σε 280 τον Νοέμβριο του 2017 μπορεί, επίσης, να βοήθησε τους χρήστες να επιλέξουν πιο μακροσκελή hashtag όπως το #datajournalism.

Εικόνα 7.1.3: Παραδείγματα κορυφαίων σχετιζόμενων hashtag (από την 1η Ιανουαρίου 2018 έως τις 13 Αυγούστου 2018). Πηγή: NodeXL.

Διασκεδαστικές ανακαλύψεις για το #ddj

Ενώ τα ευρήματά μας αποτελούν συχνά δείγματα δυναμικής δημοσιογραφίας και παρουσιάζουν όμορφες οπτικοποιήσεις, μερικές φορές είναι και απλά αστεία. Εν είδει επιλόγου, θα αναφερθούμε εν συντομία σε μερικά από τα πιο διασκεδαστικά στοιχεία που ανακαλύψαμε χρησιμοποιώντας το hashtag #ddj κατά το περασμένο έτος.

Μια αξιαγάπητη και έξυπνη οπτική μελέτη (Xaquín, 2017) παρουσίασε τους όρους με τους οποίους κάνουν αναζήτηση οι άνθρωποι σε διαφορετικές χώρες, όταν θέλουν να επισκευάσουν κάτι. Σε πολλές θερμότερες χώρες, αναζητούν πληροφορίες για ψυγεία, οι Βορειοαμερικάνοι και Ανατολικοασιάτες για τουαλέτες, ενώ οι άνθρωποι στη βόρεια και ανατολική Ευρώπη φαίνεται να χρειάζονται οδηγίες για το πώς αλλάζουν λάμπες. Στη συνέχεια, ένα διάγραμμα, που βρέθηκε στη συλλογή Sally L. Steinberg Collection of Doughnut Ephemera του Smithsonian, υποστηρίζει ότι το μέγεθος της τρύπας των ντόνατ συρρικνώθηκε σταδιακά με την πάροδο των ετών (Edwards, 2018). Σε ένα διαφορετικό άρθρο, ο γραφίστας Νάιτζελ Χολμς εικονογράφησε και εξήγησε περίεργα υπέροχους διαγωνισμούς σε όλο τον κόσμο, από αγώνες σαλιγκαριών μέχρι το κουβάλημα συζύγων, σε ένα βιβλίο με τίτλο *Crazy Competitions* (Yau, 2018).

Σε ένα άλλο κομμάτι της συλλογής μας, οι γυναίκες παγκοσμίως γνωρίζουν ήδη ότι οι τσέπες στα γυναικεία τζιν είναι πρακτικά μικροσκοπικές, και το *Pudding* έχει προσφέρει τα αδιαμφισβήτητα δεδομένα και την ανάλυση που το αποδεικνύουν (Diehm & Thomas, 2018). Τέλος, υπάρχει πραγματικά περίοδος αιχμής της τεκνοποίησης; Μια ανάλυση από το Visme, με βάση τα στοιχεία του ΟΗΕ για τις γεννήσεις ζώντων παιδιών, φαίνεται να το υποστηρίζει. Διαπίστωσαν συσχέτιση μεταξύ τριών διαφορετικών μεταβλητών: τους κορυφαίους μήνες γέννησης, τις εποχές του έτους και το γεωγραφικό πλάτος της χώρας (απόσταση από τον ισημερινό) που μπορεί να επηρεάζουν τους ρυθμούς αναπαραγωγής σε διαφορετικές χώρες (Chibana, 2017).

Παραπομπές

Arthur, C. (2010) «Analysing data is the future for journalists, says Tim Berners-Lee», *The Guardian*, 22 Νοεμβρίου. Διαθέσιμο σε: https://www.theguardian.com/media/2010/nov/22/data-analysis-tim-berners-lee (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Chibana, N. (2017) «Do humans have mating seasons? This heat map reveals the surprising link between birthdays and seasons», *Visual Learning Center0 by Visme*, 15 Αυγούστου. Διαθέσιμο σε: https://visme.co/blog/most-common-birthday/ (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Diehm, J., & Thomas, A. (2018) «Pockets», *The Pudding*, Αύγουστος. Διαθέσιμο σε: https://pudding.cool/2018/08/pockets/ (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Edwards, P. (2018) «Have donut holes gotten smaller? This compelling vintage chart says yes», *Vox*, 1 Ιουνίου. Διαθέσιμο σε: https://www.vox.com/2015/9/20/9353957/donut-hole-size-chart (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Gallego, C. S. (2018) «How to investigate companies found in the offshore leaks database», *ICIJ*, 23 Ιανουαρίου. Διαθέσιμο σε: https://www.icij.org/blog/2018/01/investigate-companies-found-offshore-leaks-database/ (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

GIJN (άνευ ημερομηνίας) «Top 10 in Data Journalism», *Global Investigative Journalism Network*. Διαθέσιμο σε: https://gijn.org/series/top-10-data-journalism-links/ (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Grossenbacher, T. (2019) «(Big) data journalism with Spark and R», *timogrossenbacher.ch*, 8 Μαρτίου. Διαθέσιμο σε: https://timogrossenbacher.ch/2019/03/big-data-journalism-with-spark-and-r/ (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Howard, A. (2014) «Data-driven journalism fuels accountability and insight in the 21st century», *TechRepublic*, 3 Μαρτίου. Διαθέσιμο σε: https://www.techrepublic.com/article/data-driven-journalism-fuels-accountability-and-insight-in-the-21st-century/ (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Smith, M. (2014) *#ddj NodeXL SNA* [Twitter], 22 Ιανουαρίου. Διαθέσιμο σε: https://twitter.com/marc_smith/status/425801408873385984 (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Xaquín G. V. (2017) «How to fix a toilet and other things we couldn't do without search», *How to Fix a Toilet*, 1 Σεπτεμβρίου. Διαθέσιμο σε: http://how-to-fix-a-toilet.com/ (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Yau, N. (2018) «Nigel Holmes new illustrated book on Crazy Competitions», *FlowingData*, 21 Μαΐου. Διαθέσιμο σε: https://flowingdata.com/2018/05/21/nigel-holmes-new-illustrated-book-on-crazy-competitions/ (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Οι συγγραφείς

Η Γιούνις Ω είναι συντονίστρια προγράμματος στο GIJN, όπου παρουσιάζει μια εβδομαδιαία ανασκόπηση του Top 10 #ddj με τα πιο δημοφιλή tweet για τη δημοσιογραφία δεδομένων.
Ο Μαρκ Σμιθ είναι κοινωνιολόγος που δημιουργεί και εφαρμόζει εργαλεία για τη μελέτη των μέσων κοινωνικής δικτύωσης.

7.2
Αρχειοθέτηση της Δημοσιογραφίας Δεδομένων

Meredith Broussard

Περίληψη: Αυτό το κεφάλαιο εξετάζει τις προκλήσεις της αρχειοθέτησης έργων δημοσιογραφίας δεδομένων και τα βήματα που μπορούν να ακολουθήσουν οι ομάδες δεδομένων για να διασφαλίσουν ότι τα έργα τους θα διατηρηθούν και στο μέλλον.

Λέξεις-κλειδιά: δημοσιογραφία δεδομένων, πρακτικές αρχειοθέτησης, αρχεία, ψηφιακά αρχεία, κατεστραμμένοι σύνδεσμοι, αρχειοθέτηση Ιστού

Στην πρώτη έκδοση του βιβλίου *The Data Journalism Handbook,* που δημοσιεύτηκε το 2012, ο πρωτοπόρος της δημοσιογραφίας δεδομένων Στιβ Ντόιγκ έγραψε ότι μία από τις αγαπημένες του ιστορίες δεδομένων ήταν το έργο *Murder Mysteries* του Τομ Χάργκροβ[192]. Στο έργο του, το οποίο δημοσιεύτηκε από το Scripps Howard News Service, ο Χάργκροβ εξέτασε λεπτομερή δημογραφικά δεδομένα σχετικά με 185.000 ανεξιχνίαστες δολοφονίες και δημιούργησε έναν αλγόριθμο, για να υποδείξει ποιες δολοφονίες μπορεί να συνδέονται μεταξύ τους. Οι αλληλένδετες δολοφονίες θα μπορούσαν να υποδηλώνουν έναν κατά συρροή δολοφόνο. «Αυτό το έργο τα έχει όλα», έγραψε ο Ντόιγκ: «σκληρή δουλειά, μια βάση δεδομένων καλύτερη από αυτή της κυβέρνησης, έξυπνη ανάλυση με τη χρήση τεχνικών των κοινωνικών επιστημών και διαδραστική παρουσίαση των δεδομένων στο Διαδίκτυο, ώστε οι αναγνώστες να μπορούν να τα εξερευνήσουν και μόνοι τους».

Μέχρι τη στιγμή της δεύτερης έκδοσης του τόμου *The Data Journalism Handbook*, έξι χρόνια αργότερα, η διεύθυνση URL του έργου ήταν πια «σπασμένη» (projects.scrippsnews.com/magazine/murder-mysteries). Το έργο είχε χαθεί από το Διαδίκτυο, επειδή ο εκδότης του είχε κλείσει: η Scripps Howard News Service είχε περάσει από πολλαπλές συγχωνεύσεις και αναδιαρθρώσεις και, τελικά, συγχωνεύθηκε με τον οργανισμό Gannett, εκδότη του δικτύου τοπικών ειδήσεων της USA Today.

Γνωρίζουμε ότι οι άνθρωποι αλλάζουν δουλειές και οι εταιρείες μέσων ενημέρωσης έρχονται και παρέρχονται. Ωστόσο, αυτό έχει καταστροφικές συνέπειες για τα έργα δημοσιογραφίας δεδομένων (για περισσότερα σχετικά με το θέμα αυτό, βλ. π.χ. Boss & Broussard, 2017· Broussard, 2014, 2015a, 2015b· Fisher & Klein, 2016).

Τα έργα δεδομένων είναι πιο εύθραυστα από τις «απλές» ιστορίες κειμένου και εικόνων που δημοσιεύονται στην έντυπη έκδοση μιας εφημερίδας ή ενός περιοδικού.

Συνήθως, η απώλεια των συνδέσμων δεν είναι σημαντικό πρόβλημα για τους αρχειοφύλακες· είναι εύκολο να χρησιμοποιήσει κανείς το Lexis-Nexis ή το ProQuest ή κάποιον άλλο πάροχο βάσεων δεδομένων, για να βρει ένα αντίγραφο όλων όσων έχουν δημοσιευτεί, ας πούμε, από την έντυπη έκδοση των *New York Times*, οποιαδήποτε ημέρα του 21ού αιώνα. Για τις ιστορίες δεδομένων, όμως, η απώλεια συνδέσμων υποδηλώνει ένα βαθύτερο πρόβλημα. Οι ιστορίες

192. http://www.murderdata.org/

δημοσιογραφίας δεδομένων δεν διατηρούνται σε παραδοσιακά αρχεία. Ως εκ τούτου, μπορούν να εξαφανιστούν από τον Ιστό. Εάν οι ειδησεογραφικοί οργανισμοί και οι βιβλιοθήκες δεν αναλάβουν δράση, οι ιστορικοί του μέλλοντος δεν θα είναι σε θέση να διαβάσουν όλες τις δημοσιεύσεις της Boston Globe σε οποιαδήποτε ημέρα του 2017. Αυτό έχει σοβαρές συνέπειες για τους μελετητές και για τη συλλογική μνήμη του κλάδου. Η δημοσιογραφία αναφέρεται συχνά ως το «πρώτο προσχέδιο της Ιστορίας». Εάν αυτό το πρώτο προσχέδιο είναι ημιτελές, πώς θα κατανοήσουν οι μελλοντικοί μελετητές τη σημερινή εποχή; Ή, εάν οι ιστορίες εξαφανιστούν από το Διαδίκτυο, πώς θα διατηρήσουν οι δημοσιογράφοι τα προσωπικά τους portfolio;

Αυτό είναι ένα ανθρώπινο πρόβλημα, όχι μόνο ένα υπολογιστικό πρόβλημα. Για να κατανοήσουμε καλύτερα τους λόγους για τους οποίους η δημοσιογραφία δεδομένων δεν αρχειοθετείται για τις επόμενες γενιές, μπορούμε να ξεκινήσουμε με το πώς αρχειοθετούνται οι «κανονικές» ειδήσεις. Όλοι οι ειδησεογραφικοί οργανισμοί χρησιμοποιούν ένα λογισμικό που ονομάζεται σύστημα διαχείρισης περιεχομένου («content management system», CMS), το οποίο επιτρέπει στον οργανισμό να προγραμματίζει και να διαχειρίζεται τα εκατοντάδες κομμάτια που δημιουργεί καθημερινά και, επίσης, εξασφαλίζει συνέπεια στην οπτική εμφάνιση και αισθητική κάθε κομματιού που δημοσιεύεται. Ιστορικά, οι παλαιότεροι ειδησεογραφικοί οργανισμοί χρησιμοποιούσαν άλλο CMS για την έντυπη έκδοση και άλλο για τη διαδικτυακή. Το διαδικτυακό CMS επιτρέπει στον ειδησεογραφικό οργανισμό να ενσωματώνει διαφημίσεις σε κάθε σελίδα, που είναι ένας από τους τρόπους με τους οποίους ο ειδησεογραφικός οργανισμός κερδίζει χρήματα. Το έντυπο CMS επιτρέπει στους σχεδιαστές των σελίδων προς εκτύπωση να διαχειρίζονται διαφορετικές εκδοχές διάταξης του εντύπου και, στη συνέχεια, να στέλνουν τις σελίδες στο τυπογραφείο για εκτύπωση και δέσιμο. Συνήθως, το βίντεο βρίσκεται σε διαφορετικό CMS. Τις αναρτήσεις στα μέσα κοινωνικής δικτύωσης μπορεί να τις διαχειρίζεται μια διαφορετική εφαρμογή, όπως το SocialFlow ή το Hootsuite –ή και όχι. Η τροφοδοσία αρχειακού υλικού προς το Lexis-Nexis και τους άλλους μεγάλους παρόχους τείνουν να συνδέονται με το έντυπο CMS. Εκτός εάν κάποιος στον ειδησεογραφικό οργανισμό θυμηθεί να συνδέσει και το διαδικτυακό CMS, οι ειδήσεις που κυκλοφορούν πρώτα σε ψηφιακή μορφή δεν περιλαμβάνονται στις ψηφιακές ενημερώσεις που λαμβάνουν οι βιβλιοθήκες και τα αρχεία. Αυτό είναι μια υπενθύμιση ότι η αρχειοθέτηση δεν είναι ουδέτερη, αλλά εξαρτάται από σκόπιμες ανθρώπινες επιλογές σχετικά με το τι έχει σημασία (και τι όχι) για το μέλλον.

Οι περισσότεροι άνθρωποι ρωτούν σε αυτό το σημείο «τι γίνεται με το Internet Archive;» Το Internet Archive είναι ένας θησαυρός και η ομάδα κάνει αξιοθαύμαστη δουλειά, καταγράφοντας στιγμιότυπα των ειδησεογραφικών ιστότοπων. Η τεχνολογία τους συγκαταλέγεται μεταξύ των πιο προηγμένων λογισμικών ψηφιακής αρχειοθέτησης. Ωστόσο, η προσέγγισή τους δεν καταγράφει τα πάντα. Το Internet Archive συλλέγει μόνο δημόσια διαθέσιμες ιστοσελίδες. Οι ειδησεογραφικοί οργανισμοί που απαιτούν είσοδο ή που περιλαμβάνουν paywall, ως μέρος της οικονομικής τους στρατηγικής, δεν μπορούν να διασωθούν αυτόματα στο Internet Archive. Οι ιστοσελίδες με στατικό περιεχόμενο ή απλές .html είναι ευκολότερο να διατηρηθούν. Οι σελίδες αυτές μπορούν εύκολα να καταγραφούν στο Internet Archive. Δυναμικό περιεχόμενο, όπως JavaScript ή οπτικοποίηση δεδομένων ή οτιδήποτε άλλο που κάποτε αναφερόταν ως «Web 2.0», είναι πολύ πιο δύσκολο να διατηρηθεί και δεν αποθηκεύεται συχνά στο Internet Archive. «Υπάρχουν πολλά διαφορετικά είδη δυναμικών σελίδων, μερικές από τις οποίες αποθηκεύονται εύκολα σε ένα αρχείο και μερικές από τις οποίες καταρρέουν εντελώς», αναφέρεται σε μια από τις συχνές ερωτήσεις του Internet Archive. «Όταν μια δυναμική σελίδα προβάλλεται σε τυπική .html, το αρχείο λειτουργεί άψογα. Όταν μια δυναμική σελίδα περιέχει φόρμες, JavaScript ή άλλα στοιχεία που απαιτούν αλληλεπίδραση με τον κεντρικό υπολογιστή προέλευσης, το αρχείο δεν θα περιέχει τη λειτουργικότητα του αρχικού ιστοτόπου».

Οι δυναμικές οπτικοποιήσεις δεδομένων και οι ειδησεογραφικές εφαρμογές, τα πιο σύγχρονα είδη ιστοριών δημοσιογραφίας δεδομένων, δεν μπορούν να καταγραφούν από την υπάρχου-

σα τεχνολογία αρχειοθέτησης στο Διαδίκτυο. Επίσης, για διάφορους θεσμικούς λόγους, αυτού του είδους οι ιστορίες τείνουν να αναπτύσσονται εκτός CMS. Έτσι, ακόμη και αν ήταν δυνατή η αρχειοθέτηση οπτικοποιήσεων δεδομένων και ειδησεογραφικών εφαρμογών (κάτι που γενικά δεν γίνεται με αυτήν την προσέγγιση), οποιαδήποτε αυτοματοποιημένη ροή δεν θα τις κατέγραφε, επειδή δεν βρίσκονται μέσα στο CMS. Είναι ένα περίπλοκο πρόβλημα. Δεν υπάρχουν εύκολες απαντήσεις. Συνεργάζομαι με μια ομάδα δημοσιογράφων δεδομένων, βιβλιοθηκονόμων και επιστημόνων πληροφορικής που προσπαθούν να αναπτύξουν τεχνολογία για την επίλυση αυτού του ακανθώδους προβλήματος. Δανειζόμαστε μεθόδους από την αναπαραγώγιμη επιστημονική έρευνα, για να διασφαλίσουμε ότι οι άνθρωποι θα μπορούν να διαβάζουν τις σημερινές ειδήσεις στους αυριανούς υπολογιστές. Προσαρμόζουμε ένα εργαλείο που ονομάζεται ReproZip, που συλλέγει τον κώδικα, τα δεδομένα και το περιβάλλον διακομιστή και το οποίο χρησιμοποιείται σε πειράματα υπολογιστικής επιστήμης.Πιστεύουμε ότι το ReproZip μπορεί να ενσωματωθεί με ένα εργαλείο όπως το Webrecorder.io για τη συλλογή και τη διατήρηση ειδησεογραφικών εφαρμογών, οι οποίες είναι τόσο ιστορίες όσο και λογισμικό. Επειδή τα έργα δημοσιογραφίας δεδομένων για το Διαδίκτυο και τις κινητές συσκευές εξαρτώνται από και υπάρχουν σε σχέση με ένα ευρύ φάσμα άλλων περιβαλλόντων μέσων, βιβλιοθηκών, χαρακτηριστικών του προγράμματος περιήγησης και οντοτήτων του Ιστού (τα οποία μπορεί, επίσης, να αλλάζουν συνεχώς), πιστεύουμε ότι θα μπορέσουμε να χρησιμοποιήσουμε το ReproZip, για να συλλέξουμε και να διατηρήσουμε τις απομακρυσμένες βιβλιοθήκες και τον κώδικα που επιτρέπουν τη λειτουργία σύνθετων αντικειμένων δημοσιογραφίας δεδομένων στον Ιστό. Θα χρειαστούν ακόμη ένα-δυο χρόνια, για να αποδειχθεί η υπόθεσή μας.

Εν τω μεταξύ, υπάρχουν μερικά συγκεκριμένα πράγματα που κάθε ομάδα δεδομένων μπορεί να κάνει, για να διασφαλίσει ότι η δημοσιογραφία δεδομένων της θα διατηρηθεί στο μέλλον.

Τραβήξτε ένα βίντεο. Αυτή η στρατηγική προέρχεται από τη διατήρηση των βιντεοπαιχνιδιών. Ακόμα και όταν η κονσόλα ενός βιντεοπαιχνιδιού δεν υπάρχει πια, υπάρχει ένα βίντεο που μπορεί να δείξει το παιχνίδι στο αρχικό του περιβάλλον. Το ίδιο ισχύει και για τις ιστορίες δημοσιογραφίας δεδομένων. Αποθηκεύστε το βίντεο σε μια κεντρική τοποθεσία με μεταδεδομένα απλού κειμένου που περιγράφουν τι δείχνει το βίντεο. Κάθε φορά που εμφανίζεται ένα νέο φορμάτ βίντεο (όπως όταν το VHS έδωσε τη θέση του στο DVD ή όταν το DVD αντικαταστάθηκε από το streaming video), αναβαθμίστε όλα τα βίντεο σε αυτό το νέο φορμάτ.

Φτιάξτε μια σμίκρυνση για τις επόμενες γενιές. Βιβλιοθήκες όπως η Django-bakery επιτρέπουν την απόδοση δυναμικών σελίδων ως στατικές σελίδες. Αυτό μερικές φορές ονομάζεται «baking out» («ψήσιμο»). Ακόμη και σε μια βάση δεδομένων με χιλιάδες εγγραφές, κάθε δυναμική εγγραφή μπορεί να «ψηθεί», δηλαδή να αποθηκευτεί, ως μια στατική σελίδα που απαιτεί ελάχιστη συντήρηση. Θεωρητικά, όλες αυτές οι στατικές σελίδες θα μπορούσαν να εισαχθούν στο σύστημα διαχείρισης περιεχομένου του οργανισμού. Το «baking out» δεν είναι απαραίτητο να γίνει από την αρχή. Ένα έργο δεδομένων μπορεί να ξεκινήσει ως δυναμικός ιστότοπος και, στη συνέχεια, να μετατραπεί σε στατικό ιστότοπο, αφού η επισκεψιμότητα μειωθεί λίγους μήνες αργότερα. Η γενική ιδέα είναι να προσαρμόσετε το έργο σας για συστήματα αρχειοθέτησης, φτιάχνοντας την απλούστερη δυνατή έκδοση και, στη συνέχεια, να βεβαιωθείτε ότι αυτή η απλή έκδοση βρίσκεται στην ίδια ψηφιακή τοποθεσία με όλες τις άλλες ιστορίες που δημοσιεύτηκαν την ίδια εποχή.

Σκεφτείτε το μέλλον. Οι δημοσιογράφοι τείνουν να σχεδιάζουν τη δημοσίευση και μετά να προχωρούν στο επόμενο θέμα. Αντ' αυτού, προσπαθήστε να σχεδιάσετε τα «γεράματα» των ιστοριών δεδομένων σας την ίδια στιγμή που σχεδιάζετε τη δημοσίευσή τους. Η ιστορία του Ματ Γουέιτ *Kill All Your Darlings* στο *Source*, το ιστολόγιο του OpenNews, είναι ένας εξαιρετικός οδηγός για το πώς να σκεφτείτε τον κύκλο ζωής μιας ιστορίας δημοσιογραφίας δεδομένων. Στο τέλος, θα πάρετε προαγωγή ή θα μετακινηθείτε σε έναν νέο οργανισμό. Αυτό που θέλετε είναι να επιβιώσουν τα έργα δημοσιογραφίας δεδομένων σας μετά την αποχώρησή σας.

Συνεργασία με βιβλιοθήκες, ιδρύματα μνήμης και επαγγελματικά αρχεία. Ως δημοσιογράφος, θα πρέπει οπωσδήποτε να κρατάτε αντίγραφα της δουλειάς σας. Ωστόσο, κανείς δεν πρόκειται να ψάξει σε ένα κουτί στην ντουλάπα σας ή στον σκληρό σας δίσκο, ή ακόμη και στην προσωπική σας ιστοσελίδα, όταν θα αναζητά δημοσιογραφικά κομμάτια στο μέλλον. Θα ψάξουν στο LexisNexis, στο ProQuest ή σε άλλα μεγάλα επαγγελματικά αποθετήρια. Για να μάθετε περισσότερα σχετικά με την επαγγελματική συντήρηση και την ψηφιακή αρχειοθέτηση, μπορείτε να διαβάσετε το βιβλίο των Καθλίν Χάνσεν και Νόρα Πολ *Future-Proofing the News: Preserving the First Draft of History* (Hansen & Paul, 2017) που είναι ο κλασικός οδηγός για την κατανόηση του τοπίου της αρχειοθέτησης ειδήσεων, καθώς και των τεχνολογικών, νομικών προκλήσεων αλλά και προκλήσεων που απορρέουν από τη δομή του οργανισμού για τη διατήρηση των ειδήσεων.

Παραπομπές

Boss, K., & Broussard, M. (2017) «Challenges of archiving and preserving born-digital news applications», *IFLA Journal*, 43(2), σσ. 150–157. DOI: https://doi.org/10.1177/0340035216686355

Broussard, M. (2014) «Future-proofing news apps», *MediaShift*, 23 Απριλίου. Διαθέσιμο σε: http://mediashift.org/2014/04/future-proofing-news-apps/ (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Broussard, M. (2015a) «Preserving news apps present huge challenges», *Newspaper Research Journal*, 36(3), σσ. 299–313. DOI: https://doi.org/10.1177/0739532915600742

Broussard, M. (2015b) «The irony of writing about digital preservation», *The Atlantic*, 20 Νοεμβρίου. Διαθέσιμο σε: https://www.theatlantic.com/technology/archive/2015/11/the-irony-of-writing-about-digital-preservation/416184/ (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Fisher, T., & Klein, S. (2016) «A conceptual model for interactive databases in news», *ProPublica/News App Model* [GitHub]. Διαθέσιμο σε: https://github.com/propublica/newsappmodel (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Hansen, K. A. & Paul, N. (2017). Future-Proofing the News: *Preserving the First Draft of History.* Λάνχαμ, Μεριλαντ: Rowman & Littlefield Publishers.

Η συγγραφέας

Η Μέρεντιθ Μπρουσάρντ είναι Αναπληρώτρια Καθηγήτρια στο Δημοσιογραφικό Ινστιτούτο Arthur L. Carter του Πανεπιστημίου της Νέας Υόρκης και συγγραφέας του *Artificial Unintelligence: How Computers Misunderstand the World* (MIT Press, 2018).

7.3

Από τον *Guardian* στο Google News Lab: Μια Δεκαετία Εργασίας στη Δημοσιογραφία Δεδομένων

Simon Rogers

Περίληψη: Μια προσωπική αφήγηση της τελευταίας δεκαετίας της δημοσιογραφίας δεδομένων υπό το πρίσμα μιας καταξιωμένης επαγγελματικής διαδρομής.

Λέξεις-κλειδιά: δημοσιογραφία δεδομένων, Datablog του *Guardian*, WikiLeaks, ανοιχτά δεδομένα, διαφάνεια, υπολογιστικά φύλλα

Όταν αποφάσισα ότι ήθελα να γίνω δημοσιογράφος, κάπου μεταξύ της πρώτης και της δεύτερης τάξης του δημοτικού σχολείου, δεν μου περνούσε ποτέ από το μυαλό ότι αυτό θα περιλάμβανε δεδομένα. Τώρα, δουλεύοντας με δεδομένα κάθε μέρα, συνειδητοποιώ πόσο τυχερός ήμουν. Σίγουρα δεν ήταν το αποτέλεσμα μιας προσεκτικά σχεδιασμένης επαγγελματικής διαδρομής. Απλά βρέθηκα στο σωστό μέρος την κατάλληλη στιγμή. Ο τρόπος με τον οποίο συνέβη αυτό λέει πολλά για την κατάσταση της δημοσιογραφίας δεδομένων το 2009. Πιστεύω ότι μας λέει, επίσης, πολλά για τη δημοσιογραφία δεδομένων το 2019.

Ο Άντριαν Χολοβάτι, ένας προγραμματιστής από το Σικάγο, ο οποίος είχε εργαστεί στη *Washington Post* και ίδρυσε το *EveryBlock*, έναν ιστότοπο ειδήσεων και συζητήσεων για τη γειτονιά, ήρθε να δώσει μια ομιλία στην αίθουσα σύνταξης στο Εκπαιδευτικό Κέντρο του *Guardian* στην οδό Φάρινγκτον στο Λονδίνο. Την εποχή εκείνη ήμουν συντάκτης ειδήσεων στην έντυπη εφημερίδα (τότε το κέντρο βάρους), έχοντας εργαστεί στο Διαδίκτυο και έχοντας αναλάβει την επιμέλεια της επιστημονικής ενότητας. Όσο περισσότερο ο Χολοβάτι μιλούσε για το πώς χρησιμοποιούσε τα δεδομένα, για να αφηγηθεί ιστορίες και να βοηθήσει τους ανθρώπους να κατανοήσουν τον κόσμο, τόσο περισσότερο κάτι ξυπνούσε μέσα μου. Όχι μόνο αυτό ήταν κάτι που θα μπορούσα να κάνω, αλλά στην πραγματικότητα ανταποκρινόταν σε αυτό που ήδη έκανα ολοένα και περισσότερο. Ίσως θα μπορούσα να είμαι δημοσιογράφος που δούλευε με δεδομένα. Ένας «δημοσιογράφος δεδομένων».

Η δουλειά μου ως συντάκτης ειδήσεων στο τμήμα γραφικών μού έδωσε την ευκαιρία να συνεργαστώ με σχεδιαστές που άλλαξαν τον τρόπο με τον οποίο βλέπω τον κόσμο, στην ταλαντούχα ομάδα του Μάικλ Ρόμπινσον. Και καθώς το χαρτοφυλάκιο των οπτικοποιήσεων μεγάλωνε, αποδείχτηκε ότι είχα συγκεντρώσει πολλούς αριθμούς: ο Ματ Μακάλιστερ, ο οποίος εγκαινίαζε το ανοιχτό API του *Guardian*, το περιέγραψε ως «χρυσορυχείο». Είχαμε καθαρίσει τα δεδομένα για το ΑΕΠ, τις εκπομπές διοξειδίου του άνθρακα, τις κυβερνητικές δαπάνες και πολλά άλλα, όλα αποθηκευμένα ως Google spreadsheet (υπολογιστικά φύλλα στο Google Drive) και έτοιμα για χρήση την επόμενη φορά που θα τα χρειαζόμασταν.

Τι θα γινόταν, εάν απλά δημοσιεύαμε αυτά τα δεδομένα σε φορμάτ ανοιχτών δεδομένων; Χωρίς αρχεία .pdf, μόνο ενδιαφέροντα, προσβάσιμα δεδομένα, έτοιμα προς χρήση από οποιονδήποτε. Και αυτό κάναμε με το Datablog του *Guardian* –αρχικά με 200 διαφορετικά σύνολα δεδομένων: ποσοστά εγκληματικότητας, οικονομικοί δείκτες, λεπτομέρειες για εμπόλεμες ζώ-

νες, ακόμη και για την εβδομάδα μόδας και για τους κακούς της σειράς Doctor Who. Αρχίσαμε να συνειδητοποιούμε ότι τα δεδομένα μπορούν να εφαρμοστούν σε όλα. Ωστόσο, κάτι τέτοιο παρέμενε παράξενο. Ο «συντάκτης δεδομένων» («data editor») δεν ήταν καθόλου μια ευρέως διαδεδομένη θέση εργασίας –πολύ λίγες αίθουσες σύνταξης είχαν οποιαδήποτε ομάδα δεδομένων. Στην πραγματικότητα, και μόνο η χρήση της λέξης «δεδομένα» σε μια σύσκεψη θα προκαλούσε καγχασμούς. Αυτό δεν ήταν «σωστή» δημοσιογραφία, σωστά;

Το 2009, όμως, ήταν η αρχή της επανάστασης των ανοιχτών δεδομένων: ο κόμβος κυβερνητικών δεδομένων των ΗΠΑ, data.gov, είχε ξεκινήσει τον Μάιο του ίδιου έτους, με μόλις 47 σύνολα δεδομένων. Χώρες και πόλεις σε ολόκληρο τον κόσμο εγκαινίαζαν πύλες ανοιχτών δεδομένων και οι ακτιβιστές ζητούσαν πρόσβαση σε ολοένα και περισσότερα. Μέσα σε έναν χρόνο, βάλαμε τους αναγνώστες μας να μας βοηθήσουν να συγκεντρώσουμε τα έξοδα χιλιάδων βουλευτών. Την ίδια χρονική περίοδο, η κυβέρνηση του Ηνωμένου Βασιλείου δημοσίευσε το τελικό σύνολο δεδομένων για τις δαπάνες της: το COINS (Combined Online Information System –«Συνδυασμένο Διαδικτυακό Σύστημα Πληροφοριών»). Η ομάδα του *Guardian* είχε δημιουργήσει ένα διαδραστικό σύστημα εξερεύνησης, για να ενθαρρύνει τους αναγνώστες να βοηθήσουν στην εξερεύνησή του[193]. Ωστόσο, μόλις προέκυψαν άρθρα από αυτά τα δεδομένα, το ερώτημα έγινε «πώς μπορούμε να έχουμε περισσότερα από αυτά;».

Δεν χρειάστηκε να περιμένουμε πολύ. Η απάντηση ήρθε από έναν νεοσύστατο τότε οργανισμό, με έδρα τη Σουηδία, ο οποίος θα μπορούσε να χαρακτηριστεί ως ριζοσπαστική ατζέντα διαφάνειας: τα WikiLeaks. Ό,τι κι αν πιστεύετε σήμερα για τα WikiLeaks, ο αντίκτυπος του οργανισμού στην πρόσφατη ιστορία της δημοσιογραφίας δεδομένων δεν μπορεί να παραγνωριστεί. Ξαφνικά, υπήρξε μια μαζική διάθεση χιλιάδων λεπτομερών αρχείων από τις εμπόλεμες ζώνες του Αφγανιστάν πρώτα, και στη συνέχεια του Ιράκ. Ήρθε με τη μορφή ενός γιγαντιαίου υπολογιστικού φύλλου, πολύ μεγάλου για να το χειριστεί αρχικά η ομάδα ερευνών του *Guardian*.

Ήταν μεγαλύτερο από τα *Pentagon Papers*, εκείνη τη δημοσιοποίηση αρχείων κατά τη διάρκεια του πολέμου του Βιετνάμ που έριξε φως στο πώς πραγματικά εξελισσόταν η σύγκρουση. Τα αρχεία ήταν, επίσης, εκτενή –συμπεριλάμβαναν κατάλογο περιστατικών με αριθμό θυμάτων, γεωγραφικές τοποθεσίες, λεπτομέρειες και κατηγορίες. Μπορούσαμε να δούμε την αύξηση των επιθέσεων με αυτοσχέδιους εκρηκτικούς μηχανισμούς στο Ιράκ, για παράδειγμα, και πόσο επικίνδυνοι είχαν γίνει οι δρόμοι σε όλη τη χώρα. Και όταν αυτά τα δεδομένα συνδυάστηκαν με τις παραδοσιακές δεξιότητες των έμπειρων πολεμικών ρεπόρτερ, τα δεδομένα άλλαξαν τον τρόπο με τον οποίο ο κόσμος έβλεπε τον πόλεμο.

Δεν ήταν δύσκολο να παραχθεί περιεχόμενο που έμοιαζε να έχει αντίκτυπο σε ολόκληρο τον κόσμο. Τα γεωδεδομένα των υπολογιστικών φύλλων προσφέρονται για χαρτογράφηση, για παράδειγμα, και υπήρχε ένα νέο δωρεάν εργαλείο που μπορούσε να βοηθήσει σε αυτό: το Google Fusion Tables. Έτσι, δημιουργήσαμε έναν γρήγορο χάρτη με κάθε περιστατικό στο Ιράκ στο οποίο υπήρξε τουλάχιστον ένας θάνατος. Μέσα σε 24 ώρες, ένα κομμάτι που χρειάστηκε μία ώρα για να φτιαχτεί προβλήθηκε σε όλο τον κόσμο, καθώς οι χρήστες μπορούσαν να εξερευνήσουν οι ίδιοι την εμπόλεμη ζώνη με τρόπο που την έκανε να μοιάζει πιο αληθινή. Και επειδή τα δεδομένα ήταν δομημένα, οι ομάδες γραφικών μπορούσαν να παράγουν εξελιγμένα, πλούσια οπτικά στοιχεία που παρείχαν πιο εμπεριστατωμένο ρεπορτάζ.

Και μέχρι το τέλος του 2011 –έναν χρόνο πριν από την πρώτη έκδοση αυτού του βιβλίου– το έργο *Reading the Riots* («Διαβάζοντας τις αναταραχές») είχε εφαρμόσει τις τεχνικές του υποβοηθούμενου από υπολογιστή ρεπορτάζ («computer-assisted reporting») του Φιλ Μάγιερ, από τη δεκαετία του 1960, σε ένα ξέσπασμα βίας σε όλη την Αγγλία (Robertson, 2011). Ο Μάγιερ είχε εφαρμόσει τεχνικές των κοινωνικών επιστημών για την κάλυψη των ταραχών στο Ντιτρόιτ στα τέλη της δεκαετίας

193. https://www.theguardian.com/politics/coins-combined-online-information-system

του 1960. Μια ομάδα με επικεφαλής τον Πολ Λιούις του *Guardian* έκανε το ίδιο με το ξέσπασμα των αναταραχών σε όλη την Αγγλία εκείνη τη χρονιά και ενσωμάτωσε τα δεδομένα ως βασικό μέρος αυτού του έργου. Αυτά ήταν θέματα βασισμένα σε δεδομένα που δημοσιεύονταν στο πρωτοσέλιδο.

Υπήρχε, όμως, και μια άλλη μεταβολή στον τρόπο που καταναλώνουμε πληροφορίες, η οποία εξελισσόταν γρήγορα. Δεν θυμάμαι να έχω ακούσει τη λέξη «viral» να χρησιμοποιείται με άλλο νόημα, πλην θεμάτων υγείας, πριν από το 2010. Δεν ισχύει το ίδιο σήμερα και η άνοδος της δημοσιογραφίας δεδομένων συνέπεσε, επίσης, με την άνοδο των μέσων κοινωνικής δικτύωσης. Χρησιμοποιούσαμε tweet, για να πουλήσουμε ιστορίες σε χρήστες σε ολόκληρο τον κόσμο και η επακόλουθη επισκεψιμότητα οδήγησε σε περισσότερους χρήστες που αναζητούσαν τέτοιου είδους ιστορίες βασισμένες σε δεδομένα. Μια οπτικοποίηση ή ένας αριθμός μπορεί να γίνει αντιληπτός μέσα σε δευτερόλεπτα από χιλιάδες ανθρώπους. Τα μέσα κοινωνικής δικτύωσης μεταμόρφωσαν τη δημοσιογραφία, αλλά η ενίσχυση της δημοσιογραφίας δεδομένων ήταν η εξέλιξη που την έκανε, από εξειδικευμένη, mainstream.

Πρώτον, άλλαξε τη δυναμική στις σχέσεις με τους καταναλωτές. Στο παρελθόν, τα λόγια ενός δημοσιογράφου θεωρούνταν ιερά και απαραβίαστα –τώρα είσαι απλώς μια φωνή ανάμεσα σε εκατομμύρια. Εάν κάνεις ένα λάθος με ένα σύνολο δεδομένων, 500 άνθρωποι είναι έτοιμοι να σου το πουν. Θυμάμαι να κάνω μακροσκελείς (και έντονες) συζητήσεις στο Twitter με ντιζάινερ γύρω από τους χρωματικούς συνδυασμούς των χαρτών και να αλλάζω αυτό που έκανα εξαιτίας των συζητήσεων αυτών. Ο διαμοιρασμός έκανε τη δουλειά μου καλύτερη.

Στην πραγματικότητα, αυτό το πνεύμα συνεργασίας είναι κάτι που εξακολουθεί να υφίσταται και σήμερα στη δημοσιογραφία δεδομένων. Εξάλλου, η πρώτη έκδοση αυτού του βιβλίου αναπτύχθηκε αρχικά από μια ομάδα ανθρώπων που συναντήθηκαν στο Mozilla Festival στο Λονδίνο –και καθώς διάφορες εκδηλώσεις γύρω από τα δεδομένα άρχισαν να ξεφυτρώνουν, το ίδιο συνέβη και με τις ευκαιρίες για τους δημοσιογράφους δεδομένων να συνεργαστούν και να μοιραστούν δεξιότητες. Εάν οι δημοσιεύσεις για το Ιράκ και τα WikiLeaks ήταν σπουδαία αρχικά παραδείγματα διατλαντικής συνεργασίας, τότε δείτε πώς αυτές οι ασκήσεις εξελίχθηκαν σε παγκόσμια ρεπορτάζ με τη συμμετοχή εκατοντάδων δημοσιογράφων. Οι διαρροές του Σνόουντεν και τα *Panama Papers* ήταν αξιοσημείωτες για τον τρόπο με τον οποίο οι δημοσιογράφοι συντονίστηκαν σε όλο τον κόσμο, για να μοιραστούν τις ιστορίες τους και να αξιοποιήσουν ο ένας το έργο του άλλου[194].

Πάρτε για παράδειγμα ένα εγχείρημα όπως το *Electionland*, το οποίο χρησιμοποίησε τεχνικές συνεργατικού ρεπορτάζ για την παρακολούθηση των προβλημάτων στην ψηφοφορία, σε πραγματικό χρόνο, την ημέρα των εκλογών. Συμμετείχα κι εγώ, παρέχοντας δεδομένα της Google και βοηθώντας στην οπτικοποίηση αυτών των ανησυχιών, σε πραγματικό χρόνο. Μέχρι σήμερα, το *Electionland* είναι η μεγαλύτερη άσκηση ρεπορτάζ μέσα σε μία ημέρα στην Ιστορία, με τη συμμετοχή άνω των 1.000 δημοσιογράφων. Υπάρχει άμεση σύνδεση αυτού που κάναμε εκείνα τα πρώτα χρόνια με το *Electionland*.

Σκοπός μου δεν είναι να απαριθμήσω έργα, αλλά να αναδείξω το ευρύτερο πλαίσιο εκείνων των προηγούμενων ετών, όχι μόνο στον *Guardian*, αλλά σε αίθουσες σύνταξης σε ολόκληρο τον κόσμο. Οι *New York Times*, οι *Los Angeles Times*, η *La Nación* στην Αργεντινή: σε ολόκληρο τον κόσμο οι δημοσιογράφοι ανακάλυπταν νέους τρόπους εργασίας, αφηγούμενοι ιστορίες που βασίζονται σε δεδομένα με καινοτόμους τρόπους. Αυτό ήταν το υπόβαθρο της πρώτης έκδοσης αυτού του βιβλίου. Η *La Nación* στην Αργεντινή είναι ένα καλό παράδειγμα. Μια μικρή ομάδα ενθουσιωδών δημοσιογράφων έμαθε πώς να οπτικοποιεί με το Tableau (ένα νέο εργαλείο τότε) και το συνδύασε με ρεπορτάζ για την ελευθερία της πληροφόρησης, ώστε να ξεκινήσει ένας κόσμος δημοσιογραφίας δεδομένων στη Λατινική και Νότια Αμερική.

194. Για περισσότερες πληροφορίες σχετικά με τις συνεργασίες μεγάλης κλίμακας γύρω από τα *Panama Papers*, βλ. το κεφάλαιο των Ντιάζ-Στρακ, Γκαλιέγο και Ρομέρα σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 4.2, «Οικοδόμηση Συνεργασιών για τα Panama και Paradise Papers»).

Η δημοσιογραφία δεδομένων μετατράπηκε από υπόθεση μερικών μοναχικών ανθρώπων σε καθιερωμένο κομμάτι πολλών μεγάλων newsroom, αλλά μια τάση ήταν σαφής ακόμη και τότε: κάθε φορά που εισάγεται μια νέα τεχνική στο ρεπορτάζ, τα δεδομένα όχι μόνο αποτελούν βασικό μέρος της, αλλά και οι δημοσιογράφοι δεδομένων είναι ακριβώς εκεί, εν τω μέσω της διαδικασίας. Μέσα σε λιγότερο από τρία χρόνια, ο πληθοπορισμός («crowdsourcing») έγινε ένα καθιερωμένο εργαλείο στην αίθουσα σύνταξης και οι δημοσιογράφοι βρήκαν δεδομένα, χρησιμοποίησαν βάσεις δεδομένων για τη διαχείριση τεράστιων όγκων εγγράφων, δημοσίευσαν σύνολα δεδομένων και εφάρμοσαν τεχνικές ανάλυσης με βάση τα δεδομένα σε σύνθετες ειδήσεις.

Αυτό δεν πρέπει να θεωρηθεί μια μεμονωμένη εξέλιξη στον τομέα της δημοσιογραφίας. Ήταν το αποτέλεσμα των τεράστιων εξελίξεων στη διεθνή διαφάνεια πέραν της δημιουργίας πυλών ανοιχτών δεδομένων. Αυτές περιλάμβαναν εκστρατείες όπως εκείνες που διεξάγονται από το Free Our Data, το Open Knowledge Foundation και από ομάδες τεχνολογίας των πολιτών («civic tech groups»), για την αύξηση της πίεσης προς την κυβέρνηση του Ηνωμένου Βασιλείου να ανοίξει ενημερωτικά σύνολα δεδομένων για δημόσια χρήση και να παρέχει API[195] για να τα εξερευνήσει ο καθένας. Περιλάμβαναν, επίσης, αυξημένη πρόσβαση σε ισχυρά, δωρεάν εργαλεία οπτικοποίησης και καθαρισμού δεδομένων, όπως τα OpenRefine, Google Fusion Tables, Many Eyes, Datawrapper, Tableau Public και άλλα. Αυτά τα δωρεάν εργαλεία, σε συνδυασμό με την πρόσβαση σε πολλά δωρεάν δημόσια δεδομένα, διευκόλυναν την παραγωγή όλο και περισσότερων οπτικοποιήσεων και έργων δεδομένων για το κοινό. Αίθουσες σύνταξης, όπως το *Texas Tribune* και η *ProPublica*, άρχισαν να αναπτύσσουν δραστηριότητες γύρω από αυτά τα δεδομένα.

Βλέπετε πώς λειτουργεί; Ένας ενάρετος κύκλος δεδομένων, εύκολης επεξεργασίας, οπτικοποίησης δεδομένων, περισσότερων δεδομένων και ούτω καθεξής. Όσο περισσότερα δεδομένα υπάρχουν, όσο περισσότερη δουλειά γίνεται με τα δεδομένα, τόσο μεγαλύτερη είναι η πίεση για τη δημοσιοποίηση περισσότερων δεδομένων. Όταν έγραψα το άρθρο *Data Journalism Is the New Punk* («Η δημοσιογραφία των δεδομένων είναι η νέα πανκ»), το νόημα ήταν το εξής: βρισκόμασταν σε ένα σημείο όπου η δημιουργικότητα μπορούσε πραγματικά να αφεθεί ελεύθερη (Rogers, 2012), αλλά και εκεί όπου το έργο θα γινόταν, τελικά, mainstream.

Τα δεδομένα δεν μπορούν να κάνουν τα πάντα. Όπως έγραψε ο Τζόναθαν Γκρέι: «Το σημερινό κύμα ενθουσιασμού για τα δεδομένα, τις τεχνολογίες δεδομένων και όλα τα πράγματα που βασίζονται στα δεδομένα μπορεί να οδηγήσει κάποιον να υποψιαστεί ότι αυτά τα μηχαναγνώσιμα, δομημένα πράγματα είναι μια ειδική περίπτωση» (Gray, 2012). Είναι μόνο ένα από τα κομμάτια του παζλ των αποδεικτικών στοιχείων που πρέπει να συναρμολογήσουν οι δημοσιογράφοι. Καθώς, όμως, παρέχονται ολοένα και περισσότερα διαθέσιμα δεδομένα, ο ρόλος αυτός αλλάζει και γίνεται ακόμη πιο σημαντικός. Η δυνατότητα πρόσβασης και ανάλυσης τεράστιων συνόλων δεδομένων ήταν το κύριο κίνητρο για την επόμενη επαγγελματική μου κίνηση.

Το 2013, μου δόθηκε η ευκαιρία να μετακομίσω στην Καλιφόρνια και να ενταχθώ στο Twitter ως ο πρώτος συντάκτης δεδομένων –και ήταν ξεκάθαρο ότι τα δεδομένα είχαν μπει στο λεξιλόγιο του κυρίαρχου εκδοτικού χώρου, σίγουρα στις Ηνωμένες Πολιτείες και στην Ευρώπη. Ορισμένοι ιστότοποι δημοσιογραφίας δεδομένων εμφανίστηκαν μέσα σε λίγες εβδομάδες, όπως το Upshot των *New York Times* και το *FiveThirtyEight* του Νέιτ Σίλβερ. Το κοινό εκεί έξω, στον κόσμο, γινόταν όλο και περισσότερο οπτικά εγγράμματο και εκτιμούσε τις εξελιγμένες οπτικοποιήσεις πολύπλοκων θεμάτων. Θα ρωτήσετε ποια είναι τα στοιχεία που έχω ότι ο κόσμος αισθάνεται άνετα με τις οπτικοποιήσεις δεδομένων; Δεν έχω πολλές πληροφορίες πέρα από την εμπειρία μου ότι η παραγωγή μιας οπτικοποίησης που θα προκαλέσει μεγάλες αντιδράσεις στο Διαδίκτυο είναι πιο δύσκολη από ό,τι παλαιότερα. Εκεί που πιο παλιά όλοι αντιδρούσαμε με «ωωω και αααα» στις οπτικοποιήσεις, τώρα δύσκολα θα αντιδράσουμε με κάτι παραπάνω από ένα ανασήκωμα των ώμων.

195. ΣτΜ: «Application Programming Interface» (διεπαφή προγραμματισμού εφαρμογών).

Όταν εντάχθηκα στο Google News Lab, για να ασχοληθώ με τη δημοσιογραφία δεδομένων το 2015, ήταν σαφές ότι ο τομέας είχε πρόσβαση σε σημαντικότερα και μεγαλύτερα σύνολα δεδομένων από ποτέ άλλοτε. Κάθε μέρα πραγματοποιούνται δισεκατομμύρια αναζητήσεις, ένα σημαντικό ποσοστό των οποίων για πρώτη φορά. Και όλο και περισσότερο οι δημοσιογράφοι παίρνουν αυτά τα δεδομένα και τα αναλύουν, μαζί με τα tweet και τα like στο Facebook[196]. Πρόκειται για την εξάντληση της σύγχρονης ζωής, που αντιστρέφεται και επιστρέφει ως γνώση για τον τρόπο με τον οποίο ζούμε σήμερα.

Η δημοσιογραφία δεδομένων είναι πλέον πιο διαδεδομένη από ποτέ. Το 2016, τα Data Journalism Awards έλαβαν 471 συμμετοχές –αριθμός ρεκόρ ως τότε. Όμως, στα βραβεία του 2018 έλαβαν σχεδόν 700 συμμετοχές –περισσότερες από τις μισές ήταν από μικρές αίθουσες σύνταξης και πολλές από όλο τον κόσμο. Και αυτές οι συμμετοχές γίνονται ολοένα και πιο καινοτόμες. Η τεχνητή νοημοσύνη («artificial intelligence») ή η μηχανική μάθηση («machine learning») έχει γίνει εργαλείο για τη δημοσιογραφία δεδομένων, όπως αποδεικνύεται από το έργο του Πίτερ Άλντους στο *Buzzfeed* (Aldhous, 2017).

Εν τω μεταξύ, η πρόσβαση σε νέες τεχνολογίες, όπως η εικονική («virtual reality») και η επαυξημένη πραγματικότητα («augmented reality»), προσφέρει δυνατότητες για την αφήγηση ιστοριών με δεδομένα με νέους τρόπους. Ως κάποιος που η δουλειά του είναι να φαντάζεται πώς θα μπορούσε να αλλάξει η δημοσιογραφία δεδομένων –και τι μπορούμε να κάνουμε για να την υποστηρίξουμε–, εξετάζω τον τρόπο με τον οποίο οι αναδυόμενες τεχνολογίες μπορούν να διευκολύνουν περισσότερους δημοσιογράφους, ώστε να τις ενσωματώσουν στη δουλειά τους. Για παράδειγμα, πρόσφατα συνεργαστήκαμε με το στούντιο ντιζάιν Datavized για την κατασκευή του TwoTone, ενός οπτικού εργαλείου για τη μετάφραση δεδομένων σε ήχο[197].

Τι κάνει ένας δημοσιογράφος δεδομένων στην Google; Έχω τη δυνατότητα να αφηγούμαι ιστορίες με μια μεγάλη και πλούσια συλλογή από σύνολα δεδομένων, καθώς και να συνεργάζομαι με ταλαντούχους ντιζάινερ, για να φανταστώ το μέλλον της οπτικοποίησης των ενημερωτικών δεδομένων και τον ρόλο των νέων τεχνολογιών στη δημοσιογραφία. Μέρος του ρόλου μου είναι να συμβάλλω στη διερεύνηση του τρόπου με τον οποίο οι νέες τεχνολογίες μπορούν να συνδυαστούν με τις σωστές περιπτώσεις χρήσης και περιστάσεις στις οποίες είναι κατάλληλες και αξιοποιήσιμες.

Ο ρόλος αυτός περιλαμβάνει, επίσης, τη διερεύνηση του τρόπου με τον οποίο οι δημοσιογράφοι χρησιμοποιούν τα δεδομένα και τις ψηφιακές τεχνολογίες, για να αφηγηθούν ιστορίες με νέους τρόπους. Για παράδειγμα, ένα πρόσφατο έργο, το Zones of Silence της *El Universal*, έδειξε τη χρήση της τεχνητής νοημοσύνης στη δημοσιογραφία, χρησιμοποιώντας γλωσσική επεξεργασία για την ανάλυση της ειδησεογραφικής κάλυψης των δολοφονιών των καρτέλ ναρκωτικών και τη σύγκρισή τους με τα επίσημα δεδομένα, ενώ το χάσμα ανάμεσα στα δύο ήταν οι περιοχές όπου η δημοσιογραφία σώπαινε. Τους βοήθησα παρέχοντάς τους πρόσβαση σε API τεχνητής νοημοσύνης και σχεδιαστικούς πόρους.

Οι προκλήσεις είναι μεγάλες, για όλους μας. Όλοι μας καταναλώνουμε πληροφορίες με ολοένα και πιο κινητούς τρόπους, γεγονός που επιφέρει τις δικές του προκλήσεις. Οι πολύπλοκες οπτικοποιήσεις πλήρους οθόνης έρχονται σε σύγκρουση με το γεγονός ότι περισσότεροι από τους μισούς από εμάς διαβάζουμε πλέον τις ειδήσεις στο τηλέφωνο ή σε άλλες κινητές συσκευές –το ένα τρίτο από εμάς διαβάζουμε τις ειδήσεις στην τουαλέτα, σύμφωνα με μελέτη του *Reuters* για την κατανάλωση ειδήσεων (Newman κ.α., 2017). Αυτό σημαίνει ότι ολοένα και περισσότερο οι ντιζάινερ σε αίθουσες σύνταξης πρέπει να σχεδιάζουν για μικροσκοπικές οθόνες και για όλο και μικρότερη διάρκεια προσοχής.

196. Για περισσότερες προσεγγίσεις πάνω στο θέμα, δείτε την ενότητα 6, «Διερευνώντας τα Δεδομένα, τις Πλατφόρμες και τους Αλγορίθμους».
197. https://twotone.io/

Έχουμε, επίσης, ένα νέο πρόβλημα που μπορεί να μας εμποδίσει να μάθουμε από το παρελθόν. Ο κώδικας πεθαίνει, οι βιβλιοθήκες σαπίζουν και, τελικά, μεγάλο μέρος της πιο φιλόδοξης δημοσιογραφικής δουλειάς απλά χάνεται. Τα έξοδα των βουλευτών του *Guardian*, το *EveryBlock* και άλλα έργα έχουν όλα υποκύψει στην εξαφάνιση της θεσμικής μνήμης. Αυτό το πρόβλημα της εξαφάνισης της δημοσιογραφίας δεδομένων ήδη αποτελεί αντικείμενο ορισμένων καινοτόμων προσεγγίσεων (όπως μπορείτε να δείτε στο κεφάλαιο της Μπρουσάρντ[198] σε αυτόν τον τόμο). Μακροπρόθεσμα, αυτό απαιτεί κατάλληλες επενδύσεις και μένει να δούμε αν η κοινότητα έχει επαρκή κίνητρα για να το πραγματοποιήσει.

Μάλιστα, αντιμετωπίζουμε ένα ευρύτερο και όλο και πιο ανησυχητικό ζήτημα: εκείνο της εμπιστοσύνης. Η ανάλυση δεδομένων ήταν πάντα αντικείμενο ερμηνείας και διαφωνίας, αλλά η καλή δημοσιογραφία δεδομένων μπορεί να το ξεπεράσει αυτό. Σε μια εποχή που η πίστη στις ειδήσεις και σε ένα κοινό σύνολο γεγονότων αμφισβητείται καθημερινά, η δημοσιογραφία δεδομένων μπορεί να μας φωτίσει τον δρόμο, φέρνοντας στο φως γεγονότα και αποδείξεις με προσιτό τρόπο.

Έτσι, παρά τις οποιεσδήποτε αλλαγές, ορισμένα πράγματα παραμένουν σταθερά σε αυτόν τον τομέα. Η δημοσιογραφία δεδομένων έχει μακρά ιστορία[199], αλλά το 2009 η δημοσιογραφία δεδομένων έμοιαζε να είναι ένας σημαντικός τρόπος για να φτάσουμε σε μια κοινή αλήθεια, κάτι που θα μπορούσαμε όλοι να υποστηρίξουμε. Τώρα αυτή η ανάγκη είναι μεγαλύτερη από ποτέ.

Παραπομπές

Aldhous, P. (2017) «We trained a computer to search for hidden spy planes. This is what it found», *BuzzFeed News*, 8 Αυγούστου. Διαθέσιμο σε: https://www.buzzfeednews.com/article/peteraldhous/hidden-spy-planes (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Gray, J. (2012) «What data can and cannot do», *The Guardian*, 31 Μαΐου. Διαθέσιμο σε: https://www.theguardian.com/news/datablog/2012/may/31/data-journalism-focused-critical (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A. L., & Nielsen, R. K. (2017). Digital News Report 2017. Οξφόρδη: *Reuters Institute for the Study of Journalism*. Διαθέσιμο σε: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital%20News%20Report%202017%20web_0.pdf (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Robertson, C. (2011) «Reading the riots: How the 1967 Detroit riots were investigated», *The Guardian*, 9 Δεκεμβρίου. Διαθέσιμο σε: https://www.theguardian.com/uk/video/2011/dec/09/reading-the-riots-detroit-meyer-video (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Rogers, S. (2012) «Anyone can do it. Data journalism is the new punk», *The Guardian*, 24 Μαΐου. Διαθέσιμο σε: https://www.theguardian.com/news/datablog/2012/may/24/data-journalism-punk (Προσπελάστηκε: 12 Δεκεμβρίου 2021).

Ο συγγραφέας

Ο Σάιμον Ρότζερς είναι συντάκτης δεδομένων στην ομάδα του Google News Lab, με έδρα το Σαν Φρανσίσκο, είναι διευθυντής των Sigma Data Journalism Awards και διδάσκει δημοσιογραφία δεδομένων στη Σχολή Medill του Πανεπιστημίου Northwestern.

198. ΣτΕ: Κεφάλαιο 7.2, «Αρχειοθέτηση της Δημοσιογραφίας Δεδομένων»
199. Βλ., για παράδειγμα, τα κεφάλαια των Άντερσον και Κοέν σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 9.1, «Γενεαλογίες της Δημοσιογραφίας Δεδομένων» και Κεφάλαιο 5.1, «Τρόποι Άσκησης της Δημοσιογραφίας Δεδομένων» αντίστοιχα).

7.4
Οι Δεσμοί της Δημοσιογραφίας Δεδομένων με την Τεχνολογία των Πολιτών

Stefan Baack

Περίληψη: Ποια είναι τα κοινά στοιχεία που μοιράζεται η δημοσιογραφία δεδομένων με άλλες μορφές εργασίας με δεδομένα και κουλτούρας δεδομένων.

Λέξεις-κλειδιά: τεχνολογία των πολιτών, έλεγχος πρόσβασης, επαγγελματικά όρια, δημοσιογραφία δεδομένων, ελευθερία πληροφόρησης, βάσεις δεδομένων

Ενώ το υποβοηθούμενο από υπολογιστή ρεπορτάζ («computer-assisted reporting») θεωρούνταν στο παρελθόν ως μια πρακτική που αφορούσε αποκλειστικά τους (ερευνητές) δημοσιογράφους, η δημοσιογραφία δεδομένων χαρακτηρίζεται από την εμπλοκή της με τον τομέα της τεχνολογίας και άλλες μορφές εργασίας με δεδομένα και κουλτούρας δεδομένων. Σε σύγκριση με το computer-assisted reporting, η ανάδυση της δημοσιογραφίας δεδομένων στις Ηνωμένες Πολιτείες και στην Ευρώπη συνέπεσε με διάφορες εξελίξεις τόσο εντός όσο και εκτός των αιθουσών σύνταξης. Αυτές περιλαμβάνουν: την αυξανόμενη διαθεσιμότητα των δεδομένων στο Διαδίκτυο, κυρίως λόγω των πρωτοβουλιών για τα ανοιχτά δεδομένα και των διαρροών· τις αίθουσες σύνταξης που προσλαμβάνουν προγραμματιστές και τους εντάσσουν στη συντακτική ομάδα για την καλύτερη διαχείριση των δεδομένων και τη δημιουργία διαδραστικών διαδικτυακών εφαρμογών· και την εμφάνιση διαφόρων κινημάτων «τεχνολογίας για το καλό» που έλκονται από τη δημοσιογραφία ως έναν τρόπο για να χρησιμοποιήσουν τις τεχνολογικές τους δεξιότητες για ένα «δημόσιο καλό». Αυτό συνέβαλε στο να μπουν οι τεχνολόγοι στις αίθουσες σύνταξης από όταν εμφανίστηκε η δημοσιογραφία δεδομένων και έγινε δημοφιλής τη δεκαετία του 2000 στη Δύση και αλλού. Ωστόσο, η εμπλοκή που προκύπτει ανάμεσα στους δημοσιογράφους δεδομένων και άλλες μορφές εργασίας με δεδομένα είναι ξεχωριστή σε κάθε διαφορετική περιοχή. Επιπλέον, η δημοσιογραφία δεδομένων συνδέεται με νέες, επιχειρηματικές μορφές δημοσιογραφίας που έχουν προκύψει ως απάντηση στον συνεχή αγώνα των οργανισμών μέσων ενημέρωσης να αναπτύξουν βιώσιμα επιχειρηματικά μοντέλα. Αυτοί οι νέοι τύποι ενημερωτικών οργανισμών (για παράδειγμα, μη κερδοσκοπικές αίθουσες σύνταξης, όπως η *ProPublica*, ή νεοσύστατες ενημερωτικές εταιρείες που χρηματοδοτούνται από επιχειρηματικά κεφάλαια, όπως το *BuzzFeed*) τείνουν να αμφισβητούν τα παραδοσιακά όρια της δημοσιογραφίας στη φιλοδοξία τους να «αναβιώσουν» ή να «βελτιώσουν» τη δημοσιογραφία, με την τεχνολογία και τα δεδομένα να παίζουν συχνά βασικό ρόλο σε αυτές τις προσπάθειες (Usher, 2017· Wagemans κ.α., 2019).

Η εμπλοκή της δημοσιογραφίας δεδομένων με άλλες μορφές εργασίας με δεδομένα και κουλτούρες δεδομένων δημιουργεί νέες εξαρτήσεις αλλά και νέες συνέργειες που επιτρέπουν νέες μορφές διατομεακής συνεργασίας. Εδώ, θέλω να φέρω ως παράδειγμα τη στενή σχέση ανάμεσα στη δημοσιογραφία δεδομένων και στην τεχνολογία των πολιτών («civic tech»), επειδή σε πολλά μέρη και τα δύο φαινόμενα εμφανίστηκαν περίπου την ίδια εποχή και, από πολύ νωρίς, διαμόρφωσαν αμοιβαία το ένα το άλλο. Η τεχνολογία των πολιτών αφορά την ανάπτυξη εργαλείων που αποσκοπούν στην ενδυνάμωση των πολιτών, διευκολύνοντάς τους να ασχοληθούν με τις κυβερνήσεις τους ή να τις υποχρεώσουν να λογοδοτήσουν. Παραδείγματα εγχειρημά-

των της τεχνολογίας των πολιτών είναι το OpenParliament, ένας δικτυακός τόπος κοινοβουλευτικής εποπτείας που, μεταξύ άλλων, καθιστά τις κοινοβουλευτικές ομιλίες πιο προσιτές· το WhatDoTheyKnow, ένας δικτυακός τόπος για την ελευθερία της πληροφόρησης που βοηθά τους χρήστες να υποβάλλουν και να βρίσκουν αιτήματα για την ελευθερία της πληροφόρησης· και το FixMyStreet, το οποίο απλοποιεί την αναφορά προβλημάτων στις τοπικές Αρχές[200].

Οι τεχνολόγοι των πολιτών και οι δημοσιογράφοι δεδομένων μοιράζονται ορισμένα σημαντικά χαρακτηριστικά. Πρώτον, πολλοί επαγγελματίες και από τις δύο ομάδες είναι προσηλωμένοι στις αρχές της κουλτούρας του ανοιχτού κώδικα και προωθούν τον διαμοιρασμό, τη χρήση εργαλείων ανοιχτού κώδικα και προτύπων δεδομένων. Δεύτερον, οι δημοσιογράφοι δεδομένων και οι τεχνολόγοι των πολιτών βασίζονται σε μεγάλο βαθμό σε δεδομένα, είτε αυτά προέρχονται από επίσημους φορείς, είτε μέσω πληθοπορισμού («crowdsourcing»), είτε από άλλες πηγές. Τρίτον, αν και διαφέρουν ως προς τα μέσα τους, και οι δύο ομάδες φιλοδοξούν να παρέχουν δημόσιες υπηρεσίες που ενδυναμώνουν τους πολίτες και καθιστούν τις Αρχές υπόλογες. Λόγω αυτού του αλληλεπικαλυπτόμενου συνόλου δεξιοτήτων σχετικά με τα δεδομένα, των συμπληρωματικών φιλοδοξιών και της κοινής δέσμευσης για διαμοιρασμό, οι τεχνολόγοι των πολιτών και οι δημοσιογράφοι δεδομένων αντιλαμβάνονται εύκολα οι μεν τους δε ως συμπληρωματικούς. Επιπλέον, η υποστήριξη από ενημερωτικούς οργανισμούς, ιδρύματα όπως το Knight Foundation και πρωτοβουλίες από τα κάτω, όπως το Hacks/Hackers, έχουν δημιουργήσει συνεχείς ανταλλαγές και συνεργασίες μεταξύ δημοσιογράφων δεδομένων και τεχνολόγων των πολιτών.

Η ένταση μεταξύ της επέκτασης και της ενίσχυσης του δημοσιογραφικού «πυρήνα»

Με βάση μια περιπτωσιολογική μελέτη στη Γερμανία και το Ηνωμένο Βασίλειο, η οποία εξέτασε τον τρόπο με τον οποίο οι δημοσιογράφοι δεδομένων και οι τεχνολόγοι των πολιτών αλληλοσυμπληρώνονται, μπορούμε να περιγράψουμε τη μεταξύ τους εμπλοκή ως περιστρεφόμενη γύρω από δύο βασικές πρακτικές: διευκόλυνση και έλεγχος πρόσβασης (Baack, 2018). Διευκόλυνση σημαίνει να δίνεται η δυνατότητα σε άλλους να αναλάβουν οι ίδιοι δράση, ενώ ο έλεγχος πρόσβασης («gatekeeping») παραπέμπει στο παραδοσιακό δημοσιογραφικό μοντέλο του φύλακα των πληροφοριών που αφορούν το κοινό. Για να καταδειχθεί η διαφορά, οι ιστότοποι κοινοβουλευτικής εποπτείας που αναπτύχθηκαν από τους τεχνολόγους των πολιτών έχουν ως στόχο να επιτρέπουν στους χρήστες τους να ενημερώνονται, για παράδειγμα, αναζητώντας κοινοβουλευτικές ομιλίες (διευκόλυνση), αλλά όχι να τους προωθούν προκαταβολικά πληροφορίες που κρίνονται σχετικές από επαγγελματίες («gatekeeping»). Η διευκόλυνση έχει να κάνει με την ενδυνάμωση του ατόμου, ενώ το gatekeeping έχει να κάνει με την καθοδήγηση της δημόσιας συζήτησης και τον αντίκτυπο. Αυτό που χαρακτηρίζει την εμπλοκή μεταξύ των δημοσιογράφων δεδομένων και των τεχνολόγων των πολιτών είναι ότι οι πρακτικές διευκόλυνσης και φύλαξης είναι συμπληρωματικές και μπορούν να αλληλοενισχύονται. Για παράδειγμα, οι εφαρμογές τεχνολογίας των πολιτών δεν διευκολύνουν μόνο τους απλούς πολίτες· οι δημοσιογράφοι δεδομένων μπορούν να τις χρησιμοποιήσουν για τις δικές τους έρευνες. Από την άλλη πλευρά, οι έρευνες των δημοσιογράφων μπορούν να τραβήξουν την προσοχή σε συγκεκριμένα ζητήματα και να ενθαρρύνουν τους πολίτες να κάνουν χρήση των υπηρεσιών διευκόλυνσης. Επιπλέον, τα δικαιώματα πληροφόρησης είναι απαραίτητα τόσο για τις πρακτικές διευκόλυνσης όσο και για τις πρακτικές φύλαξης, γεγονός που δημιουργεί πρόσθετες συνέργειες. Για παράδειγμα, οι δημοσιογράφοι δεδομένων μπορούν να χρησιμοποιήσουν τα αποκλειστικά δικαιώματά τους, για να αποκτήσουν δεδομένα τα οποία στη συνέχεια μοιράζονται με τους τεχνολόγους των πολιτών, ενώ οι δημοσιογράφοι μπορούν να επωφεληθούν

200. https://www.openingparliament.org/, https://www.whatdotheyknow.com/, https://www.fixmystreet.com/

από τη συνηγορία της τεχνολογίας των πολιτών για ισχυρότερα δικαιώματα ελευθερίας της πληροφόρησης και πολιτικές ανοιχτών δεδομένων.

Οι νέες καινοτόμες μορφές δημοσιογραφίας διαδραματίζουν ιδιαίτερο ρόλο στη σχέση μεταξύ της δημοσιογραφίας δεδομένων και της τεχνολογίας των πολιτών, καθώς είναι πιο ανοιχτές προς την επέκταση του παραδοσιακού gatekeeping με την έννοια της διευκόλυνσης, όπως την εννοεί η τεχνολογία των πολιτών. Για παράδειγμα, η *ProPublica* έχει αναπτύξει αρκετές μεγάλες βάσεις δεδομένων με δυνατότητα αναζήτησης, οι οποίες αποσκοπούν όχι μόνο στη διευκόλυνση της ενασχόλησης των απλών πολιτών με τις κυβερνήσεις τους, αλλά και στην ενίσχυση των δημοσιογραφικών ερευνών που εκπονούνται από δημοσιογράφους σε τοπικές αίθουσες σύνταξης οι οποίοι δεν διαθέτουν τους πόρους και την τεχνογνωσία για να συλλέξουν, να καθαρίσουν και να αναλύσουν οι ίδιοι τα δεδομένα. Ένα άλλο μη κερδοσκοπικό newsroom από τη Γερμανία, το *Correctiv*, ακολούθησε παρόμοια προσέγγιση και ενσωμάτωσε τον ιστότοπο για την ελευθερία της πληροφόρησης του Ιδρύματος Open Knowledge Foundation στη Γερμανία σε ορισμένες από τις εφαρμογές του. Η ενσωμάτωση αυτή επέτρεψε στους χρήστες να ζητούν απευθείας περαιτέρω πληροφορίες, οι οποίες προστίθενται αυτόματα στη βάση δεδομένων του *Correctiv* μόλις αποκτηθούν[201].

Ενώ τα παραδείγματα αυτά καταδεικνύουν ότι υπάρχει ένας αυξανόμενος αριθμός οργανισμών που διευρύνουν τις παραδοσιακές έννοιες της δημοσιογραφίας, ενσωματώνοντας πρακτικές και αξίες από άλλες κουλτούρες δεδομένων, υπάρχει επίσης και το αντίθετο: δημοσιογράφοι δεδομένων που αντιδρούν στις ομοιομορφίες των πρακτικών και των φιλοδοξιών με άλλα πεδία εργασίας δεδομένων, αγκαλιάζοντας την επαγγελματική τους ταυτότητα ως φύλακες και αφηγητές. Αυτοί οι δημοσιογράφοι δεν απορρίπτουν απαραιτήτως την τεχνολογία των πολιτών, αλλά η απάντησή τους είναι μια μεγαλύτερη εξειδίκευση της δημοσιογραφίας, πιο κοντά στις έννοιες της παραδοσιακής, ερευνητικής δημοσιογραφίας.

Οι ευκαιρίες των ρευστών ορίων

Εν ολίγοις, η εμπλοκή της δημοσιογραφίας δεδομένων με άλλα πεδία εργασίας με δεδομένα και κουλτούρας δεδομένων συμβάλλει στην αυξημένη διαφοροποίηση του τρόπου με τον οποίο γίνεται κατανοητή και ασκείται η «δημοσιογραφία», είτε προς την κατεύθυνση της επέκτασης είτε προς την κατεύθυνση της ενίσχυσης των παραδοσιακών αξιών και ταυτοτήτων. Τόσο οι ίδιοι οι δημοσιογράφοι όσο και οι ερευνητές μπορούν να θεωρήσουν τη δημοσιογραφία δεδομένων ένα φαινόμενο ενταγμένο σε ευρύτερους τεχνολογικούς, πολιτιστικούς και οικονομικούς μετασχηματισμούς. Σε αυτό το κεφάλαιο, επικεντρώθηκα στην εμπλοκή μεταξύ των δημοσιογράφων δεδομένων και των τεχνολόγων των πολιτών, αλλά θα ήθελα να επισημάνω δύο βασικά συμπεράσματα για τους δημοσιογράφους δεδομένων που είναι σημαντικά πέρα από τη συγκεκριμένη περίπτωση.

Επωφεληθείτε από τα ρευστά όρια. Οι δημοσιογράφοι τείνουν να περιγράφουν την έλλειψη επαγγελματικών ορίων έναντι άλλων τομέων ως προβληματική, αλλά οι συνέργειες μεταξύ των τεχνολόγων των πολιτών και των δημοσιογράφων δεδομένων καταδεικνύουν ότι τα ρευστά όρια μπορούν, επίσης, να λειτουργήσουν ως πλεονέκτημα. Αντί να τις αντιλαμβάνονται πρωταρχικά ως προβληματικές, οι δημοσιογράφοι δεδομένων πρέπει, επίσης, να αναρωτηθούν εάν υπάρχουν συνέργειες με άλλους τομείς εργασίας δεδομένων και πώς μπορούν να επωφεληθούν καλύτερα από αυτές. Αυτό δεν σημαίνει ότι οι δημοσιογράφοι πρέπει απαραίτητα να υιοθετήσουν πρακτικές διευκόλυνσης των ίδιων. Αν και υπάρχουν τέτοια παραδείγματα, οι δημοσιογράφοι που απορρίπτουν αυτήν την ιδέα μπορούν ακόμη να προσπαθήσουν να βρουν τρόπους να επωφεληθούν, χωρίς να θυσιάσουν την επαγγελματική τους ταυτότητα.

201. https://correctiv.org/en/

Αγκαλιάστε τη διαφορετικότητα στην επαγγελματική δημοσιογραφία. Τα ευρήματα της μελέτης μου αντικατοπτρίζουν τον τρόπο με τον οποίο η «δημοσιογραφία» παρέχεται όλο και περισσότερο από μια ποικιλία διαφορετικών, πιο εξειδικευμένων φορέων. Αυτή η διαφοροποίηση προκαλεί ανησυχίες σε ορισμένους από τους δημοσιογράφους από τους οποίους πήρα συνέντευξη. Για αυτούς, τα μέσα ενημέρωσης που υιοθετούν πρακτικές διευκόλυνσης μπορεί να αποδυναμώσουν την έννοια της «σκληρής», ερευνητικής δημοσιογραφίας. Ωστόσο, οι δημοσιογράφοι πρέπει να αναγνωρίσουν ότι είναι απίθανο να υπάρξει στο μέλλον μία και μόνη οριστική μορφή δημοσιογραφίας.

Εν κατακλείδι, μια ισχυρότερη επίγνωση τόσο των ιστορικών όσο και των σύγχρονων δεσμών με άλλες μορφές εργασίας με δεδομένα και κουλτούρας δεδομένων μπορεί να βοηθήσει τους δημοσιογράφους να προβληματιστούν σχετικά με τον δικό τους ρόλο και να έχουν μεγαλύτερη επίγνωση όχι μόνο των νέων εξαρτήσεων, αλλά και των πιθανών συνεργειών που μπορούν να χρησιμοποιηθούν για την υποστήριξη και τη δυνητική επέκταση της αποστολής τους.

Παραπομπές

Baack, S. (2018) «Practically engaged. The entanglements between data journalism and civic tech», *Digital Journalism*, 6(6), σσ. 673–692. DOI: https://doi.org/10.1080/21670811.2017.1375382

Usher, N. (2017) «Venture-backed news start-ups and the field of journalism», *Digital Journalism,* 5(9), σσ. 1116–1133. DOI: https://doi.org/10.1080/21670811.2016.1272064

Wagemans, A., Witschge, T., & Harbers, F. (2019) «Impact as driving force of journalistic and social change», *Journalism*, 20(4), σσ. 552–567. DOI: https://doi.org/10.1177/1464884918770538

Ο συγγραφέας

Ο Στέφαν Μπαάκ είναι ερευνητής των μέσων ενημέρωσης που συνεργάζεται με το Weizenbaum Institute for the Networked Society και αναλυτής έρευνας και δεδομένων στο Mozilla Foundation.

7.5
Πρακτικές Ανοιχτού Κώδικα στη Δημοσιογραφία Δεδομένων

Ryan Pitts, Lindsay Muscato

Περίληψη: Αυτό το κεφάλαιο εξετάζει τις προκλήσεις του ανοιχτού κώδικα για τη δημοσιογραφία και τα κοινά χαρακτηριστικά επιτυχημένων έργων.

Λέξεις-κλειδιά: ανοιχτός κώδικας, προγραμματισμός, κωδικοποίηση, δημοσιογραφία, ανάπτυξη εργαλείων, βιβλιοθήκες κώδικα

Φανταστείτε το εξής: δύο δημοσιογράφοι συνεργάζονται για να συλλέξουν αρχεία από κυβερνητικούς ιστοτόπους, να μετατρέψουν αυτά τα έγγραφα σε δεδομένα, να αναλύσουν αυτά τα δεδομένα για να αναζητήσουν μοτίβα και, στη συνέχεια, να δημοσιεύσουν μια οπτικοποίηση που θα αφηγείται μια ιστορία στους αναγνώστες. Κάποια εκδοχή αυτής της διαδικασίας εκτυλίσσεται στις εφημερίδες σε όλο τον κόσμο κάθε μέρα. Σε πολλές αίθουσες σύνταξης, κάθε βήμα βασίζεται, τουλάχιστον εν μέρει, σε λογισμικό ανοιχτού κώδικα, συγκεντρώνοντας δοκιμασμένα από την κοινότητα εργαλεία σε μια ροή εργασίας που είναι ταχύτερη από οποιονδήποτε τρόπο χρησιμοποιούσαμε πριν.

Δεν είναι, όμως, μόνο το λογισμικό ανοιχτού κώδικα που έχει γίνει μέρος της σημερινής ροής εργασίας της δημοσιογραφίας δεδομένων, αλλά και η *φιλοσοφία* του ανοιχτού κώδικα. Μοιραζόμαστε γνώσεις και δεξιότητες μεταξύ μας, σε εκδηλώσεις και μέσω διαύλων της κοινότητας και των μέσων κοινωνικής δικτύωσης. Δημοσιεύουμε μεθοδολογίες και δεδομένα, καλώντας τους συναδέλφους να διορθώσουν τις παραδοχές μας και δίνοντας στους αναγνώστες λόγους να εμπιστεύονται τα αποτελέσματά μας. Τέτοιες ανοιχτές, συνεργατικές προσεγγίσεις μπορούν να κάνουν τη δημοσιογραφία μας καλύτερη[202]. Κάθε φορά που ζητάμε σχολιασμό ή εξωτερικές συνεισφορές, κάνουμε το έργο μας πιο ανθεκτικό. Κάποιος άλλος μπορεί να εντοπίσει ένα πρόβλημα στον τρόπο που χρησιμοποιήσαμε τα δεδομένα σε μια ιστορία ή να συνεισφέρει ένα νέο χαρακτηριστικό που θα κάνει το λογισμικό μας καλύτερο.

Αυτές οι πρακτικές μπορούν, επίσης, να έχουν ευρύτερα οφέλη πέραν των δικών μας έργων και οργανισμών. Οι περισσότεροι από εμάς δεν θα ξεκινήσουμε ποτέ ένα μεγάλο έργο χρησιμοποιώντας μόνο εργαλεία που έχουμε κατασκευάσει οι ίδιοι και τεχνικές που έχουμε αναπτύξει μόνοι μας. Αντιθέτως, στηριζόμαστε πάνω στη δουλειά άλλων ανθρώπων, μαθαίνοντας από μέντορες, ακούγοντας σε συνέδρια και μαθαίνοντας πώς δημιουργήθηκαν τα έργα που μας αρέσουν.

Στο *OpenNews*, έχουμε συνεργαστεί με δημοσιογράφους σε έργα ανοιχτού κώδικα, έχουμε υποστηρίξει συνεργασίες προγραμματιστών και γράψαμε τον οδηγό *The Field Guide to Open Source in the Newsroom*[203]. Σε αυτό το κεφάλαιο αναφερόμαστε σε μερικά από τα πράγματα που μάθαμε σχετικά με τον ρόλο των πρακτικών ανοιχτού κώδικα στη δημοσιογραφία δεδομένων, συμπεριλαμβανομένων των κοινών προκλήσεων και των χαρακτηριστικών των επιτυχημένων έργων.

202. Βλ. επίσης τα κεφάλαια των Λίον και Μαζότε για διαφορετικές προσεγγίσεις σχετικά με τον ρόλο των πρακτικών και των φιλοσοφιών ανοιχτού κώδικα στη δημοσιογραφία δεδομένων (ΣτΕ: Κεφάλαιο 4.5, «Εξηγώντας τις Μεθόδους: Υπολογιστικά Φύλλα, Script και Notebook Προγραμματισμού», και Κεφάλαιο 4.6, «Δουλεύοντας Ανοιχτά στη Δημοσιογραφία Δεδομένων», αντίστοιχα).
203. https://opennews.org/, http://fieldguide.opennews.org/

Κοινές προκλήσεις

Η ανοιχτή συνεργασία μπορεί να είναι ικανοποιητική και διασκεδαστική και μπορείτε να μάθετε περισσότερα κατά τη διαδικασία –αλλά δεν είναι πάντα απλή! Ο σχεδιασμός για την επιτυχία σημαίνει ότι πρέπει να έχετε πλήρη επίγνωση των προκλήσεων που συχνά αντιμετωπίζουν τα έργα ανοιχτού κώδικα.

Επιχειρηματολογία. Μπορεί να είναι δύσκολο να πείσετε τους εκδότες, τις νομικές ομάδες και άλλους ότι το να «χαρίσετε» το έργο σας είναι καλή ιδέα. Μπορεί να υπάρχουν νομικές και επιχειρηματικές ανησυχίες, καθώς και ανησυχίες σχετικά με τη φήμη και τη βιωσιμότητα. Ως απάντηση, συνεργαζόμαστε με δημοσιογράφους για να τεκμηριώσουμε τα οφέλη των εργαλείων και της διαδικασίας της ανοιχτότητας, συμπεριλαμβανομένων του καλύτερου κώδικα, της καλής θέλησης της κοινότητας και της αυξημένης αξιοπιστίας[204].

Οι άνθρωποι προχωρούν μπροστά, το ίδιο και η τεχνολογία. Όταν ένα βασικό μέλος μιας ομάδας αναλαμβάνει άλλη δουλειά, ο χρόνος που έχει στη διάθεσή του, για να συντηρεί και να υποστηρίζει ένα έργο ανοιχτού κώδικα, συχνά εξανεμίζεται. Για παράδειγμα, πριν από μερικά χρόνια, οι *New York Times* κυκλοφόρησαν το Pourover, ένα framework JavaScript που υποστήριζε το γρήγορο φιλτράρισμα γιγαντιαίων συνόλων δεδομένων εντός προγράμματος περιήγησης. Το Pourover κοινοποιήθηκε ευρέως και άρχισε να δημιουργεί μια κοινότητα. Όμως, ένας από τους βασικούς προγραμματιστές έπιασε δουλειά αλλού και η ομάδα άρχισε να εξετάζει νεότερα εργαλεία για την επίλυση παρόμοιων προβλημάτων. Αυτό σε καμία περίπτωση δεν αποτελεί αρνητική κριτική για τον κώδικα ή τον προγραμματισμό του Pourover –μερικές φορές η διάρκεια ζωής ενός έργου είναι απλώς διαφορετική από ό,τι είχατε φανταστεί.

Πιέσεις επιτυχίας. Ακούγεται αντιφατικό, αλλά η διαπίστωση ότι οι άνθρωποι είναι πραγματικά ενθουσιασμένοι με κάτι που φτιάξατε μπορεί να σας φέρει δουλειά για την οποία δεν είστε έτοιμοι. Η ξαφνική, εκρηκτική δημοτικότητα αυξάνει την πίεση να συνεχίσετε να δημιουργείτε, να διορθώνετε σφάλματα και να ανταποκρίνεστε στη συνεισφορά της κοινότητας. Ο Έλιοτ Μπέντλεϊ πάλεψε με όλα αυτά μετά την κυκλοφορία της oTranscribe, μιας διαδικτυακής εφαρμογής που έγραψε, για να λύσει ένα πρόβλημα στην καθημερινή του δουλειά: την απομαγνητοφώνηση ηχητικών συνεντεύξεων. Λίγους μήνες αργότερα, είχε δεκάδες χιλιάδες ενεργούς χρήστες και ερωτήσεις σχετικά με το μέλλον του έργου.

Χαρακτηριστικά επιτυχημένων έργων

Υπάρχουν πολλά σπουδαία παραδείγματα ανοιχτού κώδικα στη δημοσιογραφία –ξεκινώντας με έργα που κυκλοφόρησαν από μια αίθουσα σύνταξης και υιοθετήθηκαν από πολλές άλλες μέχρι εκείνα που ήταν συνεργατικά από την αρχή. Οι πιο επιτυχημένες προσπάθειες που έχουμε δει μοιράζονται ένα ή περισσότερα χαρακτηριστικά, τα οποία περιγράφουμε παρακάτω.

Λύνουν ένα πρόβλημα που αντιμετωπίζετε κάθε μέρα. Κατά πάσα πιθανότητα, κάποιος άλλος βρίσκεται αντιμέτωπος με το ίδιο εμπόδιο ή το ίδιο σύνολο επαναλαμβανόμενων εργασιών με εσάς. Καλύπτοντας την ποινική δικαιοσύνη σε εθνικό επίπεδο, το *Marshall Project* παρακολουθεί εκατοντάδες ιστοτόπους για αλλαγές και ανακοινώσεις. Το να επισκέπτεται κανείς ξανά και ξανά μια λίστα διευθύνσεων URL δεν είναι καλή χρήση του χρόνου ενός δημοσιογράφου, αλλά είναι μια εξαιρετική χρήση για έναν διακομιστή στο cloud. Το Klaxon παρακολουθεί αυτές τις ιστοσελίδες και στέλνει μια ειδοποίηση κάθε φορά που κάποια από αυτές αλλάζει –είναι τόσο γρήγορο που η αίθουσα σύνταξης συχνά έχει πληροφορίες ακόμη και πριν από την επίσημη ανακοίνωσή τους[205]. Αυτού του είδους η παρακολούθηση είναι χρήσιμη για κάθε είδους ρεπορτάζ –και όταν το *Marshall Project* έλυσε ένα πρόβλημα για τους δημοσιογράφους του,

204. http://fieldguide.opennews.org/en/latest/Chapter01-Choosing-Open-Source/
205. https://newsklaxon.org/

το έλυσε και για άλλους οργανισμούς. Διαθέτοντας το Klaxon ως έργο ανοιχτού κώδικα, οι προγραμματιστές του βοηθούν δεκάδες αίθουσες σύνταξης στο ρεπορτάζ και λαμβάνουν ως αντάλλαγμα συνεισφορές κώδικα που βελτιώνουν το εργαλείο τους.

Επιλύουν ένα πρόβλημα που δεν είναι διασκεδαστικό. Η ομάδα δεδομένων/οπτικοποιήσεων του NPR χρειαζόταν έναν τρόπο να κάνει τα γραφικά να αλλάζουν διαστάσεις μαζί με τις σελίδες στις οποίες ήταν ενσωματωμένα. Πρόκειται για ένα σημαντικό χαρακτηριστικό, καθώς οι αναγνώστες χρησιμοποιούν ολοένα και περισσότερο τις κινητές συσκευές για να έχουν πρόσβαση στο περιεχόμενο των ειδήσεων, αλλά δεν είναι απαραίτητα ένα διασκεδαστικό πρόβλημα. Όταν το NPR κυκλοφόρησε το Pym.js, μια βιβλιοθήκη ανοιχτού κώδικα που έλυσε το πρόβλημα, αυτή δεν άργησε να υιοθετηθεί ευρέως από τη δημοσιογραφική κοινότητα.

Έχουν εξαιρετική τεκμηρίωση. Υπάρχει τεράστια διαφορά ανάμεσα στο να πετάς τον κώδικα στο Διαδίκτυο και στο να εξηγείς πραγματικά τι είναι ένα έργο και πώς χρησιμοποιείται. Οι προθεσμίες έχουν την τάση να καθιστούν τη συγγραφή τεκμηρίωσης («documentation») χαμηλή προτεραιότητα, αλλά ένα έργο δεν μπορεί να ευδοκιμήσει χωρίς αυτήν. Οι νέοι χρήστες χρειάζονται ένα σημείο εκκίνησης, και εσείς, επίσης, θα ευχαριστείτε τον εαυτό σας, όταν θα επανεξετάζετε τη δική σας δουλειά αργότερα. Το Wherewolf είναι μια μικρή υπηρεσία JavaScript που μπορείτε να χρησιμοποιήσετε, για να βρείτε πού βρίσκεται μια διεύθυνση μέσα σε ένα σύνολο ορίων (π.χ. σχολικές περιφέρειες ή όρια κομητείας). Παρόλο που ο κώδικας δεν έχει χρειαστεί ενημέρωση εδώ και αρκετό καιρό, η κοινότητα των χρηστών εξακολουθεί να αυξάνεται, τουλάχιστον εν μέρει επειδή η τεκμηρίωσή του είναι λεπτομερής και γεμάτη παραδείγματα.

Καλωσορίζουν τους συνεργάτες. Το California Civic Data Coalition διαθέτει μια σουίτα εργαλείων ανοιχτού κώδικα που βοηθούν τους δημοσιογράφους να χρησιμοποιούν τα δεδομένα χρηματοδότησης των προεκλογικών εκστρατειών της πολιτείας. Το έργο ξεκίνησε ως συνεργασία μεταξύ μερικών προγραμματιστών σε δύο αίθουσες σύνταξης, αλλά έχει αναπτυχθεί χάρη στις συνεισφορές φοιτητών, ασκούμενων, ανθρώπων που ασχολούνται με τα δεδομένα των πολιτών, ενδιαφερόμενων πολιτών και ακόμη και δημοσιογράφων χωρίς καμία εμπειρία στον προγραμματισμό. Αυτό δεν ήταν τυχαίο: η πρωτοβουλία διαθέτει έναν χάρτη πορείας με τα χαρακτηριστικά που πρέπει να κατασκευαστούν και τα σφάλματα που πρέπει να διορθωθούν, δημιουργούν αναρτήσεις με εκκρεμείς εργασίες για διαφορετικά επίπεδα εξειδίκευσης, και εμφανίζονται σε συνέδρια και σχεδιάζουν σπριντ που είναι ανοιχτά σε όλους.

Υπάρχουν πολλοί τρόποι για να μετρηθεί η επιτυχία ενός έργου ανοιχτού κώδικα μιας αίθουσας σύνταξης. Θέλετε να δημιουργήσετε μια κοινότητα και να ζητήσετε συνεισφορά; Χρειάζεστε μια επιπλέον γνώμη για τη δουλειά σας; Ή μήπως φτιάξατε κάτι που λύνει ένα πρόβλημα για εσάς και απλά αισθάνεστε καλά που γλιτώνετε άλλους ανθρώπους από την ταλαιπωρία; Εσείς αποφασίζετε τι σημαίνει επιτυχία για εσάς. Ανεξάρτητα από ό,τι θα επιλέξετε, η ανάπτυξη ενός σχεδίου που θα σας οδηγήσει εκεί θα έχει μερικά κοινά στοιχεία: να είστε ξεκάθαροι σχετικά με τον στόχο σας, ώστε να μπορείτε να δημιουργήσετε ένα τίμιο στρατηγικό πλάνο για τον εαυτό σας και να ορίσετε τις σωστές προσδοκίες για τους άλλους· να γράφετε τεκμηρίωση με φιλικό τρόπο, που να είναι βασισμένη σε παραδείγματα και να μπορεί να προσελκύσει νέους χρήστες και να εξηγήσει πώς ελήφθησαν οι αποφάσεις στην πορεία· να υιοθετείτε έναν συνεργατικό τρόπο εργασίας, ανοιχτό στους υπόλοιπους ανθρώπους. Θα μάθετε πολλά στην πράξη, οπότε βγείτε έξω και μοιραστείτε τα!

Οι συγγραφείς

Ο Ράιαν Πιτς είναι επικεφαλής του προγράμματος τεχνολογίας στο OpenNews, μια μη κερδοσκοπική ομάδα που συνδέει προγραμματιστές, σχεδιαστές και αναλυτές δεδομένων με πόρους (και μεταξύ τους!), με τρόπους που καθιστούν τη δημοσιογραφία πιο δίκαιη και συνεργατική.
Η Λίντσι Μουσκάτο είναι συντάκτρια, συγγραφέας και διαχειρίστρια δημοσιογραφικών έργων, πιο πρόσφατα στο OpenNews και σήμερα στη Vox Media.

7.6
Φεουδαρχία Δεδομένων: Πώς οι Πλατφόρμες Διαμορφώνουν Διασυνοριακά Ερευνητικά Δίκτυα

Ştefan Cândea

Περίληψη: Φεουδαρχία δεδομένων: πώς οι πλατφόρμες διαμορφώνουν τη διασυνοριακή ερευνητική δημοσιογραφία και ανοίγουν τον δρόμο για τον αποικισμό της.

Λέξεις-κλειδιά: διασυνοριακή έρευνα, πολιτική οικονομία δικτύωσης, κοινωνικοτεχνολογικός έλεγχος πρόσβασης, ριζοσπαστικός διαμοιρασμός, φεουδαρχία δεδομένων, πλατφόρμες

Η πλατφορμοποίηση της διασυνοριακής ερευνητικής δημοσιογραφίας είναι ένα φαινόμενο που αναπτύσσεται και υποστηρίζεται από τον ίδιο τεχνο-θετικισμό με την τρέχουσα τάση της πλατφορμοποίησης της κοινωνίας (Dijck κ.α., 2018). Οι πλατφόρμες για τη φιλοξενία δεδομένων για διασυνοριακές έρευνες άρχισαν να κερδίζουν έδαφος γύρω στο 2010, στο πλαίσιο της διεξαγωγής ερευνών με δεδομένα κατόπιν διαρροών. Ίσως το πιο αξιοσημείωτο παράδειγμα δημοσιογραφικής συνεργασίας μεγάλης κλίμακας που βασίζεται σε πλατφόρμα είναι τα βραβευμένα με Πούλιτζερ *Panama Papers*.

Προκειμένου να οργανώσει την αναζήτηση δεδομένων και την παρουσίαση ρεπορτάζ για τους 500 δημοσιογράφους που συμμετείχαν στην έρευνα για τα *Panama Papers*, η Διεθνής Σύμπραξη Ερευνητών Δημοσιογράφων («International Consortium of Investigative Journalists», ICIJ) ανέπτυξε μια πλατφόρμα με την ονομασία Global I-Hub (Wilson-Chapman, 2017)[206]. Ο Ράιλ (2017) περιγράφει την πλατφόρμα ως «ειδικά αναπτυγμένη τεχνολογία [...] που χρησιμοποιείται για την εξέταση και τη διανομή πληροφοριών, τη σύνδεση των δημοσιογράφων μεταξύ τους σε μια διαδικτυακή αίθουσα σύνταξης και τη διασφάλιση ότι οι δημοσιογράφοι εργάζονται ως μια παγκόσμια ομάδα». Αποκαλείται «το εικονικό γραφείο του ICIJ [...] ένα Facebook για δημοσιογράφους» τόσο από το συντακτικό όσο και από το ερευνητικό προσωπικό του ICIJ (Hare, 2016· Raab, 2016).

Τα δεδομένα και οι διασυνοριακές έρευνες υποτίθεται ότι ταιριάζουν απόλυτα και ενδυναμώνουν τις ανεξάρτητες δημοσιογραφικές συνεργασίες (Coronel, 2016· Houston, 2016). Οργανισμοί όπως το ICIJ, το Organised Crime and Corruption Reporting Project (OCCRP) και άλλοι, προσφέρουν σε μια προσεκτικά επιλεγμένη ομάδα εκατοντάδων δημοσιογράφων σε όλο τον κόσμο, δωρεάν (ή, καλύτερα, επιδοτούμενη) πρόσβαση σε αποκλειστικά σύνολα δεδομένων που είναι διαθέσιμα για αναζήτηση σε μια ιδιωτική ηλεκτρονική πλατφόρμα, που δεν είναι προσβάσιμη στον έξω κόσμο. Προσφέρουν, επίσης, μια πλατφόρμα για τη δημοσίευση και τη διαφήμιση των ιστοριών που παράγουν αυτοί οι δημοσιογράφοι.

206. Για μια διαφορετική προσέγγιση σχετικά με την πλατφόρμα I-Hub, δείτε το κεφάλαιο των Ντιάζ-Στρακ, Γκαλιέγο και Ρομέρα σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 4.2, «Οικοδόμηση Συνεργασιών για τα *Panama* και *Paradise Papers*»).

Για τους οργανισμούς αυτούς, η χρήση τέτοιων πλατφορμών επιτρέπει την επίτευξη κλιμάκωσης και αποδοτικότητας. Για τους μεμονωμένους δημοσιογράφους, το να έχουν αποκλειστική και ασφαλή πρόσβαση σε έναν ενιαίο χώρο με πλήθος δεδομένων από διαρροές, αρχεία εταιρειών, αποτελέσματα αιτημάτων για την ελευθερία της πληροφόρησης, αρχειακό υλικό, σημειώσεις δημοσιογράφων, προηγούμενες ιστορίες, ψηφιοποιημένους φακέλους δίωξης και δικαστικά αρχεία –για να αναφέρουμε μόνο μερικά– είναι σκέτη νιρβάνα. Αυτό ισχύει κυρίως για όσους εργάζονται απομονωμένοι και δεν διαθέτουν τους πόρους για ταξίδια και για την αποθήκευση και επεξεργασία δεδομένων.

Αν και αυτά τα βραχυπρόθεσμα οφέλη αναγνωρίζονται, δεν έχει ακόμη αναπτυχθεί κριτική έρευνα για το πώς αυτές οι πλατφόρμες έρευνας διαμορφώνουν τη θέση και το έργο των μεμονωμένων δημοσιογράφων που τις χρησιμοποιούν και των δικτύων στα οποία ανήκουν.

Η ύπαρξη πολύ λίγων φορέων που διαχειρίζονται τέτοιες πλατφόρμες και μεγάλου αριθμού δημοσιογράφων που εξαρτώνται από αυτές στο πεδίο της διασυνοριακής δημοσιογραφίας έχει συνέπειες. Μία από αυτές θα μπορούσε να εκληφθεί ως αυτό που στο τοπίο των «μεγάλων εταιρειών τεχνολογίας» έχει ονομαστεί «υπερσύγχρονη μορφή φεουδαρχίας» που βασίζεται στην ιδιοκτησία δεδομένων (Morozov, 2016). Η έννοια αυτή στρέφει την προσοχή στον τρόπο με τον οποίο ο πλήρης έλεγχος των δεδομένων και των αλληλεπιδράσεων των χρηστών περνάει στα χέρια λίγων εταιρειών που δεν αντιμετωπίζουν κανέναν ανταγωνισμό.

Αυτό το μοντέλο δημιουργεί μια σειρά από ανησυχίες. Μία σημαντική ανησυχία είναι ο έλεγχος πρόσβασης. Η πρόσβαση σε αυτές τις πλατφόρμες προστατεύεται για πολλούς και καλούς λόγους από πολλά επίπεδα ασφαλείας και δεν είναι ελεύθερη σε όλους τους δημοσιογράφους. Το ουσιαστικό ερώτημα είναι ποιος παίρνει την απόφαση για το ποιος συμπεριλαμβάνεται και ποιος αποκλείεται, ποιοι είναι οι κανόνες που διέπουν αυτές τις αποφάσεις, καθώς και ποιες είναι οι εντάσεις και οι συγκρούσεις που μπορεί να προκύψουν από αυτές. Η συμμετοχή σε τέτοιες πλατφόρμες συνήθως ρυθμίζεται από μια βασική συμφωνία τήρησης του απορρήτου ή μια συμφωνία συνεργασίας, όπου απαριθμούνται λεπτομερώς τα καθήκοντα του δημοσιογράφου ή του μέσου ενημέρωσης που αποκτά πρόσβαση, συνήθως με ελάχιστη αναφορά στα δικαιώματά τους. Τέτοια συστήματα και τα σχήματα που τα διέπουν δεν έχουν σχεδιαστεί με γνώμονα τις αρχές της συνιδιοκτησίας, αλλά μάλλον ως κεντρικά ελεγχόμενες δομές, με ενσωματωμένα χαρακτηριστικά την επιτήρηση των δραστηριοτήτων των χρηστών και την αστυνόμευση των παραβιάσεων των συμφωνιών.

Επιπλέον, η υιοθέτηση αυτού του μοντέλου στην ερευνητική δημοσιογραφία, όπως και στην υπόλοιπη «οικονομία του διαμοιρασμού», ενέχει τον κίνδυνο να δημιουργήσει ένα πρεκαριάτο στον χώρο της ερευνητικής δημοσιογραφίας. Ενδεικτικές αυτού του κινδύνου είναι οι αυτοπεριγραφές ορισμένων από τους οργανισμούς που διαχειρίζονται αυτές τις πλατφόρμες. Για παράδειγμα, το OCCRP περιγράφει τον εαυτό του ως «το Airbnb ή το Uber των δημοσιογράφων» που θέλουν να κάνουν «σπουδαίες διασυνοριακές έρευνες» (OCCRP, 2017).

Πράγματι, συχνά οι δημοσιογράφοι εργάζονται χωρίς αμοιβή σε δεδομένα που διαρρέουν και ανήκουν σε αυτούς τους οργανισμούς, ενώ πρέπει να πληρώνουν για την πρόσβαση στα δεδομένα αυτά με τις ιστορίες τους και ανά πάσα στιγμή κινδυνεύουν να διαγραφούν από την πλατφόρμα. Παρά τις δυσμενείς αυτές συνθήκες, οι δημοσιογράφοι πρέπει να δραστηριοποιούνται ολοένα και περισσότερο σε αυτές τις πλατφόρμες για να παραμείνουν στο παιχνίδι.

Για τους λόγους αυτούς, το επιχειρηματικό μοντέλο ενός μεγάλου διαμεσολαβητή-δικτύου ερευνών σήμερα μπορεί να θεωρηθεί ότι μοιάζει με αυτό μιας ψηφιακής πλατφόρμας της οικονομίας της περιστασιακής απασχόλησης. Η πρόσβαση στην πλατφόρμα μπορεί να ανακληθεί ανά πάσα στιγμή, η διαχείριση δεν είναι προς συζήτηση, η επιτήρηση της δραστηριότητας των χρηστών είναι ενσωματωμένη και «τα χρήματα είναι καλύτερο να μείνουν έξω από την εξίσωση» (Buzenberg, 2015). Η απλήρωτη εργασία και οι αλληλεπιδράσεις του «ριζοσπαστικού

διαμοιρασμού» εκατοντάδων δημοσιογράφων «πωλούνται» σε χορηγούς, χωρίς τα κέρδη να μοιράζονται. Η κυριότητα των διαρροών δεδομένων και της ανταλλαγής πληροφοριών που εμπλουτίζουν τις εν λόγω διαρροές, επίσης, δεν μοιράζεται με τους χρήστες. Τα δεδομένα που παράγονται από τις πληροφορίες που ανταλλάσσονται μεταξύ των χρηστών μοιράζονται μόνο υπό τη μορφή χαρακτηριστικών που καθιστούν την πλατφόρμα πιο αποτελεσματική και, έτσι, φέρνουν περισσότερες αλληλεπιδράσεις, περισσότερους χρήστες και, κατ' επέκταση, περισσότερους χορηγούς. Το πραγματικό κόστος των υπηρεσιών είναι άγνωστο στους χρήστες.

Τι μπορεί να γίνει για να διορθωθεί αυτή η τρέχουσα τάση στον κόσμο της ερευνητικής δημοσιογραφίας; Ένα βασικό πρώτο βήμα είναι να αναγνωριστεί ότι η κοινή χρήση δεδομένων με βάση την πλατφόρμα στα δίκτυα ερευνητικής δημοσιογραφίας πρέπει να συνοδεύεται από συζητήσεις σχετικά με τους κανόνες διαχείρισης και τον τεχνολογικό σχεδιασμό, καθώς και τη συνιδιοκτησία των δεδομένων και των ψηφιακών εργαλείων. Τα δίκτυα αυτά πρέπει να αναπτύξουν και να υιοθετήσουν δημόσιους κώδικες δεοντολογίας και να διαθέτουν μηχανισμούς λογοδοσίας για την αντιμετώπιση κάθε είδους καταχρήσεων. Η απουσία αυτών μπορεί να ενισχύσει τις επισφαλείς συνθήκες εργασίας των μεμονωμένων δημοσιογράφων, αντί να αναστατώσει τους παραδοσιακούς παράγοντες των μέσων ενημέρωσης.

Δεύτερον, ο στόχος δεν θα πρέπει να είναι η ανάπτυξη ενός μικρού αριθμού διασυνοριακών ερευνητικών δικτύων με χιλιάδες άτομα το καθένα. Αντίθετα, ο στόχος θα πρέπει να είναι να βρεθεί ένα καλό μοντέλο που να μπορεί να εφαρμοστεί σε ένα πλήθος ανεξάρτητων δικτύων που θα μπορούν να συνεργάζονται μεταξύ τους. Έτσι, αντί για ένα ενιαίο δίκτυο 150 συνεργαζόμενων μέσων ενημέρωσης, μια πιο επιθυμητή προσέγγιση θα ήταν να υπάρχουν δέκα δίκτυα των 15 εταίρων το καθένα. Το τελευταίο θα ήταν ανάλογο με τις αρχές ενός υγιούς συστήματος μέσων, συμπεριλαμβανομένου του θεμιτού ανταγωνισμού και του πλουραλισμού των μέσων ενημέρωσης. Χωρίς τέτοιες προσεγγίσεις, το συμμετοχικό δυναμικό των διασυνοριακών ερευνητικών δικτύων θα αποτύχει να υλοποιηθεί και, τροφοδοτούμενες από το αποτέλεσμα δικτύου, μερικές πλατφόρμες θα παγιωθούν σε ένα παγκόσμιο σύστημα φεουδαρχίας των ερευνητικών δεδομένων.

Παραπομπές

Buzenberg, W. E. (2015) «Anatomy of a global investigation: Collaborative, data-driven, without borders», *Shorenstein Center,* 6 Ιουλίου. Διαθέσιμο σε: https://shorensteincenter.org/anatomy-of-a-global-investigation-william-buzenberg/ (Προσπελάστηκε: 13 Δεκεμβρίου 2021).

Coronel, S. (2016) «Coronel: A golden age of global muckraking at hand», *Global Investigative Journalism Network*, 20 Ιουνίου. Διαθέσιμο σε: https://gijn.org/2016/06/20/a-golden-age-of-global-muckraking/ (Προσπελάστηκε: 13 Δεκεμβρίου 2021).

Dijck, J. van, Poell, T., & Waal, M. de. (2018). *The platform society: Public values in a connective world.* Νέα Υόρκη: Oxford University Press.

Hare, K. (2016) «How ICIJ got hundreds of journalists to collaborate on the Panama Papers», *Poynter*, 4 Απριλίου. Διαθέσιμο σε: https://www.poynter.org/reporting-editing/2016/how-icij-got-hundreds-of-journalists-to-collaborate-on-the-panama-papers/ (Προσπελάστηκε: 13 Δεκεμβρίου 2021).

Houston, B. (2016) «Panama papers showcase power of a global movement», *Global Investigative Journalism Network*, 14 Απριλίου. Διαθέσιμο σε: https://gijn.org/2016/04/13/panama-papers-showcase-power-of-a-global-movement/ (Προσπελάστηκε: 13 Δεκεμβρίου 2021).

Morozov, E. (2016) «Tech titans are busy privatising our data», *The Guardian*, 24 Απριλίου. Διαθέσιμο σε: https://www.theguardian.com/commentisfree/2016/apr/24/the-new-feudalism-silicon-valley-overlords-advertising-necessary-evil (Προσπελάστηκε: 13 Δεκεμβρίου 2021).

OCCRP (2017). *2016 Annual Report. OCCRP*. Διαθέσιμο σε: https://www.occrp.org/documents/AnnualReport2017.pdf (Προσπελάστηκε: 13 Δεκεμβρίου 2021).

Raab, B. (2016) «Behind the Panama Papers: A Q&A with International Consortium of Investigative Journalists director Gerard Ryle», *Ford Foundation*, 8 Απριλίου. Διαθέσιμο σε: https://www.fordfoundation.org/ideas/equals-change-blog/posts/behind-the-panama-papers-a-qa-with-international-consortium-of-investigative-journalists-director-gerard-ryle/ (Προσπελάστηκε: 13 Δεκεμβρίου 2021).

Ryle, G. (2017) «Paradise Papers: More documents, more reporters, more revelations», *ICIJ*, 5 Νοεμβρίου. https://www.icij.org/blog/2017/11/more-documents-more-journalists-and-bigger-revelations/

Wilson-Chapman, A. (2017) «Panama Papers a "notable security success"», *ICIJ*, 29 Αυγούστου. Διαθέσιμο σε: https://www.icij.org/blog/2017/08/panama-papers-notable-security-success/ (Προσπελάστηκε: 13 Δεκεμβρίου 2021).

Ο συγγραφέας

Ο Στεφάν Κάντεα είναι διδακτορικός ερευνητής στο CAMRI του Πανεπιστημίου Γουέστμινστερ, μέλος του ICIJ και συντονιστής του European Investigative Collaborations (EIC).

7.7
Κύρια Άρθρα Βάσει Δεδομένων; Σκέψεις για τη Χρήση Μετρήσεων Κοινού

Caitlin Petre

Περίληψη: Με βάση την εθνογραφική μελέτη της Κέιτλιν Πέτρε για το Chartbeat, τη Gawker Media και τους *New York Times*, το κεφάλαιο αυτό διερευνά τον ρόλο των μετρήσεων στη σύγχρονη παραγωγή ειδήσεων και προσφέρει συμβουλές για τις αίθουσες σύνταξης που ενσωματώνουν μετρήσεις στη συντακτική πρακτική.

Λέξεις-κλειδιά: μετρήσεις, αναλύσεις, αίθουσες σύνταξης, δημοσιογραφική πρακτική, εθνογραφία, συντακτική πρακτική

Στις 23 Αυγούστου 2013, ο σατιρικός ειδησεογραφικός ιστότοπος *The Onion* δημοσίευσε ένα άρθρο γνώμης, το οποίο υποτίθεται ότι γράφτηκε από την αρχισυντάκτρια του διαδικτυακού τμήματος του CNN, Μέρεντιθ Άρτλεϊ, με τίτλο *Let me explain why Miley Cyrus' VMA performance was our top story this morning* («Επιτρέψτε μου να εξηγήσω γιατί η εμφάνιση της Μάιλι Σάιρους στα VMA ήταν η κορυφαία είδηση σήμερα το πρωί»)[207]. Η απάντηση, όπως εξηγούσε το άρθρο, ήταν «αρκετά απλή»:

> Ήταν μια προσπάθεια να σας αναγκάσουμε να κάνετε κλικ στο CNN.com, ώστε να αυξήσουμε την επισκεψιμότητά μας, η οποία με τη σειρά της θα μας επέτρεπε να αυξήσουμε τα διαφημιστικά μας έσοδα. Δεν υπήρχε τίποτα, και εννοώ τίποτα, σε αυτήν την ιστορία που να σχετίζεται με τις σημαντικές ειδήσεις της ημέρας, με την καταγραφή γεγονότων που έχουν σημασία για την ανθρωπότητα ή με την ιδέα ότι η ίδια η δημοσιογραφία μπορεί να αποτελέσει δύναμη θετικής αλλαγής στον κόσμο. [...] Αλλά, ρε φίλε, μας έφερε μεγάλη διαδικτυακή κίνηση (Artley, 2013).

Το δημοσίευμα συνέχιζε αναφέροντας συγκεκριμένες μετρήσεις όπως οι προβολές σελίδων και τα ποσοστά εγκατάλειψης ως παράγοντες που παρακίνησαν το CNN να δώσει στην ιστορία της Σάιρους εξέχουσα θέση στην αρχική σελίδα.

Φυσικά, δεν ήταν η Άρτλεϊ που έγραψε στην πραγματικότητα το άρθρο, αλλά παρ' όλα αυτά χτύπησε φλέβα στους κύκλους των μέσων ενημέρωσης –ειδικά από τη στιγμή που μια ιστορία σχετικά με την περιβόητη εμφάνιση της Σάιρους στα MTV Video Music Awards είχε καταλάβει την πρώτη θέση στο CNN.com και, όπως επιβεβαίωσε αργότερα η πραγματική Μέρεντιθ Άρτλεϊ, έφερε την υψηλότερη επισκεψιμότητα από κάθε άλλο θέμα στον ιστότοπο εκείνη την ημέρα. Το ψεύτικο άρθρο γνώμης μπορεί να ερμηνευτεί όχι μόνο ως αποδοκιμασία προς το CNN, αλλά και ως σχόλιο για τη θλιβερή κατάσταση της κρίσης των ειδήσεων στην εποχή των διαδικτυακών μετρήσεων.

207. Αυτό το κομμάτι αποτελεί απόσπασμα και προσαρμογή από το *The Traffic Factories: Metrics at Chartbeat, Gawker Media, and The New York Times* («Τα Εργοστάσια της Επισκεψιμότητας: Μετρήσεις στο Chartbeat, στη Gawker Media και στους *New York Times*»), που δημοσιεύθηκε αρχικά από το Tow Center for Digital Journalism στη Σχολή Δημοσιογραφίας του Πανεπιστημίου Columbia το 2015. Αναδημοσιεύεται, κατόπιν άδειας.

Οι εταιρείες μέσων ενημέρωσης ανέκαθεν κατέβαλλαν προσπάθειες για τη συλλογή δεδομένων σχετικά με τα δημογραφικά στοιχεία και τη συμπεριφορά του κοινού τους. Όμως, οι δυνατότητες παρακολούθησης του Διαδικτύου, καθώς και η δυνατότητα αποθήκευσης και ανάλυσης τεράστιων ποσοτήτων δεδομένων, σημαίνουν ότι οι μετρήσεις κοινού έχουν γίνει πολύ πιο εξελιγμένες τα τελευταία χρόνια. Εκτός από τις προαναφερθείσες προβολές σελίδων («page views») και τα ποσοστά εγκατάλειψης («bounce rates»), τα εργαλεία ανάλυσης παρακολουθούν μεταβλητές όπως τα ποσοστά επιστροφής των επισκεπτών, τους ιστοτόπους αναφοράς («referral sites»), το βάθος κύλισης («scroll depth») και τον χρόνο παραμονής σε μια σελίδα. Πολλά από αυτά τα δεδομένα διαβιβάζονται στους ειδησεογραφικούς οργανισμούς σε πραγματικό χρόνο.

Τα dashboard μετρήσεων είναι πλέον σχεδόν πανταχού παρόντα στις σύγχρονες αίθουσες σύνταξης και οι έντονες συζητήσεις σχετικά με το πώς και πότε πρέπει να λαμβάνονται υπόψη είναι σχεδόν τόσο διαδεδομένες όσο και οι ίδιες οι μετρήσεις. Δεν αποτελεί έκπληξη το γεγονός ότι οι μετρήσεις έχουν γίνει ένα καυτό θέμα στη δημοσιογραφία. Η παρουσία τους καλεί μια σειρά από πάντα διαρκείς εντάσεις στα εμπορικά ειδησεογραφικά μέσα να έρθουν στο προσκήνιο. Μεταξύ αυτών: ποια είναι η θεμελιώδης αποστολή της δημοσιογραφίας και πώς μπορούν οι ειδησεογραφικοί οργανισμοί να γνωρίζουν πότε επιτυγχάνουν αυτήν την αποστολή; Πώς μπορούν οι εταιρείες των μέσων να συμβιβάσουν την επιταγή για κέρδος με τις απαιτήσεις της κοινωνίας; Στον βαθμό που η διάκριση μεταξύ δημοσιογράφου και κοινού εξακολουθεί να έχει νόημα, τι είδους σχέση πρέπει να έχουν οι δημοσιογράφοι με τους αναγνώστες τους; Οι μετρήσεις κοινού είναι πλέον πανταχού παρούσες στους ειδησεογραφικούς οργανισμούς, αλλά έχει γίνει ελάχιστη εμπειρική έρευνα σχετικά με το πώς παράγονται τα δεδομένα ή πώς επηρεάζουν την κουλτούρα της αίθουσας σύνταξης και την καθημερινή εργασία των δημοσιογράφων.

Με την υποστήριξη του Tow Center for Digital Journalism του Πανεπιστημίου Columbia, ανέλαβα ένα εθνογραφικό ερευνητικό πρόγραμμα μακράς διάρκειας, προκειμένου να κατανοήσω τον τρόπο με τον οποίο η χρήση των μετρήσεων αλλάζει τη συμπεριφορά των δημοσιογράφων, και τι σημαίνει αυτό για τη δημοσιογραφία. Τα βασικά ερευνητικά μου ερωτήματα περιλάμβαναν τα εξής:

- Πώς παράγονται οι μετρήσεις; Δηλαδή, πώς οι προγραμματιστές, οι επιστήμονες δεδομένων, οι ντιζάινερ, οι υπεύθυνοι προϊόντων, οι υπεύθυνοι μάρκετινγκ και οι πωλητές που κατασκευάζουν και πωλούν αυτά τα εργαλεία αποφασίζουν ποιες πτυχές της συμπεριφοράς του κοινού θα πρέπει να μετρηθούν και πώς θα τις μετρήσουν; Ποιες ιδέες –τόσο για εκείνους των οποίων τη συμπεριφορά μετρούν (καταναλωτές ειδήσεων) όσο και για εκείνους που θα χρησιμοποιήσουν το εργαλείο τους (δημοσιογράφοι)– ενσωματώνονται σε αυτές τις αποφάσεις; Πώς γνωστοποιούν οι εταιρείες ανάλυσης την αξία των μετρήσεων στους ειδησεογραφικούς οργανισμούς;
- Πώς ερμηνεύονται οι μετρήσεις; Παρά τις αντίθετες τοποθετήσεις τους, τα επιχειρήματα ότι οι μετρήσεις είναι καλές ή κακές για τη δημοσιογραφία έχουν ένα κοινό σημείο: τείνουν να υποθέτουν ότι το νόημα των μετρήσεων είναι σαφές και ξεκάθαρο. Όμως, ένας αριθμός από μόνος του δεν σημαίνει τίποτα χωρίς ένα εννοιολογικό πλαίσιο με το οποίο μπορεί να ερμηνευτεί. Ποιος δίνει νόημα στις μετρήσεις και με ποιο τρόπο;
- Πώς χρησιμοποιούνται πραγματικά οι μετρήσεις στη δημοσιογραφία; Καθορίζουν τα δεδομένα τον τρόπο με τον οποίο οι αίθουσες σύνταξης ειδήσεων αναθέτουν, γράφουν και προβάλλουν τα θέματά τους; Με ποιους τρόπους, εάν υπάρχουν, τα δεδομένα παίζουν ρόλο στις αποφάσεις για το προσωπικό, όπως αυξήσεις, προαγωγές και απολύσεις; Τα δεδομένα παίζουν περισσότερο ρόλο στην καθημερινή εργασία ή στη μακροπρόθεσμη στρατηγική; Και πώς οι απαντήσεις σε αυτά τα ερωτήματα διαφέρουν μεταξύ των διαφόρων πλαισίων του κάθε οργανισμού;

Για να απαντήσω σε αυτά τα ερωτήματα, διεξήγαγα μια εθνογραφική μελέτη του ρόλου των μετρήσεων στις σύγχρονες ειδήσεις εξετάζοντας τρεις περιπτωσιολογικές μελέτες: Chartbeat,

Gawker Media και *New York Times*. Μέσω ενός συνδυασμού παρατήρησης και συνεντεύξεων με διευθυντές παραγωγής, επιστήμονες δεδομένων, δημοσιογράφους, μπλόγκερ, συντάκτες και άλλους, η πρόθεσή μου ήταν να αναδείξω τις παραδοχές και τις αξίες που διέπουν τις μετρήσεις κοινού, την επίδραση των μετρήσεων στην καθημερινή εργασία των δημοσιογράφων και τους τρόπους με τους οποίους οι μετρήσεις αλληλεπιδρούν με την κουλτούρα ενός οργανισμού. Παρακάτω, θα συνοψίσω μερικές από τις κεντρικές διαπιστώσεις μου.

- Τα dashboard των αναλύσεων και των μετρήσεων έχουν σημαντικές συναισθηματικές διαστάσεις που πολύ συχνά παραβλέπονται. Οι μετρήσεις, και το ευρύτερο φαινόμενο των «μεγάλων δεδομένων» («big data»), τμήμα του οποίου αποτελούν, περιγράφονται συνήθως ως μια δύναμη εξορθολογισμού: δηλαδή, επιτρέπουν στους ανθρώπους να λαμβάνουν αποφάσεις με βάση ψύχραιμες, αντικειμενικές πληροφορίες –και όχι με βάση την αναξιόπιστη διαίσθηση ή κρίση τους. Αν και αυτή η περιγραφή δεν είναι εσφαλμένη, είναι ελλιπής. Η δύναμη και η ελκυστικότητα των μετρήσεων βασίζονται, σε μεγάλο βαθμό, στην ικανότητα των δεδομένων να προκαλούν συγκεκριμένα συναισθήματα, όπως ενθουσιασμό, απογοήτευση, επικύρωση και επιβεβαίωση. Η Chartbeat γνώριζε ότι αυτή η συναισθηματική αξία ήταν ένα ισχυρό κομμάτι της ελκυστικότητας αυτών των dashboard και η εταιρεία συμπεριέλαβε τέτοια χαρακτηριστικά στην πλατφόρμα της, ώστε να προκαλέσει τη συναισθηματική ανταπόκριση των χρηστών. Για παράδειγμα, το dashboard σχεδιάστηκε για να εκφράζει σεβασμό στη δημοσιογραφική κρίση, να μετριάζει το πλήγμα της χαμηλής επισκεψιμότητας και να παρέχει ευκαιρίες για πανηγυρισμό στις αίθουσες σύνταξης.
- Ο αντίκτυπος ενός εργαλείου ανάλυσης εξαρτάται από τον οργανισμό που το χρησιμοποιεί. Συχνά θεωρείται ότι η ίδια η ύπαρξη ενός εργαλείου ανάλυσης θα αλλάξει τον τρόπο λειτουργίας μιας αίθουσας σύνταξης με συγκεκριμένους τρόπους. Ωστόσο, διαπίστωσα ότι το πλαίσιο του κάθε οργανισμού είχε μεγάλη επιρροή στη διαμόρφωση της πιθανότητας και του τρόπου που οι μετρήσεις επηρεάζουν την παραγωγή ειδήσεων. Για παράδειγμα, η Gawker Media και οι *New York Times* είναι και οι δύο πελάτες της Chartbeat, αλλά το εργαλείο χρησιμοποιείται με πολύ διαφορετικούς τρόπους σε κάθε περιβάλλον. Στο Gawker, οι μετρήσεις ήταν ιδιαίτερα ορατές και είχαν μεγάλη απήχηση. Στους *New York Times*, ήταν λιγότερο σημαντικές και συχνά χρησιμοποιούνταν για να επιβεβαιώσουν αποφάσεις που είχαν ήδη ληφθεί από τους συντάκτες. Αυτό υποδηλώνει ότι είναι αδύνατο να γνωρίζουμε πώς οι μετρήσεις επηρεάζουν τη δημοσιογραφία χωρίς να εξετάσουμε πώς χρησιμοποιούνται σε συγκεκριμένες αίθουσες σύνταξης.
- Για τους συντάκτες, η κουλτούρα που βασίζεται στις μετρήσεις μπορεί να είναι ταυτόχρονα πηγή άγχους και καθησυχασμού. Είναι, επίσης, εκπληκτικά συμβατή με την αντίληψη της δημοσιογραφικής ελευθερίας. Ενώ οι συντάκτες της Gawker Media θεωρούσαν αγχωτικές τις πιέσεις της επισκεψιμότητας, πολλοί επηρεάστηκαν ψυχολογικά πολύ περισσότερο από το διαδικτυακό βιτριόλι στα σχόλια και στα μέσα κοινωνικής δικτύωσης. Σε ένα κλίμα διαδικτυακής εχθρότητας, ή ακόμη και παρενόχλησης, οι συντάκτες στράφηκαν μερικές φορές στις μετρήσεις ως μια καθησυχαστική υπενθύμιση της επαγγελματικής τους επάρκειας. Είναι ενδιαφέρον ότι οι αρθρογράφοι και οι συντάκτες γενικά δεν θεώρησαν τα συστήματα αξιολόγησης της εταιρείας με βάση την επισκεψιμότητα εμπόδιο στη δημοσιογραφική τους αυτονομία. Αυτό υποδηλώνει ότι οι δημοσιογράφοι σε εταιρείες μέσων ενημέρωσης που λειτουργούν μόνο στο Διαδίκτυο, όπως η Gawker Media, μπορεί να έχουν διαφορετικές αντιλήψεις σχετικά με τη δημοσιογραφική ελευθερία και τους περιορισμούς από ό,τι οι συνάδελφοί τους στα παραδοσιακά μέσα ενημέρωσης.

Εν κατακλείδι, έχω τις ακόλουθες προτάσεις για τους ειδησεογραφικούς οργανισμούς:

- Οι ειδησεογραφικοί οργανισμοί θα πρέπει να δώσουν προτεραιότητα στη στρατηγική σκέψη όσον αφορά σε θέματα που σχετίζονται με την ανάλυση (π.χ. ο ενδεδειγμένος ρόλος των μετρήσεων στον οργανισμό και οι τρόποι με τους οποίους τα δεδομένα «συνομιλούν» με τους

δημοσιογραφικούς στόχους του οργανισμού). Οι περισσότεροι δημοσιογράφοι ήταν πολύ απασχολημένοι με τις καθημερινές τους εργασίες για να σκεφτούν εκτενώς ή αφαιρετικά για τον ρόλο των μετρήσεων στον οργανισμό τους ή ποιες μετρήσεις πλαισίωναν καλύτερα τους δημοσιογραφικούς τους στόχους. Ως αποτέλεσμα, είχαν την τάση να συμβουλεύονται, να ερμηνεύουν και να χρησιμοποιούν μετρήσεις με έναν ad hoc τρόπο. Ωστόσο, αυτά τα δεδομένα είναι πάρα πολύ ισχυρά για να τα κάνουμε χρήση τους στο πόδι. Οι αίθουσες σύνταξης θα πρέπει να δημιουργούν ευκαιρίες –είτε εσωτερικά είτε σε συνεργασία με εξωτερικούς ερευνητές– για στοχαστική, σκόπιμη σκέψη μακριά από τις καθημερινές πιέσεις της παραγωγής σχετικά με τον καλύτερο τρόπο χρήσης των στοιχείων των μετρήσεων.

- Κατά την επιλογή μιας υπηρεσίας αναλύσεων, οι διευθυντές της αίθουσας σύνταξης θα πρέπει να κοιτάξουν πέρα από τα εργαλεία και να εξετάσουν ποιοι στρατηγικοί στόχοι, επιχειρηματικές επιταγές και αξίες του προμηθευτή συμπληρώνουν καλύτερα τους στρατηγικούς στόχους, τις επιχειρηματικές επιταγές και τις αξίες της αίθουσας σύνταξης. Έχουμε την τάση να βλέπουμε τους αριθμούς –και, κατ' επέκταση, τα dashboard– ως επίσημες και αμερόληπτες αντανακλάσεις του εμπειρικού κόσμου. Κατά την επιλογή μιας υπηρεσίας αναλύσεων, ωστόσο, είναι σημαντικό να θυμάστε ότι οι εταιρείες ανάλυσης έχουν τις δικές τους επιχειρηματικές επιταγές.
- Κατά τη χάραξη εσωτερικών πολιτικών για τη χρήση των μετρήσεων, οι διευθυντές των newsroom θα πρέπει να εξετάζουν τις πιθανές επιπτώσεις των δεδομένων επισκεψιμότητας όχι μόνο στο συντακτικό περιεχόμενο αλλά και στους εργαζομένους στη σύνταξη. Από τη στιγμή που οι κατατάξεις έχουν μια περίοπτη θέση στον τοίχο μιας αίθουσας σύνταξης ή ενός ιστοτόπου, μπορεί να είναι δύσκολο να περιοριστεί η επιρροή τους. Η κατάταξη με βάση την επισκεψιμότητα μπορεί να επισκιάσει άλλες μορφές αξιολόγησης, ακόμη και αν δεν ήταν αυτή η πρόθεση.
- Αν και οι προσπάθειες για την ανάπτυξη καλύτερων μετρήσεων είναι απαραίτητες και χρήσιμες, οι εφημερίδες και οι εταιρείες ανάλυσης θα πρέπει να είναι προσεκτικές ως προς τους περιορισμούς των μετρήσεων. Καθώς οι οργανωτικές προτεραιότητες και τα συστήματα αξιολόγησης βασίζονται ολοένα και περισσότερο σε μετρήσεις, υπάρχει ο κίνδυνος να μπερδευτεί το ποσοτικά μετρήσιμο με το αξιόλογο. Δεν μπορούν –ή δεν πρέπει– τα πάντα να γίνονται αντικείμενο καταμέτρησης. Η αίθουσα σύνταξης, οι εταιρείες ανάλυσης, οι χρηματοδότες και οι ερευνητές των μέσων ενημέρωσης θα μπορούσαν να λάβουν υπόψη πώς ορισμένα από τα πιο συναρπαστικά και απαραίτητα χαρακτηριστικά της δημοσιογραφίας, όπως η κοινωνική της αποστολή, δεν είναι εύκολο να μετρηθούν. Σε μια εποχή που οι αναλύσεις δεδομένων αποκτούν όλο και μεγαλύτερη αξία, πρέπει να προσέξουμε να μην εξισώνουμε το μετρήσιμο με το αξιόλογο.

Παραπομπές

Artley, M. (2013) «Let me explain why Miley Cyrus' VMA performance was our top story this morning», *The Onion*, 26 Αυγούστου. Διαθέσιμο σε: https://www.theonion.com/let-me-explain-why-miley-cyrus-vma-performance-was-our-1819584893 (Προσπελάστηκε: 14 Δεκεμβρίου 2021).

Η συγγραφέας

Η Κέιτλιν Πέτρε είναι Επίκουρη Καθηγήτρια Δημοσιογραφίας και Σπουδών Μέσων Μαζικής Ενημέρωσης στο Πανεπιστήμιο Rutgers. Η δουλειά της εξετάζει τις κοινωνικές και πολιτικές επιπτώσεις ενός κόσμου που είναι όλο και περισσότερο κορεσμένος από δεδομένα.

8.

Μαθαίνοντας τη Δημοσιογραφία Δεδομένων από Κοινού

8.1

Δημοσιογραφία Δεδομένων, Ψηφιακή Οικουμενικότητα και Καινοτομία στην Περιφέρεια

Anita Say Chan

Περίληψη: Ο «μύθος της ψηφιακής οικουμενικότητας» εκδηλώνεται όχι μόνο στα μέσα με τα οποία επιβάλλει στις δημόσιες αφηγήσεις και φαντασιώσεις να περιστρέφονται σταθερά και αποκλειστικά γύρω από τα λεγόμενα «κέντρα» καινοτομίας, αλλά και στα μέσα με τα οποία αποστρέφει παράλληλα την προσοχή από την ψηφιακή δυναμική πέρα από τα συγκεκριμένα κέντρα –μια δυναμική που δομεί αποικιοκρατικές σχέσεις ως προς τα δεδομένα και την περιφέρεια, τις οποίες τόσο οι δημοσιογράφοι όσο και οι μελετητές των παγκόσμιων ψηφιακών πολιτισμών πρέπει να προσέχουν να μην αναπαράγουν.

Λέξεις-κλειδιά: ψηφιακή οικουμενικότητα, κέντρο, περιφέρεια, αποικιακές σχέσεις, αποαποικιακή πληροφορική, τοπική καινοτομία

Η «ψηφιακή οικουμενικότητα» («digital universalism») είναι το διάχυτο αλλά λανθασμένο πλαίσιο που διαμορφώνει τις παγκόσμιες φαντασιώσεις γύρω από την ψηφιακή τεχνολογία και το οποίο θεωρεί ως δεδομένο ότι μια ενιαία, παγκόσμια αφήγηση που προωθείται από «κέντρα» καινοτομίας μπορεί να αντιπροσωπεύσει επαρκώς τις μορφές της ψηφιακής ανάπτυξης που βρίσκονται σε εξέλιξη σε ολόκληρο τον κόσμο σήμερα. Θεωρεί ως δεδομένο, δηλαδή, ότι τα συγκεκριμένα κέντρα της σύγχρονης «καινοτομίας» και του τεχνολογικού σχεδιασμού θα καθορίσουν αναπόφευκτα το ψηφιακό μέλλον που έρχεται να εξαπλωθεί σε ολόκληρο τον κόσμο για την πλειονότητα των «υπολοίπων του ψηφιακού κόσμου» («digital rest»). Και το γεγονός αυτό αντανακλάται μέσω της περιστασιακής παραδοχής ότι οι καλύτερες, πιο «νόμιμες» τοποθεσίες από τις οποίες μπορεί κανείς να μελετήσει και να παρατηρήσει τον τεχνολογικό μετασχηματισμό, την ψηφιακή παραγωγικότητα και πρακτική, ή την καινοτομία και την έρευνα με βάση την πληροφορία, προέρχονται από τέτοια κέντρα. Τα κυριότερα από αυτά: τα εργαστήρια, τα γραφεία και οι ερευνητικοί χώροι που βρίσκονται στη Σίλικον Βάλεϊ και τα διάσπαρτα αντίστοιχά τους σε άλλες πρωτεύουσες της καινοτομίας παγκοσμίως, τα οποία συγκεντρώνουν την ελίτ της ψηφιακής τεχνογνωσίας.

Από αυτά τα κέντρα πιθανολογείται πως προέρχεται η ψηφιακή κουλτούρα και εκεί βρίσκει την καθαρότερη μορφή και εκδήλωσή της –μόνο και μόνο για να αναπαραχθεί αλλού. Εκεί είναι που αναδύονται τα οράματα για το ψηφιακό μέλλον στις πιο ακριβείς ή ιδανικές προσεγγίσεις του. Και εκεί κυριαρχούν οι τεχνολογικές εξελίξεις –και συνεπώς οι ψηφιακές πολιτιστικές εξελίξεις– ως οι πιο δυναμικές, ζωντανές και εμπνευσμένες. Υποτίθεται, με άλλα λόγια, ότι υπάρχουν κάποιες πιο «αυθεντικές» και παραγωγικές τοποθεσίες από τις οποίες μπορεί κανείς να ξεκινήσει να μελετά τον ψηφιακό πολιτισμό –παρά τις μοναδικά παγκόσμιες διαστάσεις του– και να παρατηρεί τη δυναμική του.

Ως νεαρή ερευνήτρια, η μελέτη και η συγγραφή για τον ψηφιακό πολιτιστικό ακτιβισμό και την πολιτική στο Περού και στη Λατινική Αμερική, από τις αρχές της δεκαετίας του 2000 και μετά, διαμόρφωσε την εμπειρία μου με θεμελιώδη τρόπο, κυρίως μέσα από την φαινομενικά αθώα

ερώτηση που άκουγα συχνά: «γιατί να πάει κανείς στο Περού ή στη Λατινική Αμερική, για να μελετήσει τον ψηφιακό πολιτισμό;». Δεν υπήρχαν «καλύτερα» μέρη από τα οποία θα μπορούσα να μελετήσω τον ψηφιακό πολιτισμό; Δεν θα ήταν προτιμότερο να επενδύσω τον χρόνο μου παρακολουθώντας και καταγράφοντας τη δραστηριότητά του σε μια τοποθεσία όπως η Σίλικον Βάλεϊ;

Για τους αμφισβητίες αυτού του τύπου, το Περού, αναπόφευκτα, παραπέμπει σε έναν ορεινό νοτιοαμερικανικό λαό που κάποτε ήταν η καρδιά του πολιτισμού των Ίνκας: πατρίδα του Μάτσου Πίτσου, των ψηλών εκτάσεων της οροσειράς των Άνδεων και των μεγάλων πληθυσμών των κοινοτήτων που μιλούν τις γλώσσες Κετσούα και Αϊμάρα. Αυτό το Περού μπορεί να είναι γνωστό ως ένας ιδανικός χώρος από τον οποίο μπορεί κανείς να ανατρέξει στην παράδοση του παρελθόντος, στον πολιτισμό των ιθαγενών ή στην αφθονία της γενναιόδωρης φύσης, αλλά θεωρείται ότι έχει ελάχιστα να μας πει για τη δυναμική της σύγχρονης ψηφιακής κουλτούρας, τις ροές υψηλής τεχνολογίας ή τις σχετικές μελλοντικές εξελίξεις. Μέρη όπως το Περού θα μπορούσαν να ξεκλειδώσουν ένα μονοπάτι γεμάτο από τα κειμήλια και τους θησαυρούς ενός τεχνολογικού παρελθόντος που θα πρέπει να παλέψουμε κυριολεκτικά για να μην ξεχάσουμε, ενώ τόποι όπως η Σίλικον Βάλεϊ είναι το μέρος όπου ξεκλειδώνονται αναπόφευκτα τα μυστικά ενός τεχνολογικού μέλλοντος στου οποίου το μονοπάτι δεν έχουμε ακόμη περπατήσει.

Σιωπηλά, λοιπόν, μια τέτοια ερώτηση επιβεβαιώνει μια περιστασιακή βεβαιότητα γύρω από την ιδέα ότι το ψηφιακό μέλλον, που φαντάζονται επιλεγμένοι πληθυσμοί τεχνολόγων σε επίλεκτα κέντρα σχεδιασμού, αντιπροσωπεύει και το υπόλοιπο παγκόσμιο γίγνεσθαι και ότι το παρόν, που εκτυλίσσεται σήμερα στα κέντρα καινοτομίας, πρέπει σίγουρα να είναι το μέλλον της περιφέρειας.

Η δύναμη του «μύθου της ψηφιακής οικουμενικότητας», λοιπόν, εκδηλώνεται όχι μόνο στα μέσα με τα οποία επιβάλλει στις δημόσιες αφηγήσεις και φαντασιώσεις να περιστρέφονται σταθερά και αποκλειστικά γύρω από τα λεγόμενα «κέντρα» καινοτομίας, αλλά και στα μέσα με τα οποία αποστρέφει παράλληλα την προσοχή στην ψηφιακή δυναμική πέρα από τα συγκεκριμένα κέντρα. Ως εκ τούτου, περιορίζει την ποικιλομορφία και την παγκόσμια κυκλοφορία των αφηγήσεων γύρω από τις πραγματικές ψηφιακές δυναμικές, που εκδηλώνονται σε μια σειρά από τόπους, αδιαφορεί για τις ποικίλες μορφές ψηφιακής δημιουργικότητας και ενισχύει τεχνητά την προβολή των μητροπόλεων της «καινοτομίας» ως αποκλειστικών τόπων ψηφιακής παραγωγικότητας.

Υπάρχει μια ιδιαίτερη αποικιοκρατική έννοια της περιφέρειας που δημιουργείται εδώ, την οποία τόσο οι δημοσιογράφοι όσο και οι μελετητές των παγκόσμιων ψηφιακών πολιτισμών πρέπει να προσέχουν να μην αναπαράγουν: δηλαδή, αυτή της «περιφέρειας» ως απλού δορυφόρου ενός μέλλοντος που επινοήθηκε a priori κάπου αλλού. Πράγματι, η περιφέρεια δεν είναι, σε καμία περίπτωση, τόσο παθητική και δεν πάσχει από έλλειψη εφευρετικότητας. Οι ζωντανές και δυναμικές υπαίθριες αγορές ή τα Internet café γεμάτα με μεταχειρισμένους, ανακυκλωμένους, και επανασυναρμολογημένους υπολογιστές και εξαρτήματα είναι καινοτομίες του Παγκόσμιου Νότου, που επέκτειναν την πρόσβαση στο Διαδίκτυο, με χαμηλό κόστος, και αναβάθμισαν την κυκλοφορία του διεθνούς και τοπικού περιεχομένου των μέσων ενημέρωσης, σε διαφορετικούς πληθυσμούς, τόσο σε αγροτικές όσο και σε αστικές ζώνες. Αυτά τα τεχνολογικά τεχνάσματα και οι τοπικοί αυτοσχεδιασμοί αποτελούν ένα καθημερινό κομμάτι του τεχνολογικού τοπίου της περιφέρειας, η ζωντάνια του οποίου μόνο εν μέρει μπορεί να γινεί κατανοητή, όταν συγκρίνεται με τις τυποποιημένες εμπορικές αλυσίδες ψηφιακών αγαθών ή τους προμηθευτές ηλεκτρονικών υπολογιστών και τους παρόχους υπηρεσιών Διαδικτύου. Όπως παρατήρησαν οι κοινωνικοί επιστήμονες Ντάνιελ Μίλερ και Ντον Σλέιτερ στη μελέτη τους για το Τρινιντάντ, «το Διαδίκτυο δεν είναι ένας μονολιθικός κυβερνοχώρος» (Miller & Slater, 2001), αλλά αντίθετα υπάρχει ως μια παγκοσμίως εκτεταμένη τεχνολογία με διάφορες τοπικές πραγματικότητες, πρακτικές υιοθέτησης και πολιτισμικές πολιτικές που περιβάλλουν τις ποικίλες τοπικές εκδοχές του. Υπάρχουν, πράγματι, περισσότεροι από ένας τρόποι για να φανταστούμε πώς θα μπορούσε να είναι η ψηφιακή πρακτική και η ψηφιακή σύνδεση.

Στο Περού, ήταν εύκολα ορατές οι ενδείξεις μιας ζωντανής ψηφιακής κουλτούρας που έφερε μια σειρά διαφορετικών φορέων και συμφερόντων σε απροσδόκητη και συχνά αντιφατική εγγύτητα. Συλλογικότητες, που φαινομενικά υποστήριζαν το ελεύθερο λογισμικό και οι οποίες είχαν βοηθήσει να πραγματοποιηθεί το 2003 στην αρχαία πρωτεύουσα των Ίνκας, το Κούσκο, το πρώτο συνέδριο για τη χρήση του ελεύθερου λογισμικού στη Λατινική Αμερική υπό την αιγίδα του ΟΗΕ (ένα γεγονός ορόσημο), προσπάθησαν να επαναπροσδιορίσουν την υιοθέτηση των ανοικτών τεχνολογιών. Επιδίωξαν να την επαναπροσδιορίσουν ως ζήτημα όχι μόνο ατομικής ελευθερίας και ελεύθερης επιλογής, όπως ήταν για τους υποστηρικτές του ελεύθερου λογισμικού και του λογισμικού ανοιχτού κώδικα («free/libre and open-source software», FLOSS) στις Ηνωμένες Πολιτείες, αλλά ως ζήτημα πολιτιστικής ποικιλομορφίας, κρατικής διαφάνειας και πολιτικής κυριαρχίας από τη μονοπωλιακή δύναμη των πολυεθνικών εταιρειών στον Παγκόσμιο Νότο.

Οι αίθουσες διδασκαλίας «ψηφιακής καινοτομίας», που εγκαταστάθηκαν σε αγροτικά σχολεία από το κράτος, θα μετατρέπονταν, μόλις μερικά χρόνια αργότερα, στο μεγαλύτερο δίκτυο χώρων ανάπτυξης για την πρωτοβουλία κύρους του Ινστιτούτου Τεχνολογίας της Μασαχουσέτης, (MIT), *One Laptop per Child* («Ενας Φορητός Υπολογιστής Ανά Παιδί», OLPC) –και όλα αυτά στο όνομα της καθολικής ψηφιακής ενσωμάτωσης. Και οι τίτλοι πνευματικής ιδιοκτησίας, που χρησιμοποιήθηκαν πρόσφατα και επιθετικά από τα κρατικά προγράμματα για τα «παραδοσιακά» αγαθά, υπόσχονταν να μετατρέψουν τους παραγωγούς και τους τεχνίτες της υπαίθρου σε νέες τάξεις «εργαζομένων της πληροφορίας», έτοιμες για εξαγωγές, στο πλαίσιο των αυξανόμενων πρωτοβουλιών της χώρας που βασίζονται στην κοινωνία της πληροφορίας.

Οι υποστηρικτές του FLOSS και οι ακτιβιστές της υψηλής τεχνολογίας στο Κούσκο, οι «αίθουσες καινοτομίας» που προωθούνται από το κράτος στα αγροτικά σχολεία και οι παραδοσιακοί τεχνίτες ως νέοι παγκόσμιοι «εργάτες της πληροφορίας» δεν ήταν τα συμβατικά συμφέροντα, ούτε οι πρωταγωνιστές στις περισσότερες ιστορίες που γράφτηκαν στα κέντρα του ψηφιακού πολιτισμού. Παρακολουθώντας τις ιστορίες τους να ξετυλίγονται ήταν σαν να βλέπεις τις λεπτομέρειες της καθεμιάς να ξεπερνούν τα όρια των υφιστάμενων πλαισίων και των κυρίαρχων αφηγήσεων του ψηφιακού πολιτισμού. Οι παγκόσμιες φαντασιώσεις σχετικά με την πληροφορική στη νέα χιλιετία, άλλωστε, έχουν κάνει τους χάκερ της Σίλικον Βάλεϊ, τις εμμονές και τις φιλοδοξίες των μηχανικών της υψηλής τεχνολογίας και τη στρατηγική επιχειρηματικότητα των ανταγωνιστικών επιχειρηματιών της τεχνολογίας, αντικείμενο δημοφιλών ταινιών του Χόλιγουντ και λογαριασμών του Twitter με φανατικούς ακολούθους. Πρόκειται για ένα καστ ολοένα και πιο αναγνωρίσιμων ηθοποιών, ηρώων και αντιηρώων. Όμως, για να αποτυπωθούν οι δυναμικές αλληλεπιδράσεις και τα φορτισμένα πειράματα στον ψηφιακό πολιτισμό στο Περού, απαιτείται προσοχή σε πλήθος άλλων θεμάτων που διακυβεύονται, παραγόντων και εξελίξεων –που, λειτουργώντας στο πλαίσιο της ψηφιακότητας, προσπάθησαν να οικοδομήσουν νέους συνδέσμους και ανταλλαγές μεταξύ της υπαίθρου και της πόλης, της υψηλής τεχνολογίας και της παράδοσης, και διαφορετικούς προσανατολισμούς γύρω από το παγκόσμιο, με ισχυρές δεσμεύσεις απέναντι στο τοπικό.

Οι δημοσιογράφοι δεδομένων έχουν σήμερα στη διάθεσή τους ένα αυξανόμενο πλήθος ψηφιακών εργαλείων και τεχνολογικών πόρων για να παρακολουθήσουν, να καταγράψουν και να επαναφέρουν στη μνήμη ψηφιακούς πολιτισμούς και δραστηριότητες σε διάφορες τοποθεσίες σε ολόκληρο τον κόσμο. Ακόμη και πριν από το κύμα κοινωνικών διαμαρτυριών στη Μέση Ανατολή που ξεκίνησε στις αρχές του 2011, τα δικτυακά ψηφιακά μέσα προσέφεραν νέες δυνατότητες παγκόσμιας μετάδοσης για κινήματα που υιοθέτησαν στρατηγικές χρήσεις των μέσων κοινωνικής δικτύωσης σε διαφορετικά πλαίσια όπως το Μεξικό (Schulz, 2007), το Ιράν (Burns & Eltham, 2009· Grossman, 2009), οι Φιλιππίνες (Uy-Tioco, 2003· Vicente, 2003) και η Ουκρανία (Morozov, 2010).

Στον απόηχο της Αραβικής Άνοιξης το 2011, κινήματα –από τους «Αγανακτισμένους» της Ισπανίας μέχρι το βορειοαμερικανικό Occupy– χρησιμοποίησαν στρατηγικά την οργάνωση και τον ακτιβισμό δια των hashtag στις πλατφόρμες κοινωνικής δικτύωσης. Πιο πρόσφατα, κινήματα

από τις ΗΠΑ ξεκίνησαν το #MeToo και το #BlackLivesMatter, ενώ παράλληλα διεξάγονται παγκόσμιες κινητοποιήσεις –το #NiUnaMenos της Λατινικής Αμερικής, το #BringBackOurGirls της Νιγηρίας, το #Sosblakaustralia της Αυστραλίας, τις κινητοποιήσεις #IdleNoMore των Πρώτων Εθνών του Καναδά και #UmbrellaRevolution του Χονγκ Κονγκ.

Οι διευρυνόμενες ροές ενημέρωσης, που δημιουργούνται από τους χρήστες, αυτών των κινημάτων πολλαπλασιάζουν τις πρακτικές της κοινωνίας των πολιτών για τα δεδομένα και αποκεντρώνουν τις κυρίαρχες εφαρμογές των «μεγάλων δεδομένων» («big data») στις πλατφόρμες κοινωνικής δικτύωσης που είναι στραμμένες προς τη σκιαγράφηση των χρηστών με γνώμονα την αγορά. Αντίθετα, αξιοποιούν πρακτικές δεδομένων για νέες μορφές αφηγηματικής δυνατότητας που ξεφεύγουν από τα καθιερωμένα κέντρα των μέσων ενημέρωσης και της παραγωγής ειδήσεων, ενώ παραχωρούν εξίσου τα αρχεία δεδομένων τους –και τα διαδικτυακά αποδεικτικά στοιχεία για το παγκόσμιο κοινό τους– σε γεωγραφικά διασκορπισμένους ντοκιμαντερίστες, δημοσιογράφους και διοργανωτές.

Ωστόσο, η αύξηση των ψηφιακών πόρων και των αποθετηρίων δεδομένων –από τα διαδικτυακά αρχεία «δεδομένων» των κοινωνικών κινημάτων στις πλατφόρμες κοινωνικής δικτύωσης έως τις παράλληλες μορφές δημιουργικού ακτιβισμού δεδομένων– δημιουργεί, επίσης, νέους κινδύνους για τους δημοσιογράφους δεδομένων. Ο κυριότερος από αυτούς είναι ο κίνδυνος από τη σαγηνευτική ικανότητα των μεγάλων δεδομένων και των μέσων κοινωνικής δικτύωσης να αξιοποιούν την αφθονία των δεδομένων και των πληροφοριών που συλλέγουν ως ένα μέσο για να πείσουν τα ακροατήρια ότι η εκτεταμένη παρακολούθηση των δεδομένων τους συγκεντρώνει και δημιουργεί την καλύτερη δυνατή μορφή τεκμηρίωσης της παρούσας ανθρώπινης δραστηριότητας και κοινωνικής εμπειρίας –καθώς και την αξιολόγηση και πρόβλεψη των μελλοντικών πολιτικών ή οικονομικών συνεπειών τους.

Η χρονική αυθαιρεσία της προβολής που κάνουν οι ψηφιακοί οικουμενιστές, ότι οι μορφές του ψηφιακού «παρόντος» που καλλιεργούνται σήμερα στα κέντρα καινοτομίας μπορούν και θα αντιπροσωπεύουν με ακρίβεια τα ψηφιακά μέλλοντα των παγκόσμιων περιφερειών, βρίσκει ένα νέο συμπλήρωμα στους γεμάτους αυτοπεποίθηση ισχυρισμούς των βιομηχανιών δεδομένων για τις προγνωστικές ικανότητες της αλγοριθμικής επεξεργασίας δεδομένων. Τέτοιες διακηρύξεις παραμένουν, ακόμη και παρά τις προφανείς σύγχρονες αποτυχίες των κυρίαρχων αναλυτών πολιτικών δεδομένων, των εταιρειών κοινωνικής δικτύωσης και των ειδησιολόγων στη Δύση να προβλέψουν με ακρίβεια τις μεγάλες παγκόσμιες πολιτικές ανατροπές των τελευταίων ετών –από τις προεδρικές εκλογές του 2016 στις ΗΠΑ, το Brexit και το σκάνδαλο Cambridge Analytica έως την «αιφνιδιαστική» άνοδο των κινημάτων «alt-right»[208] σε όλη τη Δύση.

Ανεξάρτητα από το πόσο εκτεταμένα και εντυπωσιακά μεγάλα μπορεί να είναι τα σύνολα δεδομένων, οι σημερινοί δημοσιογράφοι δεδομένων θα πρέπει να επαγρυπνούν και να είναι επιφυλακτικοί στο να καταστήσουν τις καταγραφές και τα αρχεία δεδομένων τη μοναδική ή την κυρίαρχη μορφή τεκμηρίωσης, εκπροσώπησης και αξιολόγησης των ποικίλων μορφών κοινωνικής πραγματικότητας που το κοινό περιμένει από τους ίδιους να μεταδώσουν. Παράλληλα με τις αυξανόμενες εκκλήσεις από τους μελετητές της Λατινικής Αμερικής και της μετα-αποικιοκρατίας για επέκταση των μεθόδων έρευνας και τεκμηρίωσης, προκειμένου να διευρυνθεί η έννοια του τι και ποιος αντιπροσωπεύει τις πληροφορίες, την τεχνολογία και τις κουλτούρες των νέων μέσων ενημέρωσης στο πλαίσιο μιας «αποαποικιακής πληροφορικής» («decolonial computing» · Chan, 2018), οι δημοσιογράφοι δεδομένων που ασκούν κριτική στα ψηφιακά

208. ΣτΕ: Συντομογραφία της έκφρασης «alternative right», που συχνά αποδίδεται ως «εναλλακτική Δεξιά». Συνοπτικά, η ιδέα της «λευκής φυλής» η οποία αυτοδιαβρώνεται στον φιλελεύθερο, παγκοσμιοποιημένο, πολυπολιτισμικό κόσμο θα μπορούσε να περιγραφεί ως ο ιδεολογικός πυρήνας του αφηγήματος της «alt-right» που συνδέεται με τον εθνικισμό, τον φασισμό, τον νεο-ναζισμό, τον αντιδραστισμό και έχει βρει τα μέσα έκφρασής της στο Διαδίκτυο.

οικουμενικά πλαίσια θα πρέπει, επίσης, να στοχεύουν στη συνειδητή διαφοροποίηση των πηγών δεδομένων και στην αποκέντρωση των μεθόδων που θα ευνοούσαν τα «μεγάλα δεδομένα» («big data») ως το αποκλειστικό ή το πιο θεμιτό κλειδί για τη χαρτογράφηση εμπειρικών γεγονότων και κοινωνικών πραγματικοτήτων. Οι κινήσεις προς μια «αποαποικιοποίηση της γνώσης» («decolonization of knowledge») υπογραμμίζουν τη σημασία των ποικίλων τρόπων με τους οποίους οι πολίτες και οι ερευνητές στον Παγκόσμιο Νότο εμπλέκονται σε πρακτικές δεδομένων από τα κάτω[209]. Οι τελευταίες δίνουν έμφαση στις πρακτικές της κοινότητας και στα ανθρωποκεντρικά μέσα αξιολόγησης και ερμηνείας των δεδομένων για την κοινωνική αλλαγή, καθώς και στις αντιστάσεις απέναντι στη χρήση των μεγάλων δεδομένων που αυξάνουν την καταπίεση, την ανισότητα και την κοινωνική επιβάρυνση.

Οι δημοσιογράφοι δεδομένων που ασκούν κριτική στις νέες προεκτάσεις του ψηφιακού καθολικισμού προς την οικουμενικότητα των δεδομένων θα πρέπει με θάρρος να βρουν συμμάχους και ηχηρές φωνές, για την ανάπτυξη υπεύθυνων πρακτικών δεδομένων, ανοιχτών στη λογοδοσία –μαζί με μελετητές των κριτικών σπουδών δεδομένων, των σπουδών αλγορίθμων, των σπουδών λογισμικού και πλατφορμών αλλά και της μετα-αποικιακής πληροφορικής. Αυτό περιλαμβάνει μια ενισχυμένη απόρριψη του φονταμενταλισμού των δεδομένων (Boyd & Crawford, 2012) και του τεχνολογικού ντετερμινισμού που εξακολουθεί να περιβάλλει τις κύριες θεωρήσεις των αλγορίθμων στην πράξη. Αυτό συνεπάγεται, επίσης, έναν θεμελιώδη επαναπροσδιορισμό του ανθρώπου μέσα στους κόσμους των δεδομένων και στις βιομηχανίες δεδομένων –που αντιστέκεται στην παρόρμηση να διαβαστούν τα μεγάλα δεδομένα και «οι αλγόριθμοι ως φετιχιστικά αντικείμενα [...] και αντιστέκεται σθεναρά στην τοποθέτηση της τεχνολογίας στη θέση του καθοδηγητή των εξηγήσεων» (Crawford, 2016). Περιλαμβάνει, επίσης, την αντιμετώπιση των υποδομών δεδομένων και των υποκείμενων αλγορίθμων, που τους δίνουν πολιτική ζωή σκόπιμα, ως διφορούμενων και προσιτών ταυτόχρονα –για την ανάπτυξη μεθοδολογιών που όχι μόνο διερευνούν νέες εμπειρικές και καθημερινές συνθήκες για την πολιτική των δεδομένων (είτε πρόκειται για την ασφάλεια των αεροδρομίων, είτε για την πιστοληπτική βαθμολόγηση, την παρακολούθηση νοσοκομείων και ασθενών ή τα μέσα κοινωνικής δικτύωσης σε ποικίλους παγκόσμιους ιστότοπους), αλλά και βρίσκουν δημιουργικούς τρόπους για να καταστήσουν τα δεδομένα παραγωγικά προς ανάλυση (Gillespie, 2013· Ziewitz, 2016).

Τέλος, αξίζει ίσως να υπενθυμίσουμε ότι η διατήρηση των συγκεκριμένων κέντρων της ψηφιακής καινοτομίας ως αποκλειστικών τόπων της ψηφιακής επινόησης ή της δημιουργίας του μέλλοντος των δεδομένων, φυσικά, παραγνωρίζει, επίσης, μια άλλη κρίσιμη λεπτομέρεια: ότι τα κέντρα του παρόντος βρίσκονταν κάποτε επίσης στην περιφέρεια. Η εστίαση στα κέντρα ως εφευρέτες μοντέλων που απλώς έρχονται να υιοθετηθούν και να αντιγραφούν κάπου αλλού προϋποθέτει την τέλεια, συνεχή επέκταση των αναπαραγωγικών λειτουργιών και δυνάμεων. Αποτυγχάνει να λάβει υπόψη του τη δυνατότητα αλλαγής εντός του ευρύτερου συστήματος (τις αποσταθεροποιήσεις και τις ανακατατάξεις των παλαιότερων κέντρων) και, επίσης, τις ανακατατάξεις των παλαιότερων περιφερειών.

Η «έκπληξη» της Αραβικής Άνοιξης το 2011 και η επιρροή της σε ένα ευρύ φάσμα παγκόσμιων τοποθεσιών στη Δύση και σε μη δυτικές χώρες, όπως και η πρόσφατη άνοδος των μη δυτικών ψηφιακών αγορών και των οικονομικών ανταγωνιστών σε έθνη που χαρακτηρίζονταν «αναπτυσσόμενα» πριν από λιγότερο από δύο δεκαετίες, καθώς και η αποσταθεροποίηση των ισχυρών δυτικών δημοκρατιών σήμερα, υπενθυμίζουν ότι η σταθερότητα των καθιερωμένων δυνάμεων και η μονιμότητα των σχέσεων κέντρου-περιφέρειας μπορεί να τεθεί υπό αμφισβήτηση. Μακριά από την απλή υστέρηση ή τη μίμηση των κέντρων, οι δυναμικές δραστηριότητες από την περιφέρεια υποδηλώνουν πώς οι παράγοντες που κάποτε κατείχαν δευτερεύουσα θέση μπορούν να

209. Για περισσότερα σχετικά με αυτό το θέμα, βλ. το κεφάλαιο των Κουκουτάι και Γουόλτερ σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 3.3, «Κυριαρχία των Αυτοχθόνων επί των Δεδομένων τους: Επιπτώσεις για τη Δημοσιογραφία Δεδομένων»).

αναδειχθούν ως νέες πηγές ξεχωριστής παραγωγικότητας. Τα ποικίλα μοτίβα τους ανατρέπουν την άδηλη υπόθεση ότι μια ενιαία, οικουμενική αφήγηση θα μπορούσε να αναπαραστήσει επαρκώς τα ξεχωριστά ψηφιακά μέλλοντα και τις φαντασιώσεις που αναδύονται σήμερα σε μια σειρά από περιφερειακές τοποθεσίες.

Παραπομπές

Boyd, D., & Crawford, K. (2012) «Critical questions for big data», *Information, Communication & Society,* 15(5), σσ. 662–679. DOI: https://doi.org/10.1080/1369118X.2012.678878

Burns, A., & Eltham, B. (2009) «Twitter free Iran: An evaluation of Twitter's role in public diplomacy and information operations in Iran's 2009 election crisis», στο P. Franco & M. Armstrong (επιμ.), *Record of the Communications Policy & Research Forum 2009*, 19-20 Νοεμβρίου. Σίδνεϊ, Αυστραλία: Network Insight Institute, σσ. 322–334.

Chan, A. (2018) «Decolonial computing and networking beyond digital universalism», *Catalyst*, 4(2). DOI: https://doi.org/10.28968/cftt.v4i2.29844

Crawford, K. (2016) «Can an algorithm be agonistic? Ten scenes from life in calculated publics», *Science, Technology, & Human Values,* 41(1), σσ. 77–92. DOI: https://doi.org/10.1177/0162243915589635

Gillespie, T. (2013) «The relevance of algorithms», στο T. Gillespie, P. J. Boczkowski, K. A. Foot (επιμ.), *Media technologies: Essays on communication, materiality, and society.* Κέιμπριτζ, Μασαχουσέτη: MIT Press, σσ. 167–193.

Grossman, L. (2009) «Iran's protests: Why Twitter is the medium of the movement» *Time*, 17 Ιουνίου. Διαθέσιμο σε: http://content.time.com/time/world/article/0,8599,1905125,00.html (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Miller, D., & Slater, D. (2001). *The Internet: An ethnographic approach.* Οξφόρδη: Berg Publishers.

Morozov, E. (2010). *The net delusion: The dark side of Internet freedom*. Νέα Υόρκη: Public Affairs.

Schulz, M. (2007) «The role of the Internet in transnational mobilization: A case study of the Zapatista movement, 1994–2005», στο M. Herkenrath (επιμ.), *Civil society: Local and regional responses to global challenges*, Λονδίνο & Νέα Υόρκη: Transaction Publishers, σσ. 129–156.

Uy-Tioco, C. (2003, October 11) «The cell phone and EDSA 2: The role of a communication technology in ousting a president», *Critical Themes in Media Studies Conference,* Νέα Υόρκη: New School University, 11 Οκτωβρίου.

Vicente, R. (2003) «The cell phone and the crowd: Messianic politics in the contemporary Philippines», *Public Culture,* 24(47), σσ. 3–36. DOI: https://doi.org/10.1080/01154451.2003.9754246

Ziewitz, M. (2016) «Governing algorithms: Myth, mess, and methods», *Science, Technology, & Human Values*», 41(1), σσ. 3–16. DOI: https://doi.org/10.1177/0162243915608948

Η συγγραφέας

Η Ανίτα Σάι Τσαν είναι Αναπληρώτρια Καθηγήτρια στη Σχολή Επιστημών της Πληροφορίας και στο Κολλέγιο Μέσων Ενημέρωσης και συνεργάτης του Εθνικού Κέντρου Εφαρμογών Υπερυπολογισμού στο Πανεπιστήμιο του Ιλινόις, Urbana-Champaign, καθώς και συνεργάτης του Ερευνητικού Ινστιτούτου Δεδομένων και Κοινωνίας στη Νέα Υόρκη.

8.2

Δεδομενοποίηση στη Δημοσιογραφία: Στρατηγικές για Αφήγηση με Δεδομένα και Συνεργασία Βιομηχανίας-Ακαδημαϊκής Κοινότητας

Damian Radcliffe, Seth C. Lewis

Περίληψη: Πώς ανταποκρίνονται η δημοσιογραφία και η ακαδημαϊκή κοινότητα στη δεδομενοποίηση των επαγγελμάτων τους και πώς μπορούν να συνεργαστούν αποτελεσματικότερα σε εργασίες που βασίζονται σε δεδομένα;

Λέξεις-κλειδιά: δημοσιογραφία, ακαδημαϊκή κοινότητα, συνεργασία, δεδομενοποίηση, εργασία δεδομένων, συνεργασίες ερευνητών-δημοσιογράφων

Ζούμε σε έναν κόσμο που καθοδηγείται και τροφοδοτείται από δεδομένα. Τα δεδομένα επηρεάζουν όλο και περισσότερο τον τρόπο με τον οποίο λαμβάνονται πολιτικές αποφάσεις (Höchtl κ.α., 2016· Kreiss, 2016), επιδρούν στον σχεδιασμό και στη λειτουργικότητα των πόλεων που κατοικούμε (Glaeser κ.α., 2018) και παράλληλα διαμορφώνουν τα είδη των ειδήσεων, των προϊόντων και των πληροφοριών στα οποία έχουμε πρόσβαση –και τα οποία καταναλώνουμε– στην ψηφιακή εποχή (Diakopoulos, 2019· Lewis, 2017· Lewis & Westlund, 2015· Thurman κ.α., 2019· Usher, 2016· Zamith, 2018). Η πλήρης δύναμη και οι δυνατότητες των δεδομένων, για το καλό ή το κακό, μόλις τώρα αρχίζουν να γίνονται κατανοητές (Couldry & Mejias, 2019· O'Neil, 2016· Schroeder, 2018).

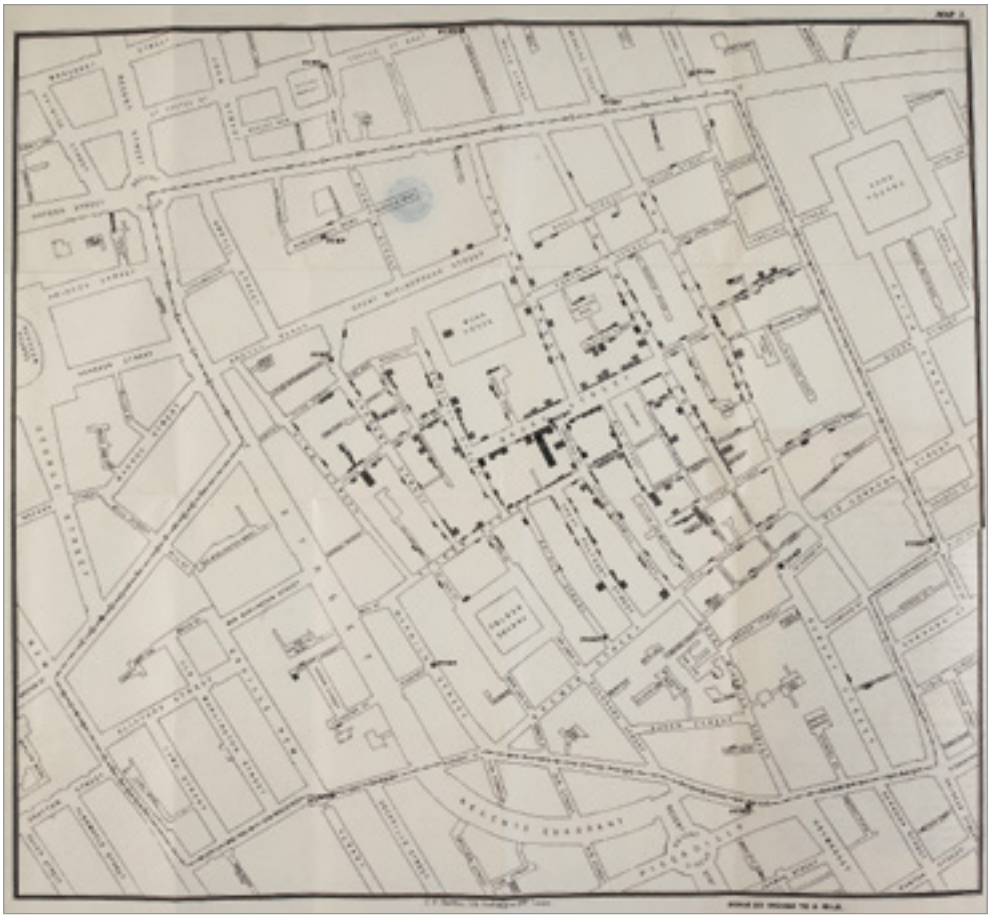
Οι κυβερνήσεις, τα πανεπιστήμια και τα ΜΜΕ χρησιμοποιούν εδώ και καιρό δεδομένα και στατιστικές, ψάχνοντας να βρουν μοτίβα για να εξηγήσουν τον κόσμο. Όμως, με την ανάπτυξη των ψηφιακών συσκευών και των μαζικών δεδομένων που αυτές παράγουν –για τα κλικ, τα like, τις κοινοποιήσεις, τις τοποθεσίες, τις επαφές και πολλά άλλα–, ο όγκος των δεδομένων που παράγονται είναι εντυπωσιακός, όπως και η αύξηση της υπολογιστικής ισχύος για την αξιοποίηση και την ανάλυση αυτών των δεδομένων σε μεγάλη κλίμακα. Το να βγάλει κανείς νόημα από όλα αυτά τα δεδομένα, σε πολλές περιπτώσεις, είναι αναμφισβήτητα η μεγαλύτερη πρόκληση και είναι βαθύτατα συνυφασμένη με ηθικούς προβληματισμούς (Crawford κ.α., 2014). Είναι ένας γρίφος που απασχολεί τους υπεύθυνους χάραξης πολιτικής, τις επιχειρήσεις, τους ερευνητές, τους ακτιβιστές, τους δημοσιογράφους και άλλους –και ο οποίος δεν θα λυθεί τόσο εύκολα με «λύσεις μεγάλων δεδομένων» («big-data solutions») ή, όπως είναι στη μόδα σήμερα, με τη λαμπερή υπόσχεση της τεχνητής νοημοσύνης (Broussard, 2018· Broussard κ.α., 2019).

Σε αυτό το κεφάλαιο, βασιζόμενοι στις αντίστοιχες παρατηρήσεις μας για την πρακτική (Radcliffe) και την έρευνα (Lewis) σχετικά με τα δεδομένα και τη δημοσιογραφία, περιγράφουμε τον τρόπο με τον οποίο ο κόσμος της δημοσιογραφίας και ο ακαδημαϊκός κόσμος ανταποκρίνονται στη δεδομενοποίηση των επαγγελμάτων τους καθώς και στην ευρύτερη δεδομενοποίηση της δημόσιας ζωής. Τελικά, ο στόχος μας είναι να διατυπώσουμε προτάσεις για το πώς αυτοί οι δύο τομείς, οι οποίοι ιστορικά είχαν μια μάλλον δύσκολη σχέση (Carlson & Lewis, 2019· Reese, 1999), θα μπορούσαν

να συνεργαστούν πιο παραγωγικά για προκλήσεις με επίκεντρο τα δεδομένα. Ο ποιητής Τζον Ντον έγραψε ότι «κανένας άνθρωπος δεν είναι νησί» («no man is an island»)[210]. Σε έναν κόσμο που καθορίζεται α πό τα δεδομένα, κανένα επάγγελμα δεν θα έπρεπε να είναι νησί.

Δημοσιογραφία και αφήγηση με βάση τα δεδομένα: πέντε στρατηγικές θεωρήσεις

Η χρήση δεδομένων για την αφήγηση ιστοριών και την κατανόηση του κόσμου γύρω μας δεν είναι εντελώς καινούρια[211]. Στη βικτοριανή Αγγλία, ο γιατρός Τζον Σνόου σχεδίασε έναν χάρτη που απεικόνιζε τα κρούσματα χολέρας στο κεντρικό Λονδίνο. Αυτό του επέτρεψε να εντοπίσει μια αντλία νερού στην Μπρόουντ Στριτ ως την πηγή ενός ιδιαίτερα θανατηφόρου, και γεωγραφικά εστιασμένου, ξεσπάσματος της νόσου το 1855 (Εικόνες 8.2.1 και 8.2.2). Η σημαίνουσα ανάλυση του Σνόου δεν διαφέρει πολύ από τους επιδημιολογικούς χάρτες που παράγονται με τα σύγχρονα εργαλεία ανάλυσης και οπτικοποίησης δεδομένων.



Εικόνα 8.2.1: Χάρτης του Λονδίνου που δημιουργήθηκε από τον γιατρό Τζον Σνόου και με τον οποίο οπτικοποιούνται τα κρούσματα χολέρας στο κεντρικό Λονδίνο το 1855. Πηγή: Βρετανική Βιβλιοθήκη.

210. ΣτΕ: Αναφορά στο ποίημα *No Man is an Island*

211. Για μια ματιά στις διαφορετικές γενεαλογίες της δημοσιογραφίας δεδομένων, βλ. το κεφάλαιο του Άντερσον σε αυτό το βιβλίο (ΣτΕ: Κεφάλαιο 9.1, «Γενεαλογίες της Δημοσιογραφίας Δεδομένων»).

therefore, determined to return to Baljik, taking with him the *Trafalgar* and *Albion*, also badly affected.

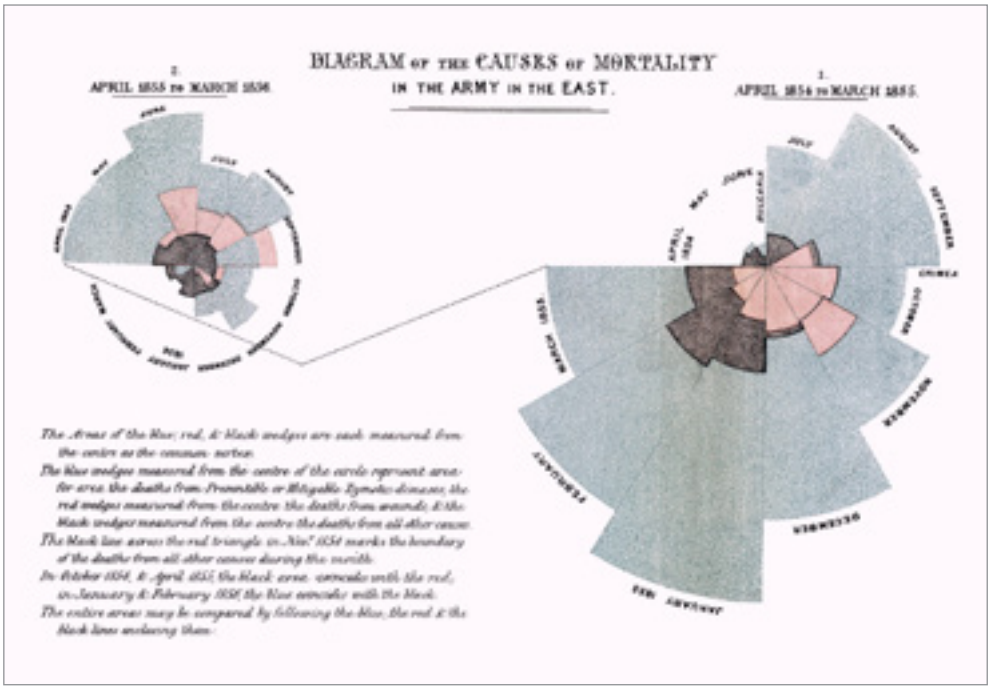
"The crew of the *Britannia* were at once sent away from the ship, in small parties, into the numerous transports that remained idle; and it appears that, by this procedure, the epidemic influences operating among them have been greatly moderated, if not extirpated."

The most terrible outbreak of cholera which ever occurred in this kingdom, is probably that which took place in Broad Street, Golden Square, and the adjoining streets, a few weeks ago. Within two hundred and fifty yards of the spot where Cambridge Street joins Broad Street, there were upwards of five hundred fatal attacks of cholera in ten days. The mortality in this limited area probably equals any that was ever caused in this country, even by the plague; and it was much more sudden, as the greater number of cases terminated in a few hours. The mortality would undoubtedly have been much greater had it not been for the flight of the population. Persons in furnished lodgings left first, then other lodgers went away, leaving their furniture to be sent for when they could meet with a place to put it in. Many houses were closed altogether, owing to the death of the proprietors; and, in a great number of instances, the tradesmen who remained had sent away their families: so that in less than six days from the commencement of the outbreak, the most afflicted streets were deserted by more than three-quarters of their inhabitants.

There were a few cases of cholera in the neighbourhood of Broad Street, Golden Square, in the latter part of August; and the so-called outbreak, which commenced in the night between the 31st August and the 1st September, was, as in all similar instances, only a violent increase of the malady. As soon as I became acquainted with the

Εικόνα 8.2.2: Κείμενο από θέμα σε εφημερίδα του 1855 που τεκμηριώνει τα κρούσματα χολέρας στο κεντρικό Λονδίνο. Πηγή: Βρετανική Βιβλιοθήκη.

Σε ένα άλλο παράδειγμα, οι οπτικοποιήσεις της Φλόρενς Νάιτινγκεϊλ «για τα αίτια της θνησιμότητας στον στρατό της Ανατολής» (The Economist, 2013) βοήθησαν στην κατάδειξη του ρόλου που έπαιξε η υγιεινή (ή η έλλειψή της) ως αιτία θανάτου των Βρετανών στρατιωτών που πολεμούσαν στον Πόλεμο της Κριμαίας (Εικόνα 8.2.3). Τα σχέδιά της εξακολουθούν να φαίνονται εξαιρετικά σύγχρονα.



Εικόνα 8.2.3: «Διάγραμμα για τα αίτια θνησιμότητας στον στρατό της Ανατολής», από τη Φλόρενς Νάιτινγκεϊλ. Πηγή: Wikimedia.

Παράλληλα με αυτές τις προσπάθειες, περίπου την ίδια εποχή, στα μέσα του 19ου αιώνα, η δουλειά του Χόρας Γκρίλεϊ για την εφημερίδα *New York Tribune* αποκάλυψε ότι ορισμένοι εκλεγμένοι αξιωματούχοι (συμπεριλαμβανομένου του νεαρού Αβραάμ Λίνκολν) αξίωναν αποζημίωση για μεγαλύτερες δαπάνες από αυτές που δικαιούνταν (Klein, 2015). Παρόλο που ο κόσμος έχει προχωρήσει (το έργο του Γκρίλεϊ εστίαζε σε αποστάσεις που συνήθως διανύονταν με άλογα), αυτό το είδος σημαντικής ερευνητικής εργασίας συνεχίζει να αποτελεί βασικό δημοσιογραφικό εργαλείο (Larcinese & Sircar, 2017· βλ. επίσης Barrett, 2009· CBC News, 2016· Reuters, 2017· Telegraph Reporters, 2009· BBC News , 2009· Rayner, 2009· HuffPost Canada, άνευ ημερομηνίας).

Αυτά τα ιστορικά παραδείγματα, σε συνδυασμό με πιο σύγχρονες περιπτωσιολογικές μελέτες (όπως αυτές που αναδεικνύονταν από τα ετήσια Data Journalism Awards), μπορούν να αποτελέσουν ισχυρές πηγές έμπνευσης για τους δημοσιογράφους[212]. Δείχνουν τον τρόπο με τον οποίο οι προσεγγίσεις που βασίζονται στα δεδομένα μπορούν να χρησιμοποιηθούν για να

212. Δείτε, επίσης, την παρουσίαση της Λούζεν για τα βραβεία στο κεφάλαιό της σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 9.2, «Χρυσά Πρότυπα Με Γνώμονα τα Δεδομένα: Τι Αποτιμά ο Κλάδος ως Δημοσιογραφία Δεδομένων που αξίζει Βραβείο και Πώς η Δημοσιογραφία Δεδομένων Εξελίσσεται μαζί με τη Δεδομενοποίηση της Κοινωνίας»).

λογοδοτήσουν οι Αρχές (ICIJ, άνευ ημερομηνίας), να υπογραμμίσουν σημαντικές κοινωνικές αδικίες (Lowenstein, 2015), καθώς και να οπτικοποιήσουν και να αναδείξουν κάτι ξεχωριστό (Los Angeles Times, 2016).

Ενώ τα δεδομένα αποτελούν από καιρό τμήμα της δημοσιογραφίας, όπως προκύπτει από την εμφάνιση του «υποβοηθούμενου από υπολογιστή ρεπορτάζ» («computer-assisted reporting») στα τέλη του 20ού αιώνα, οι πρόσφατες εξελίξεις στη διαθεσιμότητα και στην προσβασιμότητα των τεχνικών που βασίζονται στα δεδομένα έχουν ενισχύσει τις ευκαιρίες για μια δημοσιογραφία που βασίζεται αποκλειστικά στα δεδομένα (για μια ιστορική αναδρομή, βλ. Anderson, 2018· για μια επισκόπηση της δημοσιογραφίας δεδομένων, βλ. Gray κ.α., 2012). Σε αυτό το πλαίσιο, οι ειδησεογραφικοί οργανισμοί σε ολόκληρο τον κόσμο –ιδιαίτερα οι οργανισμοί με τους περισσότερους πόρους, αλλά και ολοένα και μικρότεροι ειδησεογραφικοί οργανισμοί– χρησιμοποιούν τα δεδομένα για να τεκμηριώνουν το δημοσιογραφικό τους έργο, τόσο στην αφήγηση ιστοριών με δεδομένα (Hermida & Young, 2019) όσο και στη χρήση δεδομένων (με τη μορφή ψηφιακών μετρήσεων κοινού), για να επηρεάσουν την επιλογή ιστοριών, καθώς και για να μετρήσουν και να βελτιώσουν τον αντίκτυπο του έργου τους (Zamith, 2018).

Ακολουθούν πέντε βασικά μηνύματα για τις αίθουσες σύνταξης και τους δημοσιογράφους που επιθυμούν να κάνουν περισσότερα πράγματα με τα δεδομένα:

Τα δεδομένα από μόνα τους δεν λένε ιστορίες. Χρειαζόμαστε ακόμα τους δημοσιογράφους. Παρά το πλήθος των δεδομένων στα οποία έχουμε πρόσβαση, εξακολουθούμε να χρειαζόμαστε δημοσιογράφους για να τα αξιοποιήσουν, παρέχοντας το πλαίσιο και εξετάζοντάς τα με τον ίδιο τρόπο όπως κάθε άλλη πηγή.

Όπως σημειώνει ο Στιβ Τζόνσον του Πανεπιστημίου Montclair State: «Οι αναγνώστες δεν ενδιαφέρονται για τα ακατέργαστα δεδομένα. Θέλουν την ιστορία μέσα από τα δεδομένα» (Johnson, 2013). Σχολιάζοντας δεδομένα για το κάτω Μανχάταν που παρείχε μια από τις πρώτες πύλες ανοιχτών δεδομένων, το *EveryBlock*, ο ίδιος παρατηρεί:

> Υπήρχαν ρεπορτάζ σχετικά με τα γκράφιτι που η πόλη ανέφερε ότι έσβηνε κάθε μήνα, ανά γειτονιά. Αυτό που έλειπε, όμως, ήταν το πλαίσιο και οι φωτογραφίες. Εάν είμαι δημοσιογράφος και κάνω ένα ρεπορτάζ για το γκράφιτι, θέλω να δείξω φωτογραφίες πριν και μετά. Και, το πιο σημαντικό, θέλω να ξέρω αν η πόλη πολεμά με επιτυχία τους καλλιτέχνες του γκράφιτι, δηλαδή ποιος κερδίζει. Τα ανεπεξέργαστα δεδομένα δεν παρείχαν αυτό το πλαίσιο (Johnson, 2013).

Οι πιο πρόσφατες «εκροές δεδομένων», όπως τα *Paradise Papers* και τα *Panama Papers*, υπογραμμίζουν επίσης αυτό το στοιχείο. Σε αυτήν την περίπτωση, οι πηγές έπρεπε να διασταυρωθούν και να συσχετιστούν με τα συμφραζόμενα –μια χρονοβόρα διαδικασία για την οποία πολλοί δημοσιογράφοι χρειάστηκαν μήνες εργασίας. Ωστόσο, χωρίς αυτήν την εξέταση των πηγών από τους δημοσιογράφους (σε αντίθεση με τους ενδιαφερόμενους πολίτες), ο πλήρης αντίκτυπος των δεδομένων δεν θα μπορούσε να γίνει αντιληπτός. Οι αρχές αυτές εφαρμόζονται τόσο σε τοπικό επίπεδο όσο και σε ιστορίες εθνικής και διεθνούς σημασίας (Radcliffe, 2013).

Τα δεδομένα, από μόνα τους, σπανίως είναι η πραγματική ιστορία. Αυτή θα πρέπει να αναλυθεί και να εξηγηθούν οι συνέπειές της, εάν θέλουμε να κατανοήσουμε το πλήρες νόημά της.

Δεν χρειάζεται να το κάνετε μόνοι σας. Η συνεργασία συχνά είναι το κλειδί. Η συνεργασία αποτελεί σλόγκαν της δικτυωμένης εποχής και βασικό στοιχείο της συνεχιζόμενης ανάμειξης της δημοσιογραφικής και της τεχνολογικής ευαισθησίας –συμπεριλαμβανομένης της ενσωμάτωσης των «χάκερ»-δημοσιογράφων και των «χάκερ»-προγραμματιστών (Lewis & Usher, 2014, 2016), καθώς και της ευρύτερης αλληλεπίδρασης των ειδησεογραφικών οργανισμών και των κοινοτήτων τους γύρω από κοινές ανησυχίες. Η ουσία αυτής της «δικτυωμένης δημοσιογραφίας» (Beckett, 2010· Van der Haak κ.α., 2012) ή της «σχεσιακής δημοσιογραφίας»

(Boczkowski & Lewis, 2018· Lewis, 2019) είναι η υποκείμενη πεποίθηση ότι περισσότερα μπορούν να επιτευχθούν μέσω της συνεργατικής δραστηριότητας.

Αυτή η προσέγγιση ισχύει για πολλά είδη ρεπορτάζ και ιστορίες, συμπεριλαμβανομένων εκείνων που περιλαμβάνουν μεγάλο όγκο δεδομένων. Όπως έδειξε ο *Guardian* στην ανάλυση των εξόδων των Βρετανών βουλευτών το 2009, οι ενδιαφερόμενοι πολίτες και μέρος του κοινού μπορούν να συνεργαστούν με τους δημοσιογράφους, για να αναλύσουν σύνολα δεδομένων και να δώσουν συμβουλές (Guardian News & Media, 2009· Rogers, 2009a· Rogers, 2009b). Πιο πρόσφατα, η έρευνα του Τζέιμς Χάμιλτον του Στάνφορντ (Hamilton, 2016) και άλλων (Sambrook, 2018) εντόπισε τη σημασία της συνεργασίας –τόσο μεταξύ των οργανισμών όσο και διαφορετικών κλάδων– για πολλές αίθουσες σύνταξης, όταν πρόκειται για την παραγωγή υψηλής ποιότητας και υψηλού αντικτύπου ερευνητικής δημοσιογραφίας.

Ο όγκος των δεδομένων που αντιμετωπίζουν πολλοί νέοι οργανισμοί, σε συνδυασμό με τις συνεχιζόμενες δύσκολες οικονομικές συνθήκες, σημαίνει ότι οι συνεργασίες, η χρήση ειδικών, εθελοντών και ευρέως φάσματος δεξιοτήτων συχνά αποτελούν αναγκαιότητα για πολλές αίθουσες σύνταξης. Μάλιστα, μια συνεργατική προσέγγιση είναι όλο και πιο απαραίτητη, τόσο από οικονομική όσο και από δημοσιογραφική άποψη.

Σημασία έχει ο τρόπος παρουσίασης των δεδομένων σας. Οι δημοσιογράφοι έχουν πρόσβαση σε ένα ευρύ φάσμα εργαλείων, τεχνικών και πλατφορμών μέσω των οποίων μπορούν να παρουσιάσουν τα δεδομένα και να διηγηθούν ιστορίες αναλόγως[213]. Ως αποτέλεσμα, είναι σημαντικό να καθορίσετε ποια εργαλεία είναι τα πιο κατάλληλα για την ιστορία που προσπαθείτε να πείτε.

Οπτικοποιήσεις δεδομένων, γραφήματα, infographic, διαγράμματα, πίνακες και διαδραστικά στοιχεία –όλα μπορούν να βοηθήσουν στη μεταφορά και στην προώθηση μιας ιστορίας. Ωστόσο το στοιχείο (ή τα στοιχεία) που θα επιλέξετε να χρησιμοποιήσετε, μπορεί να κάνει τη διαφορά[214].

Όπως σημείωσε η συνάδελφός μας, Νικόλ Ντάμεν, ένας τρόπος για να γίνει αυτό είναι μέσω της οπτικοποίησης των δεδομένων: «Τα οπτικά στοιχεία τραβούν την προσοχή του κοινού... [και]... [ο κόσμος] τα επεξεργάζεται 60 φορές ταχύτερα απ' ό,τι το κείμενο» (Frank κ.α., 2015). Όταν τα παραπάνω στοιχεία χρησιμοποιούνται σωστά, μπορούν να βοηθήσουν μια ιστορία να ζωντανέψει όπως δεν μπορεί να τη ζωντανέψει το κείμενο.

Το διαδικτυακό αφιέρωμα της *Washington Post* με τίτλο *The Depth of the Problem* («Το βάθος του προβλήματος»), το οποίο δείχνει πόσο βαθιά στον ωκεανό θα μπορούσε να βρίσκεται το μαύρο κουτί από την εξαφανισμένη πτήση 370 της Malaysia Airlines, είναι ένα καλό παράδειγμα (The Washington Post, άνευ ημερομηνίας). Ο αναγνώστης κατεβαίνει μέσα στη σελίδα στα 1.250 πόδια (το ύψος του Empire State Building), στα 2.600 πόδια (το βάθος των γιγάντιων καλαμαριών), κάτω από τα 12.500 πόδια (όπου βρίσκεται ο Τιτανικός), στα 15.000 πόδια –όπου πιστεύεται ότι βρίσκεται το μαύρο κουτί.

«Δεν διαβάζετε απλώς σε ποιο βάθος βρίσκεται αυτό το αεροπλάνο», λέει η Ντάμεν. «Μπορείτε να δείτε, και να συμμετάσχετε, και να βιώσετε πραγματικά, και να καταλάβετε σε τι βάθος υποψιάζονται ότι βρίσκεται αυτό το αεροπλάνο».

Ο καθορισμός της προσέγγισής σας μπορεί να επηρεαστεί τόσο από την ιστορία που θέλετε να αφηγηθείτε όσο και από τον αλφαβητισμό σχετικά με τα δεδομένα και τις προτιμήσεις του κοινού σας. Και στις δύο περιπτώσεις, οι ιστορίες σας με βάση τα δεδομένα θα πρέπει να είναι

213. https://www.journaliststoolbox.org/2018/03/11/online_journalism/

214. Βλ. https://www.import.io/post/8-fantastic-examples-of-data-journalism/, για οκτώ ιστορίες που χρησιμοποιούν διαφορετικές τεχνικές –και εξετάστε το ενδεχόμενο ανταλλαγής.

καλά σχεδιασμένες, ώστε το κοινό να μη δυσκολεύεται να καταλάβει τι απεικονίζεται ή πώς να αλληλεπιδράσει με τα δεδομένα (Radcliffe, 2017b· Radcliffe, 2017c).

Τοποθετήστε τη δουλειά σας σε ένα ευρύτερο πλαίσιο. Παράλληλα με αυτές τις θεωρήσεις, οι δημοσιογράφοι που εργάζονται με δεδομένα πρέπει, επίσης, να γνωρίζουν τις γενικότερες εξελίξεις όσον αφορά την κατανάλωση περιεχομένου και στην αντιμετώπιση της δημοσιογραφίας.

Σκεφτείτε τις κινητές συσκευές: Το 2012, το Ερευνητικό Κέντρο Pew βρήκε ότι περισσότερους από τους μισούς από όσους διέθεταν smartphone ή tablet κατανάλωναν ειδησεογραφικό περιεχόμενο μέσω αυτών των συσκευών (Mitchell κ.α., 2012· Pew Research Center, 2017)· μόλις τέσσερα χρόνια αργότερα, περισσότεροι από επτά στους δέκα Αμερικανούς (72%) ανέφεραν ότι κατανάλωναν ειδήσεις μέσω κινητών συσκευών (Mitchell κ.α., 2016). Καθώς η κατανάλωση ειδήσεων μέσω κινητών συσκευών συνεχίζει να αυξάνεται, είναι επιτακτική ανάγκη οι ειδησεογραφικοί οργανισμοί να παρέχουν μια θετική εμπειρία μέσω κινητών συσκευών για όλο το περιεχόμενό τους, συμπεριλαμβανομένου υλικού πλούσιου σε δεδομένα.

Κάντε το προσωπικό: Σε μια εποχή εξατομίκευσης και αλγοριθμικά παραγόμενων εμπειριών στα μέσα ενημέρωσης, αυτό μπορεί να περιλαμβάνει τη δημιουργία ευκαιριών για το κοινό να εξετάσει τα δεδομένα και να κατανοήσει τη σημασία τους για το ίδιο. Η έρευνα *Dollars for Docs* της *ProPublica* (Tigas κ.α., 2019), η οποία επιτρέπει στους ασθενείς να βλέπουν τις πληρωμές που πραγματοποιούν οι φαρμακευτικές εταιρείες και οι εταιρείες ιατρικών συσκευών στους γιατρούς τους, είναι ένα εφαρμοσμένο παράδειγμα αυτής της τεχνικής.

Προστασία των πηγών σας: Οι δημοσιογράφοι πρέπει να γνωρίζουν πώς να προστατεύουν τα δεδομένα όσο και πώς να τα αναλύουν. Η προστασία η δική σας και των πηγών σας μπορεί κάλλιστα να απαιτεί μια νέα προσέγγιση –συμπεριλαμβανομένων νέων δεξιοτήτων– για τον χειρισμό ευαίσθητων δεδομένων και ατόμων που προβαίνουν σε καταγγελίες (Keeble-Gagnere, 2015). Η κρυπτογράφηση σε συνδυασμό με την ανωνυμία (όπως φάνηκε στα *Panama Papers*) είναι ένας τρόπος για να γίνει αυτό.

Αξιοποίηση των νέων τεχνολογιών: Το Blockchain είναι μόνο ένα από τα εργαλεία που μπορούν να προστατεύσουν και να υποστηρίξουν τα δεδομένα και το ερευνητικό έργο (IJNET, 2018). Όπως εξηγεί ο Ουαλίντ Αλ-Σακάφ του Πανεπιστημίου Södertörn στη Σουηδία (Bouchart, 2018), «το Blockchain διατηρεί μόνιμα τα δεδομένα και αποτρέπει κάθε χειραγώγηση ή απάτη. Αυτό σημαίνει ότι, αν υπάρχουν κυβερνητικά δεδομένα, δεν μπορούν να καταργηθούν ή να αλλάξουν αφότου δημοσιευθούν». Η μηχανική μάθηση («machine learning») είναι μια άλλη τεχνολογία που χρησιμοποιείται ήδη σε αυτόν τον χώρο και η οποία πρόκειται να αναπτυχθεί ολοένα και περισσότερο (Bradshaw, 2017)[215].

Αποκατάσταση της εμπιστοσύνης του κοινού: Με την εμπιστοσύνη προς τη δημοσιογραφία να βρίσκεται σε ιστορικά χαμηλά επίπεδα, είναι καθήκον όλων των δημοσιογράφων να εργαστούν για την αποκατάσταση αυτού του προβλήματος (Knight Foundation, 2018· Nicolau & Giles, 2017). Για όσους εργάζονται με δεδομένα, αυτό σημαίνει ότι πρέπει να είστε διαφανείς σχετικά με τα δεδομένα με τα οποία εργάζεστε, να παρέχετε συνδέσμους προς τις αρχικές πηγές και να διασφαλίζετε ότι τα πρωτότυπα αρχεία δεδομένων είναι διαθέσιμα για λήψη. Δείχνοντας τη δουλειά σας –αυτό που ο Τζέι Ρόουζεν (2017) αποκαλεί «οι νέοι όροι για την εμπιστοσύνη στη δημοσιογραφία»–, επιτρέπετε στους αναγνώστες να δουν τις πρώτες ύλες τις οποίες επεξεργαστήκατε (και ερμηνεύσατε). Έτσι, ανοίγει μια πόρτα για την εμπιστοσύνη στις ειδήσεις, που βασίζεται στη διαφάνεια.

Η επιρροή των δεδομένων στη δουλειά σας (θα) είναι μεγαλύτερη από όσο νομίζετε. Τέλος, είναι αδύνατο να παραβλέψουμε τον ρόλο που διαδραματίζουν τα δεδομένα στη διαμόρφω-

215. ΣτΕ: Βλ., επίσης, Al-Saqaf, W., 2019 «Data journalism on the blockchain», *DataJournalism.com*, 4 Απριλίου. Διαθέσιμο σε: https://datajournalism.com/read/longreads/data-journalism-on-the-blockchain (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

ση των δημοσιογραφικών πράξεων. Πρέπει να θυμόμαστε ότι η δεδομενοποίηση της δημοσιογραφίας δεν επηρεάζει μόνο την αφήγηση ιστοριών με δεδομένα αλλά και το ευρύτερο δημοσιογραφικό επάγγελμα (Anderson, 2018· Usher, 2016).

Εργαλεία ανάλυσης όπως το Google Analytics, το Chartbeat και το Omniture είναι πανταχού παρόντα στις αίθουσες σύνταξης, παρέχοντας στους δημοσιογράφους περισσότερες πληροφορίες για τις αναγνωστικές συνήθειες του κοινού τους από ποτέ άλλοτε. Αυτές οι ποσοτικές γνώσεις, σε συνδυασμό με ποιοτικές γνώσεις (δείτε, για παράδειγμα, προγράμματα όπως το Metrics for News, που αναπτύχθηκε από το American Press Institute), επηρεάζουν το έργο μικρών και μεγάλων newsroom.

Όπως τονίζεται σε white paper που δημοσίευσε το Pars.ly[216] και σε πρόσφατες ακαδημαϊκές έρευνες (Belair-Gagnon & Holton, 2018· Cherubini & Nielsen, 2016· Ponsford, 2016· Radcliffe κ.α., 2017· Zamith, 2018), είναι σαφές ότι τα δεδομένα παίζουν καθοριστικό ρόλο τόσο στην τοποθέτηση των ιστοριών (συμπεριλαμβανομένου κυριολεκτικά του τρόπου με τον οποίο τοποθετούνται στις αρχικές σελίδες και προωθούνται στα μέσα κοινωνικής δικτύωσης) όσο και στη λήψη αποφάσεων σχετικά με τις ιστορίες που προβάλλονται.

Ο Λέβι Πούλκινεν, δημοσιογράφος και συντάκτης με έδρα το Σιάτλ και πρώην αρχισυντάκτης του *Seattle Post-Intelligencer*, υποστηρίζει ότι πολλά από αυτά τα δεδομένα υποδεικνύουν ότι οι αίθουσες σύνταξης πρέπει να κάνουν κάποια πράγματα διαφορετικά. «Νομίζω ότι υπάρχει μια διστακτικότητα στη βιομηχανία των εφημερίδων μεταξύ των δημοσιογράφων να αναγνωρίσουν ότι αυτό που μας λένε οι μετρήσεις είναι πως πρέπει να αλλάξουμε το περιεχόμενο», λέει ο Πούλκινεν (Radcliffe, 2017a), υποδεικνύοντας ότι το ρεπορτάζ για δημόσιες υποθέσεις (μεταξύ άλλων ειδών ρεπορτάζ) ίσως είναι ώριμο για μια αλλαγή. «Τους αρέσει, όταν μπορούμε να τους πούμε μια ολόκληρη ιστορία ή να τους πούμε μια σημαντική ιστορία [...], αλλά δεν μας χρειάζονται για να λειτουργούμε απλώς ως ένα είδος στενογράφου της κυβέρνησης» (Radcliffe, 2017a).

Προχωρώντας μπροστά: Πέντε ιδέες για τη συνεργασία βιομηχανίας-ακαδημαϊκής κοινότητας

Τα δεδομένα διαμορφώνουν και επηρεάζουν τις δημοσιογραφικές πράξεις σε όλες σχεδόν τις αίθουσες σύνταξης και τα είδη του ρεπορτάζ. Μπορεί να αποτελέσουν ένα εργαλείο για την αφήγηση συγκεκριμένων ιστοριών –όπως φαίνεται σε καταξιωμένους οργανισμούς όπως ο *Guardian* και σε νεότερα μέσα όπως το *FiveThirtyEight* και το *Quartz* (Ellis, 2015˙ Seward, 2016)– καθώς και μια σημαντική πηγή για τη λήψη εκδοτικών αποφάσεων, λαμβάνοντας υπόψη και τους πόρους.

Πέρα, όμως, από τις ξεχωριστές ιστορίες και στρατηγικές, τα δεδομένα προμηνύουν μια ευρύτερη αλλαγή στη δημοσιογραφία. Καλώς ή κακώς, η υιοθέτηση της ποσοτικοποίησης μπορεί να έχει σημαντικές επιπτώσεις σε βασικά χαρακτηριστικά των μεγάλων δεδομένων («big data») και της δημοσιογραφίας: στην επιστημολογία (τι γνωρίζει η δημοσιογραφία), στην εμπειρογνωμοσύνη (πώς η δημοσιογραφία εκφράζει αυτή τη γνώση), στα οικονομικά (η εμπορική αξία της δημοσιογραφίας) και στην ηθική (οι κοινωνικές αξίες της δημοσιογραφίας) –στα αγγλικά, τα «τέσσερα e»: epistemology, expertise, economics, ethics (Lewis & Westlund, 2015). Επομένως, οι επιπτώσεις που σχετίζονται με τα δεδομένα είναι μακροπρόθεσμες για τον τρόπο με τον οποίο διδάσκουμε, ασκούμε και μελετάμε τη δημοσιογραφία.

Πιστεύουμε ότι, πολύ συχνά, η ακαδημαϊκή κοινότητα και η βιομηχανία ειδήσεων αποτυγχάνουν να αναγνωρίσουν τις παραγωγικές δυνατότητες που θα μπορούσαν να προκύψουν από την ευρύτερη συνεργασία μεταξύ τους (όπως αναφέραμε και στην παρατήρησή μας σχετικά με

216. https://www.parse.ly/resources/case-studies/

τη συνεργατική δημοσιογραφία παραπάνω). Καθώς και οι δύο πλευρές καταπιάνονται με τις δυνατότητες που παρέχει η επεξεργασία δεδομένων, υποστηρίζουμε ότι μια πιο στενή σχέση μεταξύ δημοσιογράφων και ακαδημαϊκών θα μπορούσε να είναι αμοιβαία επωφελής. Παρακάτω περιγράφουμε πέντε σημεία από τα οποία θα μπορούσαμε να ξεκινήσουμε.

Περισσότερες συνεργασίες μεταξύ αιθουσών διδασκαλίας και αιθουσών σύνταξης. Η εργασία που εκπόνησε ο Πολ Μπράντσο προσφέρει μια σαφή ένδειξη του τρόπου με τον οποίο μπορεί να γίνει αυτό. Στο πλαίσιο του νέου Μεταπτυχιακού Προγράμματος στη Δημοσιογραφία Δεδομένων που προσφέρεται στο Πανεπιστήμιο Birmingham City στο Ηνωμένο Βασίλειο, ο Πολ και οι φοιτητές του συνεργάστηκαν με διάφορους ειδησεογραφικούς οργανισμούς, όπως η *Daily Telegraph* (Bradshaw, 2018), το BBC, το ITN, η *Manchester Evening News*, ο *Guardian* και το Κέντρο για την Ερευνητική Δημοσιογραφία («Centre for Investigative Journalism»)[217]. Για να επεκτείνουν αυτή τη διδακτική σύμπραξη για τη βελτίωση της έρευνας, αυτοί οι ειδησεογραφικοί οργανισμοί θα μπορούσαν να ανοίξουν τις διαδικασίες δημοσιογραφίας δεδομένων τους σε (συμμετοχική) παρατήρηση από εθνογράφους, με την προσδοκία ότι μια τέτοια πηγή γνώσης θα οδηγήσει όχι μόνο σε ακαδημαϊκές δημοσιεύσεις με αξιολόγηση ομότιμων, αλλά και σε εμπορικές δημοσιεύσεις, όπως αυτές που παράγονται από το Tow Center for Digital Journalism και το Ινστιτούτο Reuters για τη Μελέτη της Δημοσιογραφίας («Reuters Institute fot the Study of Journalism»).

Ανάληψη έργων εντός της αίθουσας διδασκαλίας με δυνητική ειδησεογραφική αξία. Ο Γιαν Γκούντεϊ, λέκτορας δημοσιογραφίας στο Πανεπιστήμιο του Κίνγκστον στο δυτικό Λονδίνο, έχει επίσης επιδείξει την ικανότητα να μετατρέπει τις εργασίες της τάξης του σε χειροπιαστά ρεπορτάζ, εντοπίζοντας πιθανές συγκρούσεις συμφερόντων στην τοπική αυτοδιοίκηση του Ηνωμένου Βασιλείου. Η έρευνά τους –η οποία περιλάμβανε την υποβολή, παρακολούθηση και ανάλυση 99 ξεχωριστών αιτημάτων για την ελευθερία της πληροφόρησης («freedom of information», FOI)– αποκάλυψε ότι οι οργανισμοί αυτοί επένδυαν συνταξιοδοτικά κεφάλαια σε εταιρείες υδρορωγμάτωσης, ενώ ταυτόχρονα ενεργούσαν και ως διαμεσολαβητές για τις προτάσεις σχεδιασμού που υπέβαλλε αυτή η εκκολαπτόμενη βιομηχανία (Goodey & Milmo, 2014). Σε ορισμένες περιπτώσεις, οι φοιτητές και οι καθηγητές τους μπορεί να έχουν στη διάθεσή τους περισσότερο χρόνο για να διερευνήσουν έργα δεδομένων, κάτι που τους επιτρέπει να εφαρμόσουν μορφές δημοσιογραφίας δεδομένων ανέφικτες για τους δημοσιογράφους, οι οποίοι κατακλύζονται από αδιάκοπες καθημερινές προθεσμίες.

Αντίστροφη ανάλυση αυτών των σχέσεων. Δεδομένων των προκλήσεων που αντιμετωπίζουν οι περισσότερες εφημερίδες, οι δημοσιογράφοι θα μπορούσαν να προσεγγίζουν συχνότερα φοιτητές και ακαδημαϊκούς με ιστορίες που θα επωφελούνταν από τη βοήθεια των δεύτερων. Ο Στιβ Ντόιγκ του Κρατικού Πανεπιστημίου της Αριζόνας, ο οποίος κέρδισε το 1993 το βραβείο Πούλιτζερ στην κατηγορία «Δημόσια προσφορά» με την εφημερίδα *Miami Herald*[218], για μια σειρά που έδειξε πώς οι απαξιωμένοι οικοδομικοί κώδικες και οι κακές κατασκευαστικές πρακτικές επιδείνωσαν τις ζημιές που προκάλεσε ο τυφώνας Άντριου έναν χρόνο νωρίτερα, συμβουλεύει ενεργά για τα προβλήματα του υποβοηθούμενου από υπολογιστή ρεπορτάζ («computer-assisted reporting»)[219]. Ο ίδιος κέρδισε το βραβείο Τζορτζ Πολκ (2012) για το *Decoding Prime*, μια ανάλυση ύποπτων πρακτικών νοσοκομειακής τιμολόγησης για τον οργανισμό ερευνών California Watch[220]. Πρόκειται για ένα μοντέλο παροχής συμβουλών και καθοδήγησης –με τη συμμετοχή καθηγητών και πιθανή ανάμειξη φοιτητών–, το οποίο θα μπορούσαν να μιμηθούν και άλλοι.

217. http://bcu.ac.uk/media/courses/data-journalism-ma-2018-19
218. http://www.pulitzer.org/winners/miami-herald
219. https://www.flickr.com/photos/juggernautco/sets/72157607210036175/
220. https://cronkite.asu.edu/about/faculty-and-leadership/faculty/doigbio

Ανοίξτε την πόρτα στους ερευνητές και στην ανεξάρτητη κριτική. Οι δημοσιογράφοι είναι γνωστό ότι βασίζονται τακτικά σε ακαδημαϊκούς ως πηγές για τις ειδήσεις τους, αλλά συχνά αισθάνονται λιγότερο άνετα να εκτεθούν στον ακαδημαϊκό έλεγχο. Το πρόβλημα αυτό επιδεινώνεται από τις ολοένα και πιο αυστηρές οργανωτικές οδηγίες κατά της διεξαγωγής ερευνών ή της συνομιλίας με ερευνητές χωρίς την άδεια της ανώτερης διοίκησης. Όμως, όπως ακριβώς οι δημοσιογράφοι χρειάζονται καλό πρωτογενές υλικό για να κάνουν τη δουλειά τους, έτσι και οι ακαδημαϊκοί, για να κάνουν σωστή έρευνα για τη δημοσιογραφία, πρέπει να εξασφαλίσουν καλύτερη πρόσβαση από αυτήν που έχουν σήμερα. Αυτό είναι ιδιαίτερα σημαντικό, καθώς οι ερευνητές προσπαθούν να κατανοήσουν τι σημαίνει η ψηφιοποίηση των δεδομένων για τη δημοσιογραφία (Baack, 2015) –πώς οι δημοσιογράφοι χρησιμοποιούν τις μετρήσεις (Belair-Gagnon & Holton, 2018· Christin, 2018· Ferrer-Conill & Tandoc, 2018), πώς αφηγούνται ιστορίες με νέους τρόπους (Hermida & Young, 2019· Toff, 2019· Usher, 2016) και ούτω καθεξής. Κάπως μικρότερη αμυντικότητα εκ μέρους των ειδησεογραφικών οργανισμών θα μπορούσε να συμβάλλει πολύ στην ανάπτυξη μιας αμοιβαία επωφελούς σχέσης: οι ερευνητές θα αποκτούσαν καλύτερη πρόσβαση στην κατανόηση του τρόπου με τον οποίο τα δεδομένα ταιριάζουν στη δημοσιογραφία και, με τη σειρά τους, οι ειδησεογραφικοί οργανισμοί θα μπορούσαν να κερδίσουν μια ανεξάρτητη αξιολόγηση της δουλειάς τους, αποτιμώντας καλύτερα, από μια κριτική απόσταση, τις επιδόσεις τους.

Βεβαιωθείτε ότι η έρευνά σας είναι προσβάσιμη. Από την άλλη πλευρά, οι ακαδημαϊκοί θα μπορούσαν να κάνουν πολύ περισσότερα, για να διασφαλίσουν τη διαφάνεια και την προσβασιμότητα στο έργο τους. Μέχρι τώρα, έχουν εκπονηθεί δεκάδες ακαδημαϊκές μελέτες σχετικά με τον «εκσυγχρονισμό της δεδομενοποίησης της δημοσιογραφίας», με ιδιαίτερη έμφαση στην εξέλιξη των εργαλείων για την αφήγηση ιστοριών με δεδομένα, και στον αντίκτυπό της στη δημοσιογραφική δεοντολογία και στις δημοσιογραφικές προσεγγίσεις (Ausserhofer κ.α., 2020). Οι μελέτες αυτές θα μπορούσαν να έχουν τεράστια σημασία για τους ειδησεογραφικούς οργανισμούς. Ωστόσο, πολύ συχνά βρίσκονται κλειδωμένες πίσω από τα paywall των ακαδημαϊκών περιοδικών, επισκιάζονται από την υπερβολική χρήση της επιστημονικής αργκό και εν γένει διατίθενται με τέτοιο τρόπο που δυσχεραίνει την πρόσβαση στους δημοσιογράφους, πόσο μάλλον την κατανόηση των μεταβιβάσιμων διδαγμάτων αυτών των ερευνών. Εφόσον είναι δυνατόν, οι δραστηριότητες προβολής και η αλληλεπίδραση με τον κλάδο θα μπορούσε να αποτελέσει αναπόσπαστο μέρος της διαδικασίας δημοσίευσης, ώστε τα οφέλη των πληροφοριών αυτών να έχουν απήχηση και πέραν των περιοδικών –όπως μέσω επαναδιατυπωμένων σύντομων ενημερωτικών δελτίων ή σύντομων επεξηγήσεων για εξειδικευμένα περιοδικά, όπως το Nieman Journalism Lab, ή ιστοτόπους που έχουν σχεδιαστεί για τη διάδοση ακαδημαϊκών εργασιών σε μη εξειδικευμένο κοινό, όπως ο ιστότοπος *The Conversation*.

Συμπερασματικά

Η δημοσιογραφία δεδομένων, σύμφωνα με τον διάσημο δημοσιογράφο δεδομένων Σάιμον Ρότζερς, ο οποίος σήμερα είναι συντάκτης δεδομένων στην Google, είναι «ένας μεγάλος ισοπεδωτής» (Rogers, 2012). Λόγω του αναδυόμενου χαρακτήρα της, σχεδόν οποιοσδήποτε μπορεί να τη δοκιμάσει και να αποκτήσει επάρκεια σε αυτήν. «Η δημοσιογραφία δεδομένων είναι το νέο πανκ», μας λέει (Rogers, 2012). Αυτό σημαίνει ότι «πολλές δημοσιογραφικές ομάδες ξεκινούν με τόση προϋπάρχουσα γνώση και εμπειρία, όση έχει κάποιος που χακάρει από την κρεβατοκάμαρά του» (Rogers, 2012)[221].

Η δημοσιογραφία δεδομένων, φυσικά, έχει μακρά ιστορία, με προϊστορία σε μορφές επιστή-

221. Βλ. επίσης το κεφάλαιο του Σάιμον Ρότζερς σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 7.3, «Από τον Guardian στο Google News Lab: Μια Δεκαετία Εργασίας στη Δημοσιογραφία Δεδομένων»).

μης και αφήγησης που κυκλοφορούν εδώ και πάνω από έναν αιώνα (Anderson, 2018)[222]. Ωστόσο, ως ένας εκκολαπτόμενος «κοινωνικός κόσμος» (Lewis & Zamith, 2017) στο πλαίσιο της δημοσιογραφίας –ένας χώρος για την ανταλλαγή εργαλείων, τεχνικών και βέλτιστων πρακτικών μεταξύ ειδησεογραφικών οργανισμών και σε ολόκληρο τον κόσμο–, η δημοσιογραφία δεδομένων βρίσκεται σε ένα ιδιαίτερο σημείο καμπής, εν μέσω της ευρύτερης δεδομενοποίησης της κοινωνίας στον 21ο αιώνα.

Υποστηρίζουμε ότι υπάρχει μια αντίστοιχη ευκαιρία για κριτική ενδοσκόπηση: για να εξετάσουμε τι γνωρίζουμε μέχρι στιγμής σχετικά με τη δημοσιογραφία δεδομένων, για να περιγράψουμε τι μένει να διερευνηθεί και, κυρίως, για να ακολουθήσουμε μια πορεία που αναγνωρίζει την αμοιβαία εξάρτηση της δημοσιογραφίας ως πρακτικής και ως παιδαγωγικής, της βιομηχανίας και του ακαδημαϊκού χώρου. Για να μπορέσει η δημοσιογραφία να δώσει νόημα σε έναν κόσμο που κατακλύζεται από δεδομένα, απαιτείται καλύτερη αναγνώριση, με αυτοκριτική, των περιορισμών και των γενεσιουργών δυνατοτήτων των προσεγγίσεων που βασίζονται στα δεδομένα –τι αποκαλύπτουν, τι δεν αποκαλύπτουν και πώς μπορούν να βελτιωθούν.

Παραπομπές

Anderson, C. W. (2018). *Apostles of certainty: Data journalism and the politics of doubt.* Οξφόρδη: Oxford University Press.

Ausserhofer, J., Gutounig, R., Oppermann, M., Matiasek, S., & Goldgruber, E. (2020) «The δεδομενοποίηση of data journalism scholarship: Focal points, methods, and research propositions for the investigation of data-intensive newswork», *Journalism*, 21,(7) σσ. 950–973. DOI: https://doi.org/10.1177/1464884917700667

Baack, S. (2015) «Datafication and empowerment: How the open data movement rearticulates notions of democracy, participation, and journalism», *Big Data & Society,* 2(2). DOI: https://doi.org/10.1177/2053951715594634

Barrett, D. (2009) «MPs' expenses: How they milk the system», *The Telegraph*, 17 Μαΐου. Διαθέσιμο σε: https://www.telegraph.co.uk/news/newstopics/mps-expenses/5335101/MPs-expenses-how-they-milk-the-system.html (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

BBC News (2009). «Q&A: MP expenses row explained», 18 Ιουνίου. Διαθέσιμο σε: http://news.bbc.co.uk/2/hi/uk_news/politics/7840678.stm (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Beckett, C. (2010). *The value of networked journalism*. Λονδίνο: POLIS, London School of Economics and Political Science.

Belair-Gagnon, V., & Holton, A. E. (2018) «Boundary work, interloper media, and analytics in newsrooms», *Digital Journalism*, 6(4), σσ. 492–508. DOI: https://doi.org/10.1080/21670811.2018.1445001

Boczkowski, P. J., & Lewis, S. C. (2018) «The center of the universe no more: From the self-centered stance of the past to the relational mindset of the future», στο P. J. Boczkowski & Z. Papacharissi (επιμ.), *Trump and the media*, σσ. 177–185. Κέιμπριτζ, Μασαχουσέτη: MIT Press. DOI: https://doi.org/10.7551/mitpress/11464.003.0028

Bouchart, M. (2018) *A data journalist's new year's resolutions* [Medium], 1 Φεβρουαρίου. Διαθέσιμο σε: https://medium.com/data-journalism-awards/a-data-journalists-new-year-s-resolutions-474ef92f7e8f (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

222. Βλ. επίσης το κεφάλαιο του Άντερσον σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 9.1, «Γενεαλογίες της Δημοσιογραφίας Δεδομένων»).

Bradshaw, P. (2017) «Data journalism's AI opportunity: The 3 different types of machine learning & how they have already been used», *Online Journalism Blog,* 14 Δεκεμβρίου. Διαθέσιμο σε: https://onlinejournalismblog.com/2017/12/14/data-journalisms-ai-opportunity-the-3-different-types-of-machine-learning-how-they-have-already-been-used/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Bradshaw, P. (2018) *Wanted: MA Data Journalism applicants to partner with The Telegraph* [Medium], 15 Φεβρουαρίου. Διαθέσιμο σε: https://medium.com/@paulbradshaw/wanted-ma-data-journalism-applicants-to-partner-with-the-telegraph-8abd154260f3 (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world.* Κέιμπριτζ, Μασαχουσέτη: MIT Press.

Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C.-H. (2019) «Artificial intelligence and journalism», *Journalism & Mass Communication Quarterly,* 96(3), σσ. 673–695. DOI: https://doi.org/10.1177/1077699019859901

Carlson, M., & Lewis, S. C. (2019) «Temporal reflexivity in journalism studies: Making sense of change in a more timely fashion», *Journalism*, 20(5), σσ. 642–650. DOI: https://doi.org/10.1177/1464884918760675

Cherubini, F., & Nielsen, R. K. (2016) «Editorial analytics: How news media are developing and using audience data and metrics», *Social Science Research Network*. DOI: https://doi.org/10.2139/ssrn.2739328

Christin, A. (2018) «Counting clicks: Quantification and variation in web journalism in the United States and France», *American Journal of Sociology,* 123(5), σσ. 1382–1415. DOI: https://doi.org/10.1086/696137

CBC News (2016) «*A chronology of the Senate expenses scandal*», 13 Ιουλίου. Διαθέσιμο σε: https://www.cbc.ca/news/politics/senate-expense-scandal-timeline-1.3677457 (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Couldry, N., & Mejias, U. A. (2019) «Data colonialism: Rethinking big data's relation to the contemporary subject», *Television & New Media,* 20(4), σσ. 336–349. DOI: https://doi.org/10.1177/1527476418796632

Crawford, K., Gray, M. L., & Miltner, K. (2014) «Critiquing big data: Politics, ethics, epistemology» *International Journal of Communication*, 8, σσ. 1663–1672. Διαθέσιμο σε: https://ijoc.org/index.php/ijoc/article/view/2167 (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media.* Κέιμπριτζ, Μασαχουσέτη: Harvard University Press.

Ellis, J. (2015) «Quartz maps a future for its interactive charts with Atlas», *Nieman Lab,* 23 Ιουνίου. Διαθέσιμο σε: https://www.niemanlab.org/2015/06/quartz-maps-a-future-for-its-interactive-charts-with-atlas/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Ferrer-Conill, R., & Tandoc, E. C. (2018) «The audience-oriented editor», *Digital Journalism*, 6(4), σσ. 436–453. DOI: https://doi.org/10.1080/21670811.2018.1440972

Frank, A., Yang, Y., & Radcliffe, D. (2015) «The mainstreaming of data reporting and what it means for journalism schools», *Journalism.co.uk*, 18 Δεκεμβρίου. Διαθέσιμο σε: https://www.journalism.co.uk/news/the-mainstreaming-of-data-journalism-and-what-it-means-for-journalism-schools/s2/a593608/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Glaeser, E. L., Kominers, S. D., Luca, M., & Naik, N. (2018) «Big data and big cities: The promises and limitations of improved measures of urban life», *Economic Enquiry*, 56(1), σσ. 114–137.

Διαθέσιμο σε: https://www.hbs.edu/faculty/Pages/item.aspx?num=51012 (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Goodey, J., & Milmo, C. (2014) «Exclusive: Local authorities have "conflict of interest" on fracking investments», *The Independent*, 27 Απριλίου. Διαθέσιμο σε: http://www.independent.co.uk/news/uk/politics/exclusive-local-authorities-have-conflict-of-interest-on-fracking-investments-9294590.html (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Gray, J., Chambers, L., & Bounegru, L. (επιμ.) (2012). *The data journalism handbook: How journalists can use data to improve the news.* Σεβαστούπολη, Καλιφόρνια: O'Reilly Media.

Guardian News & Media (2009) «MPs' expenses: The Guardian launches major crowdsourcing experiment», *The Guardian*, 23 Ιουνίου. Διαθέσιμο σε: https://www.theguardian.com/gnm-press-office/crowdsourcing-mps-expenses (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Hamilton, J. T. (2016). *Democracy's detectives: The economics of investigative journalism.* Κέιμπριτζ, Μασαχουσέτη: Harvard University Press.

Hermida, A., & Young, M. L. (2019). *Data journalism and the regeneration of news.* Λονδίνο: Routledge.

Höchtl, J., Parycek, P., & Schöllhammer, R. (2016) «Big data in the policy cycle: Policy decision making in the digital era», *Journal of Organizational Computing and Electronic Commerce*, 26(1–2), σσ. 147–169. DOI: https://doi.org/10.1080/10919392.2015.1125187

ICIJ (άνευ ημερομηνίας) «*The Panama Papers: Exposing the rogue offshore finance industry*». https://www.icij.org/investigations/panama-papers/

IJNET (2018) «How blockchain technology can boost freedom of the press», *International Journalists' Network,* 5 Αυγούστου. Διαθέσιμο σε: https://ijnet.org/en/story/how-blockchain-technology-can-boost-freedom-press (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Johnson, S. (2013) «Sorry EveryBlock, you never learned how to write headline», *The Hudson Eclectic,* 8 Φεβρουαρίου. Διαθέσιμο σε: https://hudsoneclectic.com/2013/02/08/sorry-everyblock-you-never-learned-how-to-write-a-headline/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Keeble-Gagnere, G. (2015) «Encryption for the working journalist: Communicating securely», *Journalism.co.uk*, 5 Νοεμβρίου. Διαθέσιμο σε: https://www.journalism.co.uk/news/encryption-for-the-working-journalist/s2/a580914/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Klein, S. (2015) «Antebellum data journalism: Or, how big data busted Abe Lincoln», *ProPublica,* 17 Μαρτίου. Διαθέσιμο σε: https://www.propublica.org/nerds/antebellum-data-journalism-busted-abe-lincoln (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Knight Foundation (2018) *10 reasons why Americans don't trust the media* [Medium], 26 Ιουνίου. Διαθέσιμο σε: https://medium.com/trust-media-and-democracy/10-reasons-why-americans-dont-trust-the-media-d0630c125b9e (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Kreiss, D. (2016). *Prototype politics: Technology-intensive campaigning and the data of democracy.* Οξφόρδη: Oxford University Press.

Larcinese, V., & Sircar, I. (2017) «Crime and punishment the British way: Accountability channels following the MPs' expenses scandal», *European Journal of Political Economy,* 47, σσ. 75–99. DOI: https://doi.org/10.1016/j.ejpoleco.2016.12.006

Lewis, S. C. (2017) «Digital journalism and big data», στο B. Franklin & S. Eldridge (επιμ.), *The Routledge companion to digital journalism studies,* σσ. 126–135. Λονδίνο: Routledge.

Lewis, S. C. (2019) «Lack of trust in the news media, institutional weakness, and relational journalism as a potential way forward», *Journalism*, 20(1), σσ. 44–47. DOI: https://doi.org/10.1177/1464884918808134

Lewis, S. C., & Usher, N. (2014) «Code, collaboration, and the future of journalism: A case study of the hacks/hackers global network», *Digital Journalism,* 2(3), σσ. 383–393. DOI: https://doi.org/10.1080/21670811.2014.895504

Lewis, S. C., & Usher, N. (2016) «Trading zones, boundary objects, and the pursuit of news innovation: A case study of journalists and programmers», *Convergence*, 22(5), σσ. 543–560. DOI: https://doi.org/10.1177/1354856515623865

Lewis, S. C., & Westlund, O. (2015) «Big data and journalism: Epistemology, expertise, economics, and ethics», *Digital Journalism*, 3(3), σσ. 447–466. DOI: https://doi.org/10.1080/21670811.2014.976418

Lewis, S. C., & Zamith, R. (2017) «On the worlds of journalism», στο P. J. Boczkowski & C. W. Anderson (επιμ.), *Remaking the news: Essays on the future of journalism scholarship in the digital age,* σσ. 111–128. Κέιμπριτζ, Μασαχουσέτη: MIT Press.

Los Angeles Times (2026) «*2016 year in review: Highlights and heartbreaks*», 29 Δεκεμβρίου. Διαθέσιμο σε: https://www.latimes.com/local/california/la-me-updates-best-year-review-2016-htmlstory.html (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Lowenstein, J. K. (2015) «How I used data-driven journalism to reveal racial disparities in U.S. nursing homes», *Storybench*, 13 Φεβρουαρίου. Διαθέσιμο σε: https://www.storybench.org/how-i-used-data-driven-journalism-to-reveal-racial-disparities-in-u-s-nursing-homes/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Mitchell, A., Gottfried, J., Barthel, M., & Shearer, E. (2016) «The modern news consumer: News attitudes and practices in the digital era», *Pew Research Center,* 7 Ιουλίου. Διαθέσιμο σε: https://www.journalism.org/2016/07/07/pathways-to-news/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Mitchell, A., Rosenstiel, T., & Christian, L. (2012) «Mobile devices and news consumption: Some good signs for journalism», στο Pew Research Center (επιμ.), *The state of the news media: An annual report on American journalism*. Διαθέσιμο σε: https://www.pewresearch.org/wp-content/uploads/sites/8/2017/05/State-of-the-News-Media-Report-2012-FINAL.pdf (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Nicolau, A., & Giles, C. (2017) «Public trust in media at all time low, research shows», *Financial Times,* 16 Ιανουαρίου. Διαθέσιμο σε: https://www.ft.com/content/fa332f58-d9bf-11e6-944b-e7eb37a6aa8e (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy.* Νέα Υόρκη: Crown.

Pew Research Center (2017) «*News consumption on digital devices*», 21 Αυγούστου. Διαθέσιμο σε: https://www.pewresearch.org/journalism/2012/03/18/mobile-devices-and-news-consumption-some-good-signs-for-journalism/5-news-consumption-on-digital-devices/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Ponsford, D. (2016) «"Heartbroken" reporter Gareth Davies says Croydon Advertiser print edition now "thrown together collection of clickbait», *Press Gazette,* 31 Ιουλίου. Διαθέσιμο σε: https://www.pressgazette.co.uk/heartbroken-reporter-gareth-davies-says-croydon-advertser-print-edition-now-thrown-together-collection-of-clickbait/ (Προσπελάστηκε! 18 Δεκεμβρίου 2021).

Radcliffe, D. (2013) «Hyperlocal media and data journalism», στο J. Mair & R. L. Keeble (επιμ.), *Data journalism: Mapping the future,* σσ. 120–132. Μπέρι Σαίν Έντμουντς, Ηνωμένο Βασίλειο: Abramis Academic Publishing.

Radcliffe, D. (2017a) «Local journalism in the Pacific Northwest: Why it matters, how it's evolving, and who pays for it», *Social Science Research Network*. Διαθέσιμο σε: https://papers.ssrn.com/abstract=3045516 (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Radcliffe, D. (2017b) «Data journalism in the US: Three case studies and ten general principles for journalists», στο J. Mair, R. L. Keeble, & M. Lucero (επιμ.), *Data journalism: Past, present and future,* σσ. 197–210. Μπέρι Σαίν Έντμουντς, Ηνωμένο Βασίλειο: Abramis Academic Publishing.

Radcliffe, D. (2017c) «10 key principles for data-driven storytelling», *Journalism.co.uk*, 28 Νοεμβρίου. Διαθέσιμο σε: https://www.journalism.co.uk/news/10-key-principles-for-data-driven-storytelling/s2/a713879/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Radcliffe, D., Ali, C., & Donald, R. (2017) «Life at small-market newspapers: Results from a survey of small-market newsrooms», *Tow Center for Digital Journalism*. DOI: https://doi.org/10.7916/D8XP7BGC

Rayner, G. (2009) «MPs' expenses: Ten ways MPs play the system to cash in on expenses and allowances», *The Telegraph*, 8 Μαΐου. Διαθέσιμο σε: https://www.telegraph.co.uk/news/newstopics/mps-expenses/5293498/MPs-expenses-Ten-ways-MPs-play-the-system-to-cash-in-on-expenses-and-allowances.html (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Reese, S. D. (1999) «The progressive potential of journalism education: Recasting the academic versus professional debate», *Harvard International Journal of Press/ Politics,* 4(4), σσ. 70–94. DOI: https://doi.org/10.1177/1081180X9900400405

Reuters (2017) «Expenses scandal an embarrassing start to 2017 for Australia's embattled Prime Minister», *The Indian Express,* 13 Ιανουαρίου. Διαθέσιμο σε: https://indianexpress.com/article/world/expenses-scandal-an-embarrassing-start-to-2017-for-australias-embattled-prime-minister/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Rogers, S. (2009a) «How to crowdsource MPs' expenses», *The Guardian,* 18 Ιουνίου. Διαθέσιμο σε: https://www.theguardian.com/news/datablog/2009/jun/18/mps-expenses-houseofcommons (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Rogers, S. (2009b) «MPs' expenses: What you've told us. So far», *The Guardian*, 19 Ιουνίου. Διαθέσιμο σε: http://www.theguardian.com/news/datablog/2009/sep/18/mps-expenses-westminster-data-house-of-commons (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Rogers, S. (2012) «Anyone can do it. Data journalism is the new punk», *The Guardian*, 24 Μαΐου. Διαθέσιμο σε: https://www.theguardian.com/news/datablog/2012/may/24/data-journalism-punk (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Rosen, J. (2017) «Show your work: The new terms for trust in journalism», *PressThink*, 31 Δεκεμβρίου. Διαθέσιμο σε: https://pressthink.org/2017/12/show-work-new-terms-trust-journalism/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Sambrook, R. (επιμ.). (2018). *Global teamwork: The rise of collaboration in investigative journalism*. Οξφόρδη: Reuters Institute for the Study of Journalism.

Schroeder, R. (2018). Social theory after the Internet: *Media, technology, and globalization.* Λονδίνο: UCL Press.

HuffPost Canada (άνευ ημερομηνίας) «*Senate expenses scandal*» Διαθέσιμο σε: https://www.huffingtonpost.ca/news/senate-expenses-scandal/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Seward, Z. M. (2016) «Atlas is now an open platform for everyone's charts and data», *Quartz*, 10 Μαΐου. Διαθέσιμο σε: https://qz.com/679853/atlas-is-now-an-open-platform-for-everyones-charts-and-data/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Telegraph Reporters (2009) «MPs' expenses: Full list of MPs investigated by The Telegraph», *The Telegraph*, 8 Μαΐου. Διαθέσιμο σε: https://www.telegraph.co.uk/politics/0/mps-expenses-full-list-mps-investigated-telegraph/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

The Economist (2013) «*Worth a thousand words*», 7 Οκτωβρίου. Διαθέσιμο σε: https://www.economist.com/christmas-specials/2013/10/07/worth-a-thousand-words (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

The Washington Post (άνευ ημερομηνίας) «*The depth of the problem*».

Thurman, N., Lewis, S. C., & Kunert, J. (2019) «Algorithms, automation, and news», *Digital Journalism,* 7(8), σσ. 980–992. DOI: https://doi.org/10.1080/21670811.2019.1685395

Tigas, M., Jones, R. G., Ornstein, C., & Groeger, L. (2019) «Dollars for docs», *ProPublica*, 17 Οκτωβρίου. Διαθέσιμο σε: https://projects.propublica.org/docdollars/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Toff, B. (2019) «The "Nate Silver effect" on political journalism: Gatecrashers, gatekeepers, and changing newsroom practices around coverage of public opinion polls», *Journalism*, 20(7), σσ. 873–889. DOI: https://doi.org/10.1177/1464884917731655

Usher, N. (2016). Interactive journalism: Hackers, data, and code. Ουρμπάνα, Ιλινόις: *University of Illinois Press.*

Van der Haak, B., Parks, M., & Castells, M. (2012) «The future of journalism: *Networked journalism», International journal of communication,* 6, σσ. 2923–2938.

Zamith, R. (2018) «Quantified audiences in news production: A synthesis and research agenda», *Digital Journalism,* 6(4), σσ. 418–435. DOI: https://doi.org/10.1080/21670811.2018.1444999

Οι συγγραφείς

Ο Ντέμιεν Ράντκλιφ κατέχει την έδρα Carolyn S. Chambers στη Δημοσιογραφία, είναι Καθηγητής Πρακτικής και συνεργάτης του Τμήματος Μελετών Μέσης Ανατολής και Βόρειας Αφρικής (Middle East and North Africa Studies, MENA) στο Πανεπιστήμιο του Όρεγκον.
Ο Σεθ Λούις (PhD, Πανεπιστήμιο του Τέξας στο Όστιν) είναι ιδρυτής και κάτοχος της έδρας Shirley Papé για τα αναδυόμενα μέσα ενημέρωσης στη Σχολή Δημοσιογραφίας και Επικοινωνίας του Πανεπιστημίου του Όρεγκον και επισκέπτης ερευνητής στο Ινστιτούτο Reuters για τη Μελέτη της Δημοσιογραφίας και στο Green Templeton College στο Πανεπιστήμιο της Οξφόρδης.

8.3
Δημοσιογραφία Δεδομένων Από Περιθωριοποιημένες Κοινότητες, Σχετικά με Αυτές και για Λογαριασμό τους

Eva Constantaras

Περίληψη: Η δημοσιογραφία δεδομένων μπορεί να διαδραματίσει σημαντικό ρόλο στη χειραφέτηση περιθωριοποιημένων κοινοτήτων για την καταπολέμηση της αδικίας, της ανισότητας και των διακρίσεων.

Λέξεις-κλειδιά: δημοσιογραφία δεδομένων, περιθωριοποιημένες κοινότητες, αδικία, ανισότητα, διακρίσεις, χειραφέτηση

Ασκώ τη δημοσιογραφία δεδομένων σε χώρες όπου θεωρείται ότι τα πράγματα πηγαίνουν γενικώς άσχημα –μιλάμε δηλαδή για χώρες που δεν περνούν απλώς μια δύσκολη φάση, δεν περνούν απλώς μια πολιτική αναποδιά, αλλά ολόκληρα πολιτικά και οικονομικά συστήματα καταρρέουν. Σε τέτοια μέρη διαβάζει κανείς ότι η διαφθορά έχει παραλύσει την κυβέρνηση, οι πολίτες είναι απελπισμένοι και η κοινωνία των πολιτών τελεί υπό καθεστώς πολιορκίας. Η κατάσταση είναι άθλια. Η παραγωγή δημοσιογραφίας δεδομένων σε μερικά από τα πιο φτωχά, χωρίς παιδεία και επικίνδυνα μέρη του κόσμου με οδήγησε σε ένα σημαντικό συμπέρασμα. Η αδικία, η ανισότητα και οι διακρίσεις είναι πανταχού παρούσες, ύπουλες και παραγνωρισμένες στις περισσότερες χώρες.

Κάποιοι δημοσιογράφοι με τους οποίους συνεργάζομαι έχουν υιοθετήσει ανεπιφύλακτα τα νέα εργαλεία για να μετρήσουν, για πρώτη φορά, το πόσο άσχημα είναι τα πράγματα, ποιος υποφέρει από την κατάσταση, ποιος φταίει και πώς μπορεί αυτή να βελτιωθεί. Σε αυτό το πλαίσιο, οι δημοσιογράφοι έχουν υιοθετήσει τα δεδομένα ως μέσο για να επηρεάσουν την πολιτική, να κινητοποιήσουν τους πολίτες και να καταπολεμήσουν την προπαγάνδα. Παρά τους περιορισμούς της ελευθεροτυπίας, η δημοσιογραφία δεδομένων θεωρείται μέσο χειραφέτησης.

Αυτό το κεφάλαιο διερευνά τη δημοσιογραφία δεδομένων από περιθωριοποιημένες κοινότητες, σχετικά με αυτές και για λογαριασμό τους. Εξετάζοντας διάφορες πτυχές της αδικίας, της ανισότητας και των διακρίσεων, καθώς και τις ευρύτερες συνέπειές τους στη ζωή των περιθωριοποιημένων κοινοτήτων, τις καθιστούμε ορατές, μετρήσιμες και ίσως ακόμη και επιλύσιμες. Αυτές οι ιστορίες προσελκύουν δημοσιογράφους βαθιά συνδεδεμένους με περιθωριοποιημένες κοινότητες. Προσεγγίζουν θέματα για τα οποία ενδιαφέρονται ομάδες που αντιμετωπίζουν θεσμικές διακρίσεις, ώστε να ενισχύσουν τη συμμετοχή των πολιτών. Διαδίδονται μέσω των τοπικών μέσων μαζικής ενημέρωσης, για να φτάσουν σε μεγάλο αριθμό ανθρώπων και να πιέσουν τις κυβερνήσεις να λάβουν καλύτερες αποφάσεις για ολόκληρη τη χώρα. Στη συνέχεια, θα αναφερθώ σε πέντε είδη ιστοριών δημοσιογραφίας δεδομένων που ασχολούνται με τα συμφέροντα και τις ανησυχίες περιθωριοποιημένων κοινοτήτων στο Αφγανιστάν, στο Πακιστάν, στην Κένυα, στην Κιργιζία και στα Βαλκάνια.

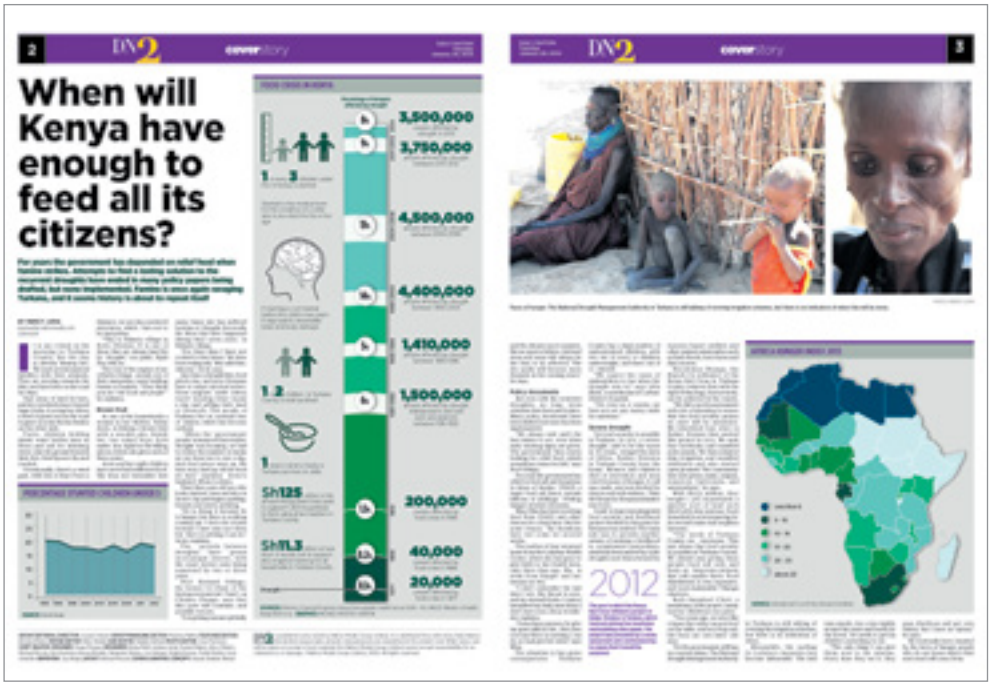
Εικόνα 8.3.1: Δημοσιογράφοι δεδομένων στο Πακιστάν αναπτύσσουν τους σκελετούς των θεμάτων τους («wireframes») με τα ευρήματα των δεδομένων του. Πηγή: *Internews*.

Γιατί οι άνθρωποι πεινάνε, αν η χώρα μας διαθέτει αρκετούς πόρους, για να τους θρέψει όλους;

Στην Κένυα, οι χορηγοί βοήθειας χρηματοδοτούσαν εντελώς λάθος επισιτιστικά προγράμματα. Ένα δωδεκάλεπτο τηλεοπτικό ρεπορτάζ της Μέρσι Γιούμα του NTV για την Τουρκάνα, μια απομονωμένη, φτωχή περιοχή της Βόρειας Κένυας, αποκάλυψε ότι ο υποσιτισμός των παιδιών αποτελεί αυξανόμενο πρόβλημα, καθώς οι ξηρασίες και οι λιμοί εμφανίζονται όλο και πιο συχνά και έντονα. Τα χρήματα πηγαίνουν σε επείγουσα επισιτιστική βοήθεια, όχι σε μακροπρόθεσμη αντιμετώπιση της ξηρασίας. Τα ίδια χρήματα που δαπανώνται για ένα έτος έκτακτης επισιτιστικής βοήθειας θα μπορούσαν να χρηματοδοτήσουν ένα πρόγραμμα διατροφικής βιωσιμότητας για ολόκληρη την κομητεία και τους σχεδόν ένα εκατομμύριο κατοίκους της, σύμφωνα με σχεδιαζόμενες πολιτικές του κοινοβουλίου. Η Γιούμα απείλησε να αποσύρει την ιστορία της, όταν οι εκδότες θέλησαν να αφαιρέσουν τα δεδομένα: η ιστορία της εξαρτιόταν από την κινητοποίηση των χορηγών βοήθειας, την οργή των πολιτών και την ταπείνωση της κυβέρνησης, κυρίως μέσω της τηλεόρασης αλλά και μέσω έντυπης έκδοσης και μιας σύνοψης που δημοσιεύτηκε στο Διαδίκτυο (Juma, 2014).

Με τη δύναμη των δεδομένων της έπεισε τους χορηγούς βοήθειας. Συγκέντρωσε δεδομένα για το κλίμα, τη γεωργία και την υγεία από κυβερνητικά υπουργεία, έρευνες δημόσιας υγείας, οργανισμούς χορηγών βοήθειας και τον Ερυθρό Σταυρό της Κένυας. Η αποστολή της οργάνωσης USAID στην Κένυα είδε την οπτικοποίηση των δεδομένων, που καταδείκνυε ότι ένα έτος επείγουσας επισιτιστικής βοήθειας της USAID θα μπορούσε να χρηματοδοτήσει ολόκληρη τη στρατηγική επισιτιστικής βιωσιμότητας του Ερυθρού Σταυρού της Κένυας για την Τουρκάνα. Η δημοσιογράφος επεσήμανε τον αντίκτυπο των καθυστερήσεων στην υγεία των παιδιών και την έντονη αντίθεση με τις χώρες που καλλιεργούν τρόφιμα σε ερήμους. Κλήθηκε να παρουσιάσει τα ευρήματά της στο γραφείο της USAID στο Ναϊρόμπι και, το 2015, η στρατηγική της USAID για τη γεωργία και την επισιτιστική ασφάλεια μετατοπίστηκε από την ανθρωπιστική βοήθεια στη βιώσιμη γεωργία[223].

223. https://www.usaid.gov/kenya/agriculture-and-food-security



Εικόνα 8.3.2: Έντυπη έκδοση του τηλεοπτικού αφιερώματος της Μέρσι Γιούμα για την επισιτιστική ασφάλεια στην Τουρκάνα. Πηγή: *Internews*.

Η ίδια εντυπωσίασε την κοινή γνώμη με την παραστατική παρουσίαση των οικογενειών που λιμοκτονούσαν στην Τουρκάνα. Πέρασε τρεις ημέρες με τις οικογένειες που παρουσιάζονται στο έργο μαζί με έναν μεταφραστή από την Τουρκάνα και έναν εικονολήπτη. Τα τηλέφωνα του σταθμού άρχισαν να χτυπούν ασταμάτητα πριν να τελειώσει η προβολή του ρεπορτάζ, από Κενυάτες οι οποίοι ζητούσαν να προσφέρουν χρήματα στις οικογένειες που παρουσιάζονταν στο ρεπορτάζ. Χάρη στη μαζική αντίδραση που προκάλεσε η είδηση από ιδιώτες και οργανισμούς, ο σταθμός δημιούργησε μέσα σε λίγες ώρες ένα ταμείο αρωγής για την κομητεία Τουρκάνα. Αυτό και τα επόμενα ρεπορτάζ για την απελπιστική κατάσταση στην οποία είχε φέρει ο λιμός τη Βόρεια Κένυα τράβηξαν σε καθημερινή βάση την προσοχή των κενυατικών μέσων ενημέρωσης, τα οποία ιστορικά έχουν επιδείξει αδιαφορία για τα δεινά των απομονωμένων και εξαθλιωμένων περιοχών της Βόρειας Κένυας. Το βασικό κοινό της δημοσιογράφου συνδέθηκε με μια δυνατή, ανθρώπινη ιστορία και όχι με τα δεδομένα που υποδηλώνουν ότι οι δωρεές θα ήταν πιο συνετό να επενδυθούν στην ανάπτυξη.

Η κυβέρνηση υπέκυψε στις πιέσεις του κοινού και των χορηγών βοήθειας. Η Επιτροπή Παρακολούθησης της Ξηρασίας ζήτησε από τη Γιούμα να μοιραστεί τα δεδομένα της ιστορίας της, επειδή, όπως ισχυρίστηκε, δεν γνώριζε ότι η κατάσταση είχε φτάσει σε τόσο απελπιστικό σημείο, αν και η ίδια υπηρεσία είχε προσπαθήσει να τη χρεώσει για πρόσβαση στα δεδομένα, όταν εκείνη ξεκίνησε την έρευνά της. Με βάση τα στοιχεία της Γιούμα για τη λειψυδρία, το Υπουργείο Υδάτων σχεδιάζει να μεταβεί στην Τουρκάνα, για να κατασκευάσει περισσότερες γεωτρήσεις. Η κυβέρνηση, μέσω του Υπουργείου Σχεδιασμού και Αποκέντρωσης, απελευθέρωσε 2,3 δισεκατομμύρια σελίνια Κένυας (27 εκατομμύρια δολάρια) για τη διανομή βοηθημάτων στην κομητεία Τουρκάνα, μια εξέλιξη που η Γιούμα παρακολούθησε στενά. Λόγω της μαζικής αντίδρασης στην ιστορία από ιδιώτες και οργανώσεις, η νομοθεσία για τη διατροφική βιωσιμότητα που ανακατεύθυνε τη βοήθεια, τελικά, προωθήθηκε στη Γερουσία τον Μάιο του

ίδιου έτους[224]. Η Γιούμα συνεχίζει να γράφει θέματα με βάση τα δεδομένα σχετικά με το χάσμα ανάμεσα στην αντίληψη του κοινού, τα προγράμματα των χορηγών βοήθειας και την πολιτική, συμπεριλαμβανομένου του θέματος *Teen Mums of Kwale*, μιας έρευνας σχετικά με τον αντίκτυπο της χρήσης αντισυλληπτικών στα ποσοστά εφηβικής εγκυμοσύνης σε ένα κατεξοχήν συντηρητικό τμήμα της χώρας (NTV Kenya, 2016).

Πώς εξασφαλίζεται ότι το δικαστικό μας σύστημα προστατεύει τους περιθωριοποιημένους;

Στο Αφγανιστάν, η ομάδα δεδομένων του *Pajhwok Afghan News* χρησιμοποίησε δεδομένα για να διερευνήσει τον αντίκτυπο δύο πολιτικών που χαρακτηρίζονται ως καθοριστικές για την πορεία της χώρας προς τη δικαιοσύνη: τον νόμο του Αφγανιστάν για την Εξάλειψη της Βίας κατά των Γυναικών (2009) και την Εθνική Στρατηγική του Αφγανιστάν για τον Έλεγχο των Ναρκωτικών (2012-2016). Εντόπισε δύο απροσδόκητα θύματα αυτών των πολιτικών: τις κακοποιημένες γυναίκες και τους εργάτες της υπαίθρου. Παρόλο που το Αφγανιστάν δεν έχει νόμο για την πρόσβαση στην πληροφόρηση, πολλοί οργανισμοί που λαμβάνουν χρηματοδότηση από χορηγούς βοήθειας, συμπεριλαμβανομένων των Υπουργείων Γυναικείων Υποθέσεων και Καταπολέμησης των Ναρκωτικών, είναι συμβατικά υποχρεωμένοι να παρέχουν αυτά τα δεδομένα.

Πέντε χρόνια μετά την έναρξη ισχύος του νόμου για την ενδοοικογενειακή βία, το *Pajhwok Afghan News* θέλησε να παρακολουθήσει την τύχη των κακοποιητών και των κακοποιημένων. Η ομάδα εξασφάλισε τα στοιχεία για τις 21.000 υποθέσεις κακοποίησης από το Υπουργείο Γυναικείων Υποθέσεων και από διάφορες υπηρεσίες του ΟΗΕ που έχουν αναλάβει την παρακολούθηση των περιστατικών, από την καταγραφή μέχρι τις τελικές ετυμηγορίες και τη διαμεσολάβηση. Διαπίστωσαν ότι στη χειρότερη χώρα του κόσμου για να είσαι γυναίκα, ο νόμος που επαινέθηκε ευρέως κατεύθυνε τις γυναίκες σε μια τοπική διαδικασία διαμεσολάβησης που εδράζεται στον παραδοσιακό σοβινισμό, η οποία συνήθως τις οδηγεί πίσω στον θύτη τους (Munsef & Salehai, 2016). Δύο χρόνια αργότερα, το Παρατηρητήριο Ανθρωπίνων Δικαιωμάτων δημοσίευσε μια μελέτη που επιβεβαίωνε τα ευρήματα του *Pajhwok*, δηλαδή ότι τόσο ο νόμος όσο και η διαμεσολάβηση άφησε απροστάτευτες τις γυναίκες του Αφγανιστάν (United Nations Assistance Mission in Afghanistan, 2018). Ακόμη και αν περισσότερες γυναίκες είχαν πρόσβαση στο δικαστικό σύστημα, το οποίο επαίρεται για το υψηλό ποσοστό καταδίκης των κακοποιών, παραμένει το ακανθώδες ζήτημα του τι πρέπει να γίνει με τις διαζευγμένες γυναίκες σε μια κοινωνία όπου οι γυναίκες δεν εργάζονται.

Παρόμοιες πρακτικές προκλήσεις προκύπτουν κατά την εφαρμογή της στρατηγικής του Αφγανιστάν για τα ναρκωτικά. Το Γραφείο των Ηνωμένων Εθνών για τον Έλεγχο των Ναρκωτικών και την Πρόληψη του Εγκλήματος εξασφάλισε πρόσβαση σε κρατούμενους που είχαν καταδικαστεί για ναρκωτικά, κάτι που συμβαίνει σπανίως, και παρέδωσε τα ανεπεξέργαστα δεδομένα της έρευνας στην ομάδα του *Pajhwok*. Η ανάλυση των ευρημάτων της έρευνας αποκάλυψε ότι η πολιτική αυτή οδήγησε στη φυλακή κυρίως φτωχούς και αναλφάβητους οδηγούς και αγρότες, ενώ οι περισσότεροι βαρόνοι των ναρκωτικών κυκλοφορούν ελεύθεροι (Barakzai & Wardak, 2016). Οι περισσότεροι, επίσης, ανέφεραν ότι σκόπευαν να επιστρέψουν αμέσως στο εμπόριο ναρκωτικών όταν αποφυλακιστούν, καθώς αυτός είναι ο μόνος τρόπος για να συντηρήσουν τις οικογένειές τους στις απομονωμένες αγροτικές περιοχές.

Αυτές οι ιστορίες εξυπηρέτησαν έναν τριπλό σκοπό για την ομάδα δεδομένων του *Pajhwok*: εξέτασαν τις πολιτικές που έχουν αναπτυχθεί από την οπτική γωνία της νομοθεσίας της Δύσης, ανέδειξαν τις συνέπειες της οικονομικής περιθωριοποίησης τόσο ως προς το φύλο όσο και ως προς την περιοχή κατοικίας και προσέφεραν περιεχόμενο δημόσιου ενδιαφέροντος με βάση τα δεδομένα στις γλώσσες νταρί, παστού και στα αγγλικά για ένα πολυποίκιλο ακροατήριο.

224. http://kenyalaw.org/kl/fileadmin/pdfdownloads/bills/2014/TheFoodSecurityBill2014.pdf

Πώς εξασφαλίζουμε ποιοτική εκπαίδευση για όλους;

Η πρόσβαση στην εκπαίδευση, που συχνά θεωρείται ένας μεγάλος εξισορροπιστής, επέτρεψε στις περιθωριοποιημένες κοινότητες να ποσοτικοποιήσουν την αποτυχία της κυβέρνησης να παράσχει βασικές δημόσιες υπηρεσίες και να ωθήσει τους τοπικούς ηγέτες σε μεταρρυθμίσεις. Με μια σειρά θεμάτων, ο προγραμματιστής-δημοσιογράφος Αμπντούλ Σαλάμ Αφρίντι δημιούργησε ένα ρεπορτάζ ειδίκευσης γύρω από την πρόσβαση στην εκπαίδευση μεταξύ των μη προνομιούχων, το οποίο τον οδήγησε στη βραχεία λίστα για τα Data Journalism Awards. Στην πρώτη του ιστορία, χρησιμοποίησε επίσημα κυβερνητικά στατιστικά στοιχεία για την εκπαίδευση και δεδομένα από πανεθνικές εκπαιδευτικές έρευνες, για να δείξει ότι οι γονείς στις απομακρυσμένες φυλετικές περιοχές του Περάσματος Χάιμπερ, οι οποίοι από απελπισία έστελναν όλο και περισσότερα παιδιά σε ιδιωτικά σχολεία, έκαναν μια κακή επένδυση. Η ιστορία του έδειξε ότι οι περισσότεροι απόφοιτοι τόσο των δημόσιων όσο και των ιδιωτικών σχολείων αποτυγχάνουν στις βασικές υποχρεωτικές εξετάσεις (Afridi, 2017a). Περαιτέρω ιστορίες για τη δημόσια εκπαίδευση στο Πέρασμα Χάιμπερ και στις Ομοσπονδιακά Διοικούμενες Φυλετικές Περιοχές, από όπου κατάγεται ο ίδιος ο Σαλάμ, εξετάζουν λεπτομερέστερα τους λόγους πίσω από την αποτυχία των σχολείων (Afridi, 2017b· Afridi, 2018).

Μια άλλη ιστορία, που βασίστηκε σε καταλόγους μαθητών για το εθνικό πρόγραμμα επαγγελματικής κατάρτισης και σε εθνικούς πίνακες εύρεσης εργασίας, αποκάλυψε ένα τεράστιο χάσμα μεταξύ των δεξιοτήτων και της ζήτησης στην αγορά. Η έρευνα αποκάλυψε ότι η χώρα εκπαιδεύει πληροφορικάριους και αισθητικούς, ενώ χρειάζεται οδηγούς και εργάτες χαλυβουργίας. Έτσι, περισσότεροι από τους μισούς αποφοίτους μένουν άνεργοι, και ο βασικός λόγος είναι αυτοί που βρίσκονται πίσω από το πρόγραμμα. Με τη χρηματοδότηση του αναπτυξιακού ταμείου της γερμανικής κυβέρνησης, GiZ, η κυβέρνηση του Πακιστάν έκανε τη δική της ανάλυση, κατέληξε στο ίδιο συμπέρασμα και αναθεώρησε γρήγορα το πρόγραμμα, προσθέτοντας νέα προσφερόμενα μαθήματα, προσαρμοσμένα σε πιο αναγκαίες δεξιότητες για την αγορά εργασίας (Afridi, 2017c).

Ένα εγγενές πλεονέκτημα του ρεπορτάζ με βάση τα δεδομένα μεταξύ των περιθωριοποιημένων κοινοτήτων είναι ότι ο δημοσιογράφος μπορεί να συνεχίσει να ασχολείται με την ιστορία και αφού ξεχαστεί το αρχικό σκάνδαλο. Κοινό χαρακτηριστικό αυτών των ιστοριών είναι ότι χρησιμοποιούν δεδομένα όχι μόνο για να παρουσιάσουν το πρόβλημα, αλλά και για να πουν τι μπορεί να γίνει για την αντιμετώπισή του. Αυτοί οι δημοσιογράφοι συγκέντρωσαν δεδομένα για να μετρήσουν το πρόβλημα, τον αντίκτυπο, τα αίτια και τη λύση. Σε παγκόσμιο επίπεδο, υπάρχει μια ώθηση προς την προσιτή δημοσιογραφία δεδομένων από περιθωριοποιημένες κοινότητες, σχετικά με αυτές και για λογαριασμό τους, ώστε να κατακτηθεί η εμπιστοσύνη τους και οι ίδιες να συμπεριληφθούν στον δημόσιο βίο.

Δημοσιογραφία δεδομένων υπό περιορισμούς

Μεγάλο μέρος της διαφωνίας στην ακαδημαϊκή κοινότητα σχετικά με τη μακροπρόθεσμη βιωσιμότητα της δημοσιογραφίας δεδομένων προέρχεται από τη διχογνωμία σχετικά με το αν ο στόχος της είναι να παράγει προβεβλημένα διαδραστικά προϊόντα ή ρεπορτάζ δημοσίου ενδιαφέροντος βασισμένα σε γεγονότα. Οι δημοσιογράφοι στις αναπτυσσόμενες χώρες χρησιμοποιούν τα δεδομένα, για να απαντήσουν σε βασικά ερωτήματα σχετικά με τις θεσμοθετημένες διακρίσεις λόγω φύλου, τα μεροληπτικά συστήματα δικαιοσύνης και την εσκεμμένη παραμέληση των λιμοκτονούντων, αλλά και για να μεταφέρουν αυτές τις πληροφορίες σε όσο το δυνατόν περισσότερους ανθρώπους. Αυτό το κάνουν γνωρίζοντας ότι πρόκειται για πολύπλοκα προβλήματα και ότι η αλλαγή των πολιτικών ως αποτέλεσμα των δικών τους ερευνών εξακολουθεί να είναι εξαιρετικά απίθανη. Οι δημοσιογράφοι δεδομένων στη Δύση, με πρόσβαση σε καλύτερους πόρους, δεδομένα και ελεύθερα μέσα ενημέρωσης, και με κυβερνήσεις που ανταποκρίνονται καλύτερα, συχνά δεν εκμεταλλεύονται την ευκαιρία να εξασφαλίσουν ότι σε τέτοιους ταραχώδεις καιρούς, αντιμετωπίζουμε τις ανάγκες πληροφόρησης των περιθωριοποιημένων πολιτών και θέτουμε τις κυβερνήσεις προ των ευθυνών τους.

Τα περισσότερα από αυτά τα προβλήματα ήταν αόρατα στο παρελθόν και θα ξαναγίνουν αόρατα, αν οι δημοσιογράφοι σταματήσουν να ασχολούνται μαζί τους. Η καλύτερη εκδοχή της δημοσιογραφίας δεδομένων προέρχεται από εκείνους για τους οποίους η κοινωνία έχει αποφασίσει ότι δεν μετρούν, έχει περιεχόμενο σχετικό με αυτούς και γίνεται για λογαριασμό τους. Ευτυχώς, η κοινωνία των πολιτών, οι ακτιβιστές, οι ακαδημαϊκοί, οι κυβερνήσεις και άλλοι συνεργάζονται, για να λογαριάσουν περισσότερο τους αποκλεισμένους. Οι δημοσιογράφοι διαδραματίζουν ζωτικό ρόλο, προκειμένου να διασφαλίσουν ότι αυτά είναι προβλήματα για τα οποία οι άνθρωποι μιλούν και εργάζονται, προκειμένου να επιλυθούν. Όλα ήταν φρικτά, είναι φρικτά και θα παραμείνουν φρικτά, αν δεν συνεχίσουμε να λογαριάζουμε τους ανθρώπους και να μιλάμε. Χρόνο με τον χρόνο, πρέπει να λογαριάζουμε τους πεινασμένους, τους κακοποιημένους, τους φυλακισμένους, τους ανεκπαίδευτους, αυτούς που δεν ακούγονται, γιατί παντού στη γη, τα πράγματα είναι φρικτά για κάποιους.

Παραπομπές

Afridi, A. S. (2017a) «In KP, parents still prefer private over public schools», *News Lens Pakistan*, 18 Φεβρουαρίου. Διαθέσιμο σε: http://www.newslens.pk/in-kp-parents-still-prefer-private-over-public-schools/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Afridi, A. S. (2017b) «Half of FATA schools functioning in dire straits», *News Lens Pakistan,* 16 Ιουνίου. Διαθέσιμο σε: http://www.newslens.pk/half-fata-schools-functioning-dire-straits/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Afridi, A. S. (2017c) «TVET Reform programmes targeting wrong skills», *News Lens Pakistan,* 16 Σεπτεμβρίου. Διαθέσιμο σε: http://www.newslens.pk/tvet-reform-programmes-targeting-wrong-skills/ (Προπελάστηκε: 18 Δεκεμβρίου 2021).

Afridi, A. S. (2018) «Despite huge investment the outlook of Education in KP remains questionable», *News Lens Pakistan*, 2 Μαρτίου. Διαθέσιμο σε: http://www.newslens.pk/despite-huge-investment-outlook-education-in-kp/ (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Barakzai, N. A., & Wardak, A. (2016) «Most jailed drug offenders are poor, illiterate», *Pajhwok Afghan* News, 28 Σεπτεμβρίου. Διαθέσιμο σε: https://www.pajhwok.com/en/2016/09/28/most-jailed-drug-offenders-are-poor-illiterate (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Juma, M. (2014) «When will Kenya have enough to feed all its citizens?», *Daily Nation*, 28 Ιανουαρίου. Διαθέσιμο σε: https://www.nation.co.ke/lifestyle/dn2/When-will-Kenya-have-enough-to-feed-all-its-citizens-/957860-2163092-p8mgj9z/index.html (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

Munsef, A. Q., & Salehai, Z. (2016) «Cases of violence against women: Is mediation the best option?», *Pajhwok Afghan News*, 11 Μαΐου. Διαθέσιμο σε: https://www.pajhwok.com/en/2016/05/11/cases-violence-against-women-mediation-best-option (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

NTV Kenya (2016) *#TeenMumsOfKwale: Primary school girls in Kwale using contraceptives to prevent unwanted pregnancies* [YouTube], 2 Οκτωβρίου. Διαθέσιμο σε: https://www.youtube.com/watch?v=xMx5lRHbw3g (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

United Nations Assistance Mission in Afghanistan (2018) «Injustice and impunity: Mediation of criminal offences of violence against women», *United Nations Office of the High Commissioner for Human Rights.* Διαθέσιμο σε: https://unama.unmissions.org/sites/default/files/unama_ohchr_evaw_report_2018_injustice_and_impunity_29_may_2018.pdf (Προσπελάστηκε: 18 Δεκεμβρίου 2021)

Η συγγραφέας

Η Εύα Κονσταντάρας είναι δημοσιογράφος δεδομένων που ειδικεύεται στη δημιουργία ομάδων δεδομένων σε οργανισμούς μέσων ενημέρωσης στον Παγκόσμιο Νότο.

8.4
Διδάσκοντας Δημοσιογραφία Δεδομένων

Cheryl Phillips

Περίληψη: Η διδασκαλία της δημοσιογραφίας δεδομένων ξεκινά με τη διδασκαλία της κριτικής σκέψης.

Λέξεις-κλειδιά: κριτική σκέψη, εκπαίδευση στη δημοσιογραφία δεδομένων, προγραμματισμός, συνεργασία, πρακτική δεδομένων, συνεργασίες ερευνητών-δημοσιογράφων

Στο Κρατικό Πανεπιστήμιο του Τέξας, η καθηγήτρια Σίντι Ρόγιαλ διδάσκει ανάπτυξη ιστοσελίδων[225]. Μερικές χιλιάδες χιλιόμετρα ανατολικά, στο Πανεπιστήμιο της Φλόριντα, η Μίντι Μακάνταμς, κάτοχος της έδρας Knight για τις Τεχνολογίες Δημοσιογραφίας και τη Δημοκρατική Διαδικασία, και ο Αναπληρωτής Καθηγητής Νόρμαν Λούις, διδάσκουν πληθώρα μαθημάτων, από προγραμματισμό μέχρι παραδοσιακή δημοσιογραφία δεδομένων και ανάπτυξη εφαρμογών. Ο Αλμπέρτο Κάιρο, κάτοχος της έδρας Knight για την Οπτική Δημοσιογραφία στη Σχολή Επικοινωνίας του Πανεπιστημίου του Μαϊάμι, διδάσκει ένα ολόκληρο πρόγραμμα που επικεντρώνεται στην οπτικοποίηση δεδομένων.

Προς τα βόρεια, οι φοιτητές του Πανεπιστημίου Columbia και του CUNY (The City University of New York) παρακολουθούν μαθήματα που διδάσκονται από επαγγελματίες δημοσιογράφους δεδομένων από το NBC και τους *New York Times*, μαθαίνοντας τα βασικά στοιχεία του ερευνητικού ρεπορτάζ μαζί με την ανάλυση δεδομένων. Στο Πανεπιστήμιο του Μέριλαντ, τα μαθήματα δικαίου των μέσων μαζικής ενημέρωσης παρακολουθούν τακτικά τη διαδικασία υποβολής αιτημάτων δημόσιων αρχείων για δημοσιογραφικά έργα. Στη Νεμπράσκα, ο Ματ Γουέιτ διδάσκει στους μαθητές να οπτικοποιούν δεδομένα χρησιμοποιώντας Lego. Στο Πανεπιστήμιο του Στάνφορντ, διδάσκουμε βασική ανάλυση δεδομένων, προγραμματισμό σε Python και R και βασική οπτικοποίηση δεδομένων, περισσότερο για την κατανόηση παρά για την παρουσίαση.

Οι καθηγητές δημοσιογραφίας δεδομένων –πολλοί από τους οποίους ξεκίνησαν ως επαγγελματίες– διδάσκουν με διάφορους τρόπους σε ολόκληρο τον κόσμο (και τα παραπάνω παραδείγματα είναι μόνο από προγράμματα στις Ηνωμένες Πολιτείες). Ποιο πρόγραμμα είναι η πραγματική δημοσιογραφία δεδομένων; Ερώτηση-παγίδα: όλα τους είναι. Πώς να τη διδάξουμε, λοιπόν;

Με τον ίδιο τρόπο που διδάσκουμε κάθε είδους δημοσιογραφικό μάθημα. Κάθε εξειδίκευση –από την αθλητική δημοσιογραφία μέχρι το επιχειρηματικό ρεπορτάζ ή το επιστημονικό ρεπορτάζ– περιλαμβάνει ειδικές δεξιότητες και γνώσεις που πρέπει να κατακτήσει κανείς. Ωστόσο, η καθεμία στηρίζεται στις βασικές αρχές της δημοσιογραφίας.

Με τον ίδιο τρόπο, η εκπαίδευση στη δημοσιογραφία δεδομένων θα πρέπει να ξεκινά από τα βασικά. Με αυτό δεν εννοώ την εκμάθηση υπολογιστικών φύλλων, αν και πιστεύω ότι αυτό

225. Για το κεφάλαιο αυτό οφείλω να ευχαριστήσω τον Τσαρλς Μπέρετ, συν-συγγραφέα του βιβλίου *Teaching Data and Computational Journalism*, το οποίο εκδόθηκε με την υποστήριξη του Πανεπιστημίου Columbia και του Ιδρύματος John S. και James L. Knight.

μπορεί να είναι ιδανικό για την κατανόηση πολλών βασικών αρχών της δημοσιογραφίας δεδομένων. Δεν υπάρχει τίποτα καλύτερο από το να κατανοήσεις την εγγενή ακαταστασία των εισαγόμενων δεδομένων, βάζοντας τους φοιτητές να συμμετάσχουν σε μια άσκηση που περιλαμβάνει την εισαγωγή πληροφοριών σε μικρά κουτάκια στην οθόνη ενός υπολογιστή. Επίσης, δεν εννοώ την εκμάθηση ενός συγκεκριμένου τρόπου προγραμματισμού, από την Python μέχρι την R, αν και πιστεύω ότι και οι δύο γλώσσες έχουν πολλά οφέλη. Δεν υπάρχει τίποτα καλύτερο από το να βλέπεις έναν φοιτητή να εκτελεί μια γραμμή κώδικα και να παίρνει ένα αποτέλεσμα για το οποίο θα χρειάζονταν τέσσερα ή περισσότερα βήματα σε ένα υπολογιστικό φύλλο.

Η εκμάθηση της δημοσιογραφίας δεδομένων ξεκινά κατανοώντας πώς να σκεφτόμαστε κριτικά για τις πληροφορίες και πώς αυτές μπορούν να συλλεχθούν, να κανονικοποιηθούν και να αναλυθούν για δημοσιογραφικούς σκοπούς. Ξεκινά εντοπίζοντας την ιστορία και θέτοντας τις ερωτήσεις που θα οδηγήσουν σε αυτήν.

Και οι καθηγητές δημοσιογραφίας πιθανότατα γνωρίζουν ήδη τη μορφή που μπορούν να πάρουν αυτές οι ερωτήσεις:

- Ποιος δημιούργησε τα δεδομένα;
- Τι υποτίθεται ότι πρέπει να περιλαμβάνουν τα δεδομένα;
- Πότε ενημερώθηκαν για τελευταία φορά τα δεδομένα;
- Ποιο τμήμα του κόσμου αντιπροσωπεύουν τα δεδομένα;
- Γιατί χρειαζόμαστε αυτά τα δεδομένα για να πούμε την ιστορία μας;
- Πώς βρίσκουμε τις απαντήσεις στις ερωτήσεις που θέλουμε να θέσουμε σε αυτά τα δεδομένα;

Έτσι, διαμορφώνονται τα προγράμματα σπουδών, χρησιμοποιώντας υπολογιστικά φύλλα ή SQL (Structured Query Language) ή Python ή R. Δεν έχει σημασία –όπως δεν έχει σημασία το γεγονός ότι κάποτε γνώριζα κάτι που λεγόταν Paradox for DOS. Αυτό που έχει σημασία είναι να ξέρουμε τα βήματα που πρέπει να ακολουθήσουμε για τη συλλογή και την ανάλυση δεδομένων. Η οπτικοποίηση είναι το κλειδί τόσο στην ανάλυση όσο και στην παρουσίαση, αλλά αν προηγηθεί η οπτική ανάλυση για την κατανόηση, τότε η παρουσίαση ακολουθεί πιο εύκολα.

Αυτό το κεφάλαιο περιέχει μια ποικιλία προσεγγίσεων και σημείων εκκίνησης σχετικά με τον τρόπο διδασκαλίας της δημοσιογραφίας δεδομένων, με βάση το ποιος είστε, τι επίπεδο προγράμματος διαθέτετε και πώς μπορείτε να οικοδομήσετε συνεργατικές προσπάθειες. Αφού παρουσιάσει την προσέγγιση της «βαλίτσας» για τη διδασκαλία της δημοσιογραφίας δεδομένων, διερευνά μοντέλα ενός μαθήματος, μοντέλα αντίστροφης τάξης, ολοκληρωμένα μοντέλα και πειράματα συνδιδασκαλίας σε διάφορα επιστημονικά πεδία.

Ένα μάθημα είναι το μόνο που μπορείτε να κάνετε: Πακετάροντας τη βαλίτσα

Όταν πηγαίνουμε για κάμπινγκ με το αυτοκίνητο, αστειευόμαστε συνέχεια, λέγοντας ότι παίρνουμε μαζί μας τα πάντα, συμπεριλαμβανομένου του νεροχύτη της κουζίνας. Το κόλπο είναι να ξέρετε τι μπορείτε να πακετάρετε και τι θα σας παραφορτώσει σε σημείο αντιπαραγωγικότητας. Ο νεροχύτης της κουζίνας είναι στην πραγματικότητα μια μικρή, πτυσσόμενη, υφασμάτινη λεκάνη.

Εάν διδάσκετε μόνο ένα μάθημα και είστε ο μοναδικός καθηγητής δημοσιογραφίας δεδομένων, μην προσπαθήσετε να στριμώξετε πάρα πολλά, συμπεριλαμβανομένης της ανάλυσης δεδομένων με υπολογιστικά φύλλα και SQL, της επεξεργασίας δεδομένων με Python, της ανάλυσης με R και του σχεδιασμού οπτικοποίησης δεδομένων με d3.js, όλα σε ένα τρίμηνο ή σε ένα εξάμηνο.

Επιλέξτε τα εργαλεία που είναι ζωτικής σημασίας. Εξετάστε το ενδεχόμενο το μάθημα να βασίζεται, τουλάχιστον εν μέρει, σε εργασίες. Όπως και να έχει, ακολουθήστε τα βήματα. Επαναλάβετε και μην το περιπλέκετε. Επικεντρωθείτε στη δημοσιογραφία που προκύπτει από τη χρήση των εργαλείων που επιλέγετε.

Το 2014 και το 2015, ο Τσαρλς Μπέρετ από το Πανεπιστήμιο Columbia και εγώ πραγματοποιήσαμε μια έρευνα και εκτενείς συνεντεύξεις με δημοσιογράφους δεδομένων και καθηγητές δημοσιογραφίας. Οι περισσότεροι από αυτούς που διδάσκουν δημοσιογραφία δεδομένων ανέφεραν ότι, ξεκινώντας με ένα υπολογιστικό φύλλο, εισάγουν την έννοια των δομημένων δεδομένων στους φοιτητές με τρόπο που είναι εύκολα κατανοητός.

Ένα άλλο βήμα είναι να αυξήσουμε την πολυπλοκότητα, ώστε να συμπεριλάβουμε και άλλες πολύτιμες τεχνικές της δημοσιογραφίας δεδομένων: να προχωρήσουμε πέρα από τις ταξινομήσεις, τα φίλτρα και τις ομαδοποιήσεις δεδομένων ή να ενώσουμε διαφορετικά σύνολα δεδομένων, για να βρούμε μοτίβα που διαφορετικά δεν θα είχαν εντοπιστεί.

Αυτό δεν σημαίνει, όμως, ότι πρέπει να προσθέσετε μυριάδες νέα εργαλεία ή ακόμη και να επιλέξετε το πιο σύγχρονο εργαλείο. Μπορείτε να εισαγάγετε τους φοιτητές σε αυτό το επόμενο επίπεδο, χρησιμοποιώντας οποιαδήποτε τεχνολογία ταιριάζει σε εσάς και στο πρόγραμμα δημοσιογραφίας του ιδρύματός σας. Εάν πρόκειται για ένα πανεπιστημιακό πρόγραμμα όπου κάθε φοιτητής διαθέτει τη MS Access, τότε χρησιμοποιήστε την, αλλά πηγαίνετε πέρα από τη διεπαφή «point-and-click», για να βεβαιωθείτε ότι οι φοιτητές κατανοούν την SQL πίσω από κάθε query[226]. Ή χρησιμοποιήστε τη MySQL. Ή χρησιμοποιήστε την Python σε ένα Jupyter Notebook. Ή χρησιμοποιήστε την R και το R Studio, το οποίο διαθέτει μερικά υπέροχα πακέτα για ερωτήματα τύπου SQL.

Ο στόχος είναι να διδάξουμε δημοσιογραφία στους φοιτητές, βοηθώντας τους παράλληλα να κατανοήσουν το τι πρέπει να συμβεί και ότι υπάρχουν πολλοί τρόποι για να επιτύχουμε παρόμοιες λειτουργίες με δεδομένα στην υπηρεσία της αφήγησης μιας ιστορίας.

Και πάλι, μην το περιπλέκετε. Μην αναγκάζετε τους φοιτητές να περάσουν από σαράντα κύματα για τα τεχνολογικά εργαλεία. Χρησιμοποιήστε τα εργαλεία, για να κάνετε τη δημοσιογραφία πιο ισχυρή και ευκολότερη στην εκπόνησή της. Για να επανέλθουμε στην αναλογία του κάμπινγκ με το αυτοκίνητο, πακετάρετε μόνο ό,τι χρειάζεστε στην τάξη σας. Μη φέρετε το αλυσοπρίονο, αν το μόνο που χρειάζεστε είναι ένα τσεκούρι ή ένα μαχαίρι τσέπης.

Όμως, μόλις καταφέρετε να εξασφαλίσετε τη μία τάξη, σκεφτείτε πέρα από αυτό το μοντέλο της μίας τάξης. Σκεφτείτε τρόπους ενσωμάτωσης στοιχείων δημοσιογραφίας δεδομένων σε όλο το τμήμα ή τη σχολή. Βρείτε κοινά κίνητρα με άλλες τάξεις. Μπορείτε να συνεργαστείτε με συναδέλφους που διδάσκουν ένα βασικό μάθημα ειδησεογραφικού ρεπορτάζ, για να διαπιστώσετε τι περισσότερο θα τους ενδιέφερε να μάθουν οι φοιτητές τους σχετικά με την ενσωμάτωση δεδομένων.

Ορισμένοι καθηγητές δημοσιογραφίας έχουν πειραματιστεί με μοντέλα «αντίστροφης τάξης», για να εξισορροπήσουν τη διδασκαλία δεξιοτήτων, την κριτική σκέψη και τον θεωρητικό προβληματισμό. Οι φοιτητές μπορούν να παρακολουθούν τα σεμινάρια με τον δικό τους ρυθμό και να επικεντρώνονται στην επίλυση προβλημάτων με τους διδάσκοντες κατά τη διάρκεια του μαθήματος, καθώς και στην εκμάθηση άλλων μεθόδων για την αντιμετώπιση ποικίλων προκλήσεων της δημοσιογραφίας δεδομένων. Για παράδειγμα, η καθηγήτρια Μακάνταμς από το Πανεπιστήμιο της Φλόριντα ακολουθεί ένα μοντέλο αντίστροφης τάξης για το μάθημά της στον σχεδιασμό διαδικτυακών εφαρμογών («web apps»).

Ένα πλεονέκτημα αυτού του τύπου διδασκαλίας είναι ότι λαμβάνει υπόψη δημοσιογράφους πολλών διαφορετικών επιπέδων ως προς τις δεξιότητές τους. Σε ορισμένες περιπτώσεις, ένα μάθημα δημοσιογραφίας μπορεί να προσελκύσει το ενδιαφέρον ενός φοιτητή που είναι έμπειρος στην επιστήμη των υπολογιστών και, ταυτόχρονα, ενός φοιτητή που δεν έχει χρησιμοποιήσει ποτέ του υπολογιστικό φύλλο.

Η διδασκαλία δημοσιογραφίας δεδομένων, όμως, υπερβαίνει τις αντίστροφες τάξεις. Αυτό σημαίνει ότι πρέπει να αναζητήσουμε άλλους τρόπους διδασκαλίας των εννοιών της δημοσιογρα-

226. ΣτΕ: Ερώτημα στη βάση δεδομένων.

φίας δεδομένων. Στο SRCCON, ένα τακτικό unconference[227], η Σάρα Κοέν, κάτοχος της Έδρας Knight στη Δημοσιογραφία Δεδομένων στο Κρατικό Πανεπιστήμιο της Αριζόνας και βραβευμένη με Πούλιτζερ δημοσιογράφος, η οποία εργαζόταν μέχρι πρότινος στους *New York Times*, υποστήριξε τη χρήση άλλων αναλογικών δραστηριοτήτων για την εμπλοκή των φοιτητών. Η Κοέν και ο Γουέιτ, καθηγητής πρακτικής στο Πανεπιστήμιο της Νεμπράσκα, παρουσίασαν την ιδέα ενός κοινού προγράμματος σπουδών με ενότητες που θα μπορούσαν να χρησιμοποιηθούν από εκπαιδευτικούς παντού. Ο στόχος είναι να δημιουργηθεί ένα σύστημα όπου οι καθηγητές δεν θα χρειάζεται να στήνουν τα πάντα από το μηδέν. Στο συνέδριο το καλοκαίρι του 2018, ήταν επικεφαλής μιας ομάδας συμμετεχόντων που συνεισέφεραν πιθανές ενότητες για την προσπάθεια αυτή. «Προσπαθούμε να μην αντιμετωπίζουμε με θρησκευτική ευλάβεια αυτά τα πράγματα [τα εργαλεία]», είπε η Κοέν στην ομάδα, υποστηρίζοντας ότι, αντίθετα, πρέπει να επικεντρωθούμε στις «θεμελιώδεις αξίες της δημοσιογραφίας και στις θεμελιώδεις αξίες της ανάλυσης δεδομένων».

Τώρα, ένα αποθετήριο στο GitHub είναι έτοιμο και λειτουργεί με συνεισφέροντες που προσθέτουν και βελτιώνουν τις ενότητες για χρήση στην εκπαίδευση της δημοσιογραφίας δεδομένων[228]. Το αποθετήριο προσφέρει, επίσης, συνδέσμους προς επιπλέον υλικό για τη διδασκαλία της δημοσιογραφίας δεδομένων, συμπεριλαμβανομένου αυτού του εγχειριδίου.

Μερικές δυνατότητες για ενότητες περιλαμβάνουν την ερμηνεία δημοσκοπήσεων ή μελετών. Η βασική αριθμητική είναι ένα σημαντικό στοιχείο των μαθημάτων δημοσιογραφίας. Η εύρεση δεδομένων στο Διαδίκτυο είναι κάτι άλλο που πυροδοτεί το ενδιαφέρον και μπορεί να δώσει ώθηση σε κάθε τάξη. Επίσης, δεν σημαίνει ότι πρέπει να αφιερώσετε όλο τον ελεύθερο χρόνο σας για τον σκοπό αυτό. Δημιουργήστε μια θεματική ενότητα ή ένα σεμινάριο μία φορά και, στη συνέχεια, αυτό μπορεί να χρησιμοποιηθεί ξανά και ξανά από άλλους. Ή αξιοποιήστε τα πολλά δωρεάν σεμινάρια που υπάρχουν ήδη εκεί έξω. Τα ετήσια συνέδρια που διοργανώνονται από τον οργανισμό Investigative Reporters and Editors (IRE) και το National Institute for Computer-Assisted Reporting (NICAR) παρέχουν ακόμη περισσότερα σεμινάρια για τα μέλη τους πάνω σε όλα τα θέματα, από τους πίνακες pivot μέχρι το scraping και τη χαρτογράφηση.

Μια φορά το τρίμηνο διδάσκω, ως επισκέπτρια για κάποιον συνάδελφό μου, σχετικά με την εύρεση δεδομένων στο Διαδίκτυο. Τα οφέλη περιλαμβάνουν τη δημιουργία ενός δικτύου φοιτητών που ενδιαφέρονται να εξερευνήσουν τη δημοσιογραφία δεδομένων και τη συμμετοχή σε μια συλλογική ατμόσφαιρα με συναδέλφους καθηγητές.

Εάν είναι δυνατόν, εξετάστε το ενδεχόμενο να δημιουργήσετε θεματικές ενότητες που θα μπορούσαν να υιοθετήσουν οι συνάδελφοι αυτοί. Οι περιβαλλοντικοί δημοσιογράφοι θα μπορούσαν να παρουσιάσουν μια ενότητα σχετικά με τις μέσες θερμοκρασίες με την πάροδο του χρόνου, χρησιμοποιώντας ένα υπολογιστικό φύλλο, για παράδειγμα. Αυτό έχει ακόμη ένα δυνητικό όφελος: δείχνετε στους συναδέλφους σας την αξία της δημοσιογραφίας δεδομένων, γεγονός που μπορεί, επίσης, να συμβάλλει στη δημιουργία ενός προγράμματος σπουδών που ενσωματώνει συστηματικά αυτές τις πρακτικές και προσεγγίσεις.

Περισσότερα σχετικά με ένα ολοκληρωμένο μοντέλο ή τη διεπιστημονική διδασκαλία

Ένα πλήρως ολοκληρωμένο μοντέλο σημαίνει ότι περισσότερα από ένα άτομα επενδύουν στη διδασκαλία των εννοιών της δημοσιογραφίας δεδομένων. Έχει, επίσης, τη δυνατότητα να ξεπεράσει τα όρια ενός δημοσιογραφικού προγράμματος. Στο Στάνφορντ, εγκαινιάσαμε το Stanford Open Policing Project και συνεργαστήκαμε με το Poynter, για να εκπαιδεύσουμε

227. ΣτΜ: Μια ανεπίσημη συνάντηση το περιεχόμενο και η δομή της οποίας καθορίζεται από τους συμμετέχοντες, αποφεύγοντας τις ιεραρχικές δομές ενός κλασικού συνεδρίου.
228. https://github.com/datajtext/DataJournalismTextbook

δημοσιογράφους στην ανάλυση δεδομένων αστυνόμευσης. Καθηγητές της μηχανικής και της δημοσιογραφίας έχουν συνεργαστεί, για να διδάξουν διεπιστημονικά μαθήματα και να εκπαιδεύσουν φοιτητές δημοσιογραφίας, φοιτητές νομικής και φοιτητές πληροφορικής. Αυτό είναι σημαντικό, επειδή οι καλύτερες ομάδες συνεργασίας στις αίθουσες σύνταξης περιλαμβάνουν άτομα από διάφορους κλάδους. Πιο πρόσφατα, τα ακαδημαϊκά ιδρύματα όχι μόνο υιοθετούν τέτοια ολοκληρωμένα μοντέλα, αλλά παράγουν έργο που φτάνει στις αίθουσες σύνταξης και ταυτόχρονα διδάσκει τους φοιτητές.

Μόλις τον περασμένο μήνα, το Ίδρυμα Scripps Howard ανακοίνωσε ότι παρέχει δύο επιχορηγήσεις ύψους τριών εκατομμυρίων δολαρίων στο Κρατικό Πανεπιστήμιο της Αριζόνας και στο Πανεπιστήμιο του Μέριλαντ, τα οποία θα ιδρύσουν κέντρα ερευνητικού ρεπορτάζ[229]. Τα κέντρα αυτά θα εκπαιδεύουν φοιτητές και θα παράγουν ερευνητικό έργο, αναλαμβάνοντας τον ρόλο του εκδότη και του εκπαιδευτή.

Τα μαθήματα που έχουν μια αποστολή και κινούνται πέρα από την αίθουσα διδασκαλίας είναι πιο συναρπαστικά για τους φοιτητές και μπορούν να προσφέρουν μια πιο ενδιαφέρουσα μαθησιακή εμπειρία. Ένα από τα πιο επιτυχημένα μαθήματα στα οποία έχω συμμετάσχει είναι το μάθημα Νόμος, Τάξη και Αλγόριθμοι («Law, Order and Algorithms»), το οποίο διδάχθηκε την άνοιξη του 2018 από εμένα και τον Επίκουρο Καθηγητή Μηχανολογίας Σαράντ Γκόελ. Ο τίτλος του μαθήματος είναι του Γκόελ, αλλά εμείς προσθέσαμε μια ανατροπή. Η ανάλογη δική μου τάξη συναντήθηκε με την τάξη του Γκόελ. Στις δύο αυτές τάξεις, διδάξαμε φοιτητές πληροφορικής και μηχανικής, φοιτητές νομικής και δημοσιογραφίας. Οι ομάδες των φοιτητών παρήγαγαν προηγμένες στατιστικές αναλύσεις, white paper και δημοσιογραφικές εργασίες από τα έργα τους. Ο Γκόελ και εγώ διδάξαμε ο καθένας στον τομέα της ειδικότητάς μας. Θέλω να πιστεύω ότι έμαθα κάτι για το δίκαιο και για τους τρόπους με τους οποίους οι αλγόριθμοι μπορούν να χρησιμοποιηθούν για το καλό και για το κακό, αλλά και ότι ο καθηγητής Γκόελ έμαθε κάτι για το τι απαιτείται, προκειμένου να κάνεις ερευνητική δημοσιογραφία και δημοσιογραφία δεδομένων.

Όσον αφορά στους φοιτητές, ο βασισμένος στις εργασίες χαρακτήρας του μαθήματος σήμαινε πως μάθαιναν ό,τι χρειάζονταν, για να επιτύχουν τους στόχους του έργου της ομάδας τους. Αυτό που αποφύγαμε ήταν να ζητήσουμε από τους φοιτητές να μάθουν τόσα πολλά εργαλεία ή τεχνικές που θα έβλεπαν μόνο σταδιακή πρόοδο. Προσπαθήσαμε να συγκεντρώσουμε μόνο ό,τι ήταν απαραίτητο για την επιτυχία, όπως σε εκείνες τις εκδρομές με το αυτοκίνητο για κάμπινγκ.

Η συγγραφέας

Η Σέριλ Φίλιπς διδάσκει δημοσιογραφία δεδομένων στο Πανεπιστήμιο του Στάνφορντ και είναι η ιδρύτρια του Big Local News (biglocalnews.org), μιας συνεργατικής προσπάθειας ανταλλαγής δεδομένων και πλατφόρμας για δημοσιογράφους.

229. Για να μάθετε περισσότερα σχετικά με τις επιχορηγήσεις για τη δημιουργία κέντρων ερευνητικής δημοσιογραφίας, βλ. Boehm, J. (2018) «Arizona State University, University of Maryland get grants to launch investigative journalism centers», *AZCentral*, 6 Αυγούστου. Διαθέσιμο σε: https://amp.azcentral.com/amp/902340002 (Προσπελάστηκε: 18 Δεκεμβρίου 2021).

8.5
Οργανώνοντας Έργα Δεδομένων με Γυναίκες και Μειονότητες στη Λατινική Αμερική

Eliana A. Vaca Muñoz

Περίληψη: Αυτό το κεφάλαιο πραγματεύεται την οργάνωση έργων δεδομένων με γυναίκες και μειονότητες στη Λατινική Αμερική.

Λέξεις-κλειδιά: αναλογικές οπτικοποιήσεις, μειονότητες, γυναίκες, Λατινική Αμερική, εκπαίδευση στη δημοσιογραφία δεδομένων, οπτικοποίηση δεδομένων

Το *Chicas Poderosas* («Δυνατά Κορίτσια») είναι ένα διακρατικό δίκτυο δημοσιογράφων, σχεδιαστών και προγραμματιστών που εργάζονται για την ανάπτυξη έργων ψηφιακών μέσων από και για γυναίκες και περιθωριοποιημένες κοινότητες σε ολόκληρη τη Λατινική Αμερική. Ως σχεδιάστρια του *Chicas Poderosas*, το έργο μου διερευνά τον ρόλο που μπορεί να διαδραματίσει ο σχεδιασμός ως παράγοντας πολιτισμού και ποικιλομορφίας, μεταξύ άλλων μέσω διεπιστημονικής και συμμετοχικής έρευνας για τη διερεύνηση της πολιτιστικής κληρονομιάς, της ταυτότητας, της οικειοποίησης εδαφών και της αναγνώρισης των γυναικών και των ευάλωτων πληθυσμών.

Το κεφάλαιο αυτό εξετάζει την οργάνωση διαφόρων πρωτοβουλιών του *Chicas Poderosas* στην Κολομβία και στην Κεντρική Αμερική. Καθώς οι κοινωνικές και πολιτιστικές ανισότητες στην Κολομβία διευρύνονται, είναι σημαντικό οι μειονότητες να ακούγονται, να μοιράζονται τις γνώσεις τους και να αντιμετωπίζονται ισότιμα. Για τον σκοπό αυτό, η ομάδα Chicas Colombia διεξήγαγε μια σειρά συνεργατικών εργαστηρίων, με επίκεντρο τη δημοσιογραφία δεδομένων και τις συναφείς πρακτικές των ψηφιακών μέσων ενημέρωσης. Στις επόμενες ενότητες, εξετάζω δύο μεθόδους που χρησιμοποιήσαμε, για να διευκολύνουμε τη συμμετοχή σε αυτά τα εργαστήρια: αναλογική συλλογή δεδομένων και αναλογική οπτικοποίηση δεδομένων. Αυτές οι προσεγγίσεις μπορεί να είναι σχετικές με τις πρακτικές και τις κουλτούρες της δημοσιογραφίας δεδομένων σε κοινότητες και περιοχές όπου η συνδεσιμότητα, οι συσκευές και ο τεχνολογικός αλφαβητισμός δεν μπορούν να θεωρούνται δεδομένα.

Αναλογική συλλογή δεδομένων

Τον Μάιο του 2016, η Chicas Colombia πήγε στη Villa España στο Κιμπντό του Τσοκό, για να συνεργαστεί με γυναίκες και εφήβους που ανήκουν στη συλλογικότητα AJODENIU (Association of Displaced Youth –«Ένωση Εκτοπισμένων Νέων»). Από το 2002 η ομάδα αυτή εργάζεται για την υπεράσπιση των συμφερόντων και των δικαιωμάτων των παιδιών που έχουν εκτοπιστεί από τις περιοχές Τσοκό, Ρίο Σούθιο, Μπογαγιά και Ουραμπά. Όλες αυτές οι περιοχές είναι δύσκολα προσβάσιμες, χωρίς Διαδίκτυο και με λίγες διαθέσιμες υπηρεσίες υποστήριξης. Ως εκ τούτου, τα εργαστήρια ξεκίνησαν με αναλογικές τεχνικές για τη συλλογή ποιοτικών δεδομένων. Με αυτά τα δεδομένα, προσπαθήσαμε να δομήσουμε ιστορίες για θέματα όπως ο αναγκαστικός εκτοπισμός και οι εφηβικές εγκυμοσύνες, καταγράφοντας προφορικές και γραπτές αφηγήσεις[230].

230. https://chicaspoderosas.org/2016/11/22/the-pacific-counts-chicas-poderosas-quibdo-colombia/

Με βάση αυτές τις προσεγγίσεις, συνεργαστήκαμε με το Αναπτυξιακό Πρόγραμμα των Ηνωμένων Εθνών στην Ονδούρα τον Σεπτέμβριο του 2018 για τη δημιουργία ενός εργαστηρίου με τα Δημοτικά Παρατηρητήρια για τη Συνύπαρξη και την Ασφάλεια των Πολιτών («Observatorios de Municipales de Convivencia y Seguridad Ciudadana»). Οι συμμετέχοντες εργάστηκαν με δεδομένα σχετικά με τους βίαιους θανάτους ανδρών και γυναικών και ενδιαφέρθηκαν να παρουσιάσουν δεδομένα κατανεμημένα ανά φύλο. Δύο από τους στόχους ήταν να δημιουργηθούν συναισθηματικά μονοπάτια, για να ξεκινήσουν συζητήσεις με την κοινότητα γύρω από αυτά τα δύσκολα θέματα μέσω συμμετοχικών δραστηριοτήτων και να χρησιμοποιηθούν οι περιορισμένοι πόροι, για να διαμοιραστούν ευαίσθητα και σημαντικά δεδομένα.

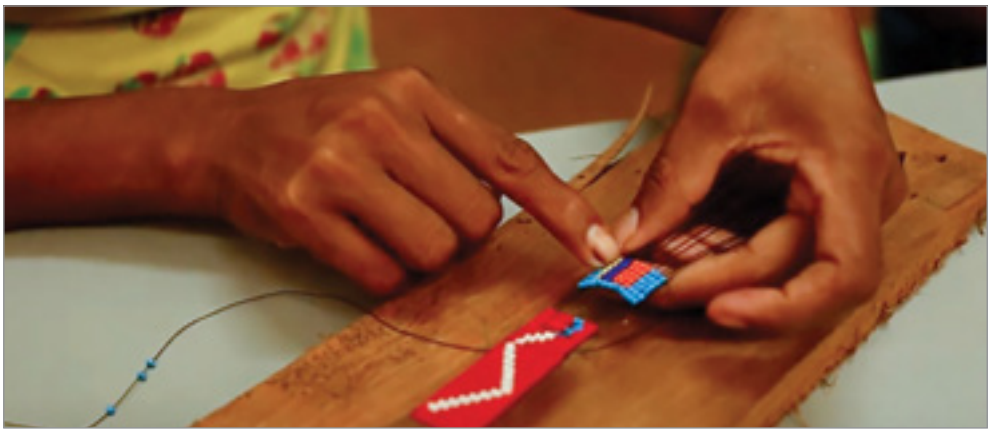
Σε αυτά τα εργαστήρια, τα αρχικά βήματα είναι δραστηριότητες που σπάνε τον πάγο με απλές και αστείες ερωτήσεις (Εικόνα 8.5.1). Στο εργαστήριο της Ονδούρας, η συζήτηση με τους συμμετέχοντες για τη βία παρουσίασε δυσκολίες λόγω των διαφορετικών κοινωνικών προτύπων, καθώς και των γλωσσικών φραγμών. Ως εκ τούτου, εστιάσαμε το εργαστήριο σε διάφορες διερευνητικές δραστηριότητες συλλογής δεδομένων, για να φέρουμε στην επιφάνεια διαφορετικές αντιλήψεις και εμπειρίες σχετικά με τη βία. Χρησιμοποιήσαμε σχέδια, εικόνες και φωτογραφίες, για να δημιουργήσουμε μαζί αφίσες. Οι συμμετέχοντες μπορούσαν να προσθέσουν σε αυτά αυτοκόλλητα ως έναν τρόπο συλλογής δεδομένων –μεταξύ άλλων για τον τρόπο με τον οποίο οραματίζονται τον εαυτό τους, για την κατανόηση των δικαιωμάτων τους και για τα διάφορα είδη ενδοοικογενειακής βίας (π.χ. σωματική, ψυχολογική, οικονομική) που έχουν βιώσει στη ζωή τους.

Αναλογική οπτικοποίηση δεδομένων

Σε μια προσπάθεια να κατανοήσουμε καλύτερα τα ζητήματα που ταλανίζουν τις κοινότητες των ιθαγενών, το 2017 σχεδιάσαμε διαδραστικά εργαστήρια με τη φυλή Εμπερά της περιοχής Βιγία ντελ Φουέρτε. Τα εργαστήρια προσπάθησαν να προσφέρουν μια εικόνα της ζωής τους, παρά τα γλωσσικά εμπόδια. Ιστορικά, οι αλληλεπιδράσεις μεταξύ της φυλής και των ξένων ήταν σε μεγάλο βαθμό ανδροκρατούμενες. Επομένως, δώσαμε προτεραιότητα στην πρόσβαση στους γυναικείους πληθυσμούς, προκειμένου να μετρήσουμε τα επίπεδα εκπαίδευσής τους και να διευκολύνουμε τις συζητήσεις σχετικά με τη χειραφέτηση.

Εικόνα 8.5.1: «Είμαι τοσο δημιουργική;»: δραστηριότητα αναλογικής συλλογής δεδομένων. Πηγή: Eliana Vaca.

Ελλείψει σύγχρονων τεχνολογιών, διερευνήσαμε παραδοσιακές εκφράσεις του πολιτισμού ως μέσο, για να αποκτήσουμε ουσιαστικότερη πρόσβαση στη ζωή των συμμετεχόντων μας. Οι εκφράσεις αυτές περιλάμβαναν παραδοσιακές πρακτικές, όπως η υφαντουργία, η χειροτεχνία με χάντρες και οι κατασκευές (Εικόνα 8.5.2).



Εικόνα 8.5.2: Παράδειγμα αναλογικής οπτικοποίησης με χάντρες, όπου διαφορετικά χρώματα αντιπροσωπεύουν διαφορετικές ομιλούμενες γλώσσες και η ποσότητα των χαντρών αντιπροσωπεύει την ευχέρεια σε καθεμία από αυτές. Πηγή: Eliana Vaca.

Τον Σεπτέμβριο του 2018 στην Ονδούρα, οργανώσαμε ένα εργαστήριο σχετικά με το πώς να «εξανθρωπίσουμε» τα δεδομένα, πραγματοποιώντας έργα ανθεκτικότητας με θύματα και πληθυσμούς σε κίνδυνο. Σχεδιάσαμε εργαστήρια αναλογικής οπτικοποίησης δεδομένων χαμηλού κόστους με τεχνικές σχεδιασμού με γνώμονα την ενσυναίσθηση, χρησιμοποιώντας ψαλίδια, χαρτιά, αυτοκόλλητα, post-it και μπαλόνια. Αυτά εξυπηρετούσαν τη διευκόλυνση της ανταλλαγής ευαίσθητων πληροφοριών με τους αρμόδιους οργανισμούς για την καλύτερη υποστήριξη των συγκεκριμένων κοινοτήτων, καθώς και τη διδασκαλία διαφορετικών μεθόδων που θα μπορούσαν να χρησιμοποιήσουν οι ευάλωτοι και χαμηλού ποσοστού αλφαβητισμού πληθυσμοί, για να μοιραστούν δεδομένα σχετικά με τη ζωή, τις εμπειρίες και τα προβλήματά τους. Για παράδειγμα, συνεργαστήκαμε με τους συμμετέχοντες για τη δημιουργία αναλογικών απεικονίσεων σχετικά με τις δολοφονίες και τις γυναικοκτονίες ανά περιοχή, τύπο και ηλικία.

Σε ένα άλλο εργαστήριο στο Μπελίζ διερευνήσαμε διαφορετικές συνεργατικές προσεγγίσεις για την οπτικοποίηση δεδομένων σχετικά με το έγκλημα και τη βία. Αρχικά, ξεκινήσαμε για να δούμε πώς τα δεδομένα από το Παρατηρητήριο Βίας του Μπελίζ θα μπορούσαν να χρησιμοποιηθούν για τον συντονισμό διαφόρων ειδών συλλογικών αποκρίσεων. Ενώ οι συμμετέχοντες διέθεταν υψηλά επίπεδα αλφαβητισμού, οι τεχνολογικοί πόροι και η συνδεσιμότητα ήταν πολύ πιο επισφαλείς, καθιστώντας δύσκολη τη χρήση βασικών διαδικτυακών εργαλείων οπτικοποίησης. Αυτό προκάλεσε πολλά ερωτήματα και προκλήσεις σχετικά με τις διαδικτυακές πρακτικές οπτικοποίησης δεδομένων, οι οποίες συχνά θεωρούνται αυτονόητες, αλλά δεν θα λειτουργούσαν στα περιβάλλοντα στα οποία βρισκόμασταν –γεγονός που υποδηλώνει, και πάλι, τη σημασία των αναλογικών προσεγγίσεων στην οπτικοποίηση δεδομένων με τη χρήση πιο εύκολα διαθέσιμων υλικών.

Η συγγραφέας

Η Ελιάνα Βάκα Μουνιός είναι ντιζάινερ η οποία ειδικεύεται στην ανάπτυξη έργων που περιλαμβάνουν ανθρώπινη συμπεριφορά και μεθοδολογίες ανθρωπολογίας για τη δημιουργία ανθρωποκεντρικών λύσεων σχεδιασμού.

9.

Χωροθέτηση της Δημοσιογραφίας Δεδομένων

9.1
Γενεαλογίες της Δημοσιογραφίας Δεδομένων

C. W. Anderson

Περίληψη: Αυτό το κεφάλαιο εξετάζει τη δημοσιογραφία δεδομένων από κριτική και ιστορική σκοπιά.

Λέξεις-κλειδιά: γενεαλογία, δημοσιογραφία δεδομένων, Ιστορία, αβεβαιότητα, υπολογιστική δημοσιογραφία, δημοσιογραφία ακριβείας

Εισαγωγή

Τι λόγο έχει κανείς να ενδιαφέρεται για την Ιστορία της δημοσιογραφίας δεδομένων; Η «Ιστορία» όχι μόνο είναι ένα μάλλον ακαδημαϊκό και αφηρημένο θέμα για τους περισσότερους ανθρώπους, αλλά επιπλέον μπορεί να φαντάζει ιδιαίτερα δευτερεύον για τους επαγγελματίες δημοσιογράφους δεδομένων που είναι απασχολημένοι με τη δουλειά τους. Είναι κατανοητό ότι οι δημοσιογράφοι, που εργάζονται υπό πιεστικές προθεσμίες και με στόχο να μεταφέρουν περίπλοκες πληροφορίες γρήγορα και κατανοητά σε όσο το δυνατόν περισσότερους αναγνώστες, δεν θέλουν να σπαταλούν πολύ χρόνο στην αυτοκριτική. Τις περισσότερες φορές, αυτή η απροθυμία για «ομφαλοσκόπηση» είναι μια αξιοθαύμαστη ιδιότητα –όταν, όμως, πρόκειται για τις πρακτικές και τις έννοιες της δημοσιογραφίας δεδομένων και του υπολογιστικού ρεπορτάζ («computational reporting»), η εχθρικότητα απέναντι στην ιστορική σκέψη μπορεί να είναι ένα μειονέκτημα που στέκεται εμπόδιο στην παραγωγή πραγματικά ποιοτικής δημοσιογραφίας.

Η δημοσιογραφία δεδομένων ίσως είναι η πιο ισχυρή μορφή συλλογικής δημοσιογραφικής εκλογίκευσης στον κόσμο σήμερα. Τουλάχιστον, μπορεί να είναι η πιο θετική και θετικιστική μορφή δημοσιογραφίας. Αυτή η *δύναμη* (η ικανότητα της δημοσιογραφίας δεδομένων να παράγει δημοσιογραφία υψηλής ποιότητας, μαζί με τη ρητορική δυναμική του μοντέλου της δημοσιογραφίας δεδομένων), η *θετικότητα* (οι περισσότεροι δημοσιογράφοι δεδομένων τρέφουν μεγάλες ελπίδες για το μέλλον του συγκεκριμένου, επιμέρους τομέα τους, πεπεισμένοι ότι αυτός βρίσκεται σε πορεία ανάπτυξης) και ο *θετικισμός* (οι δημοσιογράφοι δεδομένων πιστεύουν ακράδαντα στην ικανότητα της καθοδηγούμενης από τη μέθοδο έρευνας να καταγράφει πραγματικά και αποδείξιμα γεγονότα για τον κόσμο) δημιουργούν αυτό που θα αποκαλούσα ένα επάγγελμα με εμπειρική αυτοπεποίθηση. Μια συνέπεια αυτής της αυτοπεποίθησης, θα έλεγα, είναι ότι μπορεί, επίσης, να δημιουργήσει μια αξίωση που πηγάζει από την ιδεολογία των Ουίγων, ότι η δημοσιογραφία δεδομένων ολοένα και βελτιώνει τον κόσμο. Μια τέτοια στάση είναι δυνατόν να οδηγήσει σε αλαζονεία και σε έλλειψη κριτικής αυτοαναστοχαστικότητας, κάνοντας, τελικά, τη δημοσιογραφία να προσομοιάζει περισσότερο με τους θεσμούς τους οποίους εκείνη καλεί να λογοδοτήσουν.

Σε αυτό το κεφάλαιο θέλω να υποστηρίξω ότι, δίνοντας μεγαλύτερη προσοχή στην Ιστορία, μπορούμε στην πραγματικότητα να βελτιώσουμε τις καθημερινές εργασίες της δημοσιογραφίας δεδομένων. Κατανοώντας ότι οι διαδικασίες και οι πρακτικές τους κουβαλούν μια ιστορία, οι δημοσιογράφοι δεδομένων μπορούν να συνειδητοποιήσουν ότι σήμερα τα πράγματα θα μπορούσαν να γίνουν με άλλον τρόπο, ακριβώς επειδή κάποτε μπορεί να ήταν διαφορετικά.

Ειδικότερα, οι δημοσιογράφοι δεδομένων θα μπορούσαν να ψάξουν καλύτερα, για να βρουν έναν τρόπο να αναπαραστήσουν δημιουργικά την αβεβαιότητα στο εμπειρικό τους έργο. Θα μπορούσαν να εξετάσουν τεχνικές, με τις οποίες θα προσελκύσουν αναγνώστες διαφορετικής πολιτικής ευαισθησίας και πεποιθήσεων, που πηγαίνουν πέρα από την απλή παράθεση πραγματικών στοιχείων. Θα μπορούσαν, εν ολίγοις, να είναι ανοιχτοί σε αυτό που οι μελετητές και οι ιστορικοί των επιστημονικών και τεχνολογικών σπουδών Κάθριν Ντ' Ιγκνάζιο και Λόρεν Κλάιν έχουν αποκαλέσει «φεμινιστική οπτικοποίηση δεδομένων», η οποία επανεξετάζει τις δυαδικές σχέσεις, αγκαλιάζει τον πλουραλισμό, διερευνά την εξουσία και συνεκτιμά το πλαίσιο (D' Ignazio & Klein, 2020· βλ. επίσης το κεφάλαιο της Ντ' Ιγκνάζιο[231] σε αυτόν τον τόμο). Για να επιτύχει αυτές τις αλλαγές, η δημοσιογραφία δεδομένων θα πρέπει πράγματι να αποτυπώσει αυτήν την ισχυρή ιστορική ευαισθησία, λόγω της ίδιας της φύσης της δύναμης και της αυτοπεποίθησής της, σε μεγαλύτερο βαθμό από τις περισσότερες άλλες μορφές δημοσιογραφικής πρακτικής. Καμία μορφή Ιστορίας δεν είναι καλύτερα εφοδιασμένη για να οδηγήσει στην αυτοαναστοχαστικότητα, θα έλεγα, από τη γενεαλογική προσέγγιση της εννοιολογικής ανάπτυξης που εισήγαγε ο Μισέλ Φουκώ και υιοθέτησαν ορισμένοι ιστορικοί της επιστήμης και μελετητές των επιστημονικών και τεχνολογικών σπουδών.

Η «γενεαλογία», όπως ορίζεται από τον Φουκώ, ο οποίος στηρίζεται στο προγενέστερο έργο του Νίτσε, είναι μια μοναδική προσέγγιση για τη μελέτη της εξέλιξης των θεσμών και των εννοιών με την πάροδο του χρόνου, και μάλιστα είναι μια προσέγγιση που θα μπορούσε να διαφοροποιηθεί από την Ιστορία ως τέτοια. Η γενεαλογική ανάλυση δεν αναζητά μια ενιαία, αδιάσπαστη προέλευση πρακτικών ή ιδεών στο παρελθόν, ούτε προσπαθεί να κατανοήσει τον τρόπο με τον οποίο οι έννοιες αναπτύχθηκαν σε μια αδιάσπαστη και εξελικτική γραμμή από χθες έως σήμερα. Αντιθέτως, εστιάζει περισσότερο στην *ασυνέχεια* και στις *απροσδόκητες αλλαγές* παρά στην παρουσία του παρελθόντος μέσα στο παρόν. Όπως σημειώνει ο Νίτσε, σε ένα απόσπασμα από τη Γενεαλογία της Ηθικής που παραθέτει ο Φουκώ:

> Η «ανάπτυξη» ενός πράγματος, μιας πρακτικής ή ενός οργάνου δεν έχει καμία σχέση με την πρόοδό του προς έναν μοναδικό στόχο, ούτε καν πρόκειται για τη λογική και συντομότερη πρόοδο που επιτυγχάνεται με τη μικρότερη δυνατή δαπάνη δυνάμεων και πόρων. Είναι, μάλλον, η ακολουθία περισσότερο ή λιγότερο βαθυστόχαστων, περισσότερο ή λιγότερο ανεξάρτητων μεταξύ τους διεργασιών υπερίσχυσης που λαμβάνουν χώρα σε αυτήν την περίπτωση, μαζί με την αντίσταση που προκύπτει κάθε φορά ενάντια σε αυτήν την υπερίσχυση, τις αλλαγές της μορφής που έχουν επιχειρηθεί με σκοπό την άμυνα και την αντίδραση, καθώς και τα αποτελέσματα των επιτυχημένων αντισταθμιστικών μέτρων. Η μορφή είναι ρευστή· το «νόημα», ωστόσο, είναι ακόμη πιο ρευστό (Foucault, 1980).

Μια «γενεαλογία της δημοσιογραφίας δεδομένων» («genealogy of data journalism»), λοιπόν, θα αποκάλυπτε πώς η δημοσιογραφία δεδομένων εξελίχθηκε με τρόπους που οι δημιουργοί και οι επαγγελματίες της δεν περίμεναν ποτέ ή με τρόπους που μπορεί ακόμη και να ήταν αντίθετοι με τις επιθυμίες τους. Θα εξέταζε τους τρόπους με τους οποίους η Ιστορία μάς εκπλήσσει και μερικές φορές μας οδηγεί σε απροσδόκητες κατευθύνσεις. Αυτή η προσέγγιση, όπως υποστήριξα προηγουμένως, θα ήταν ιδιαίτερα χρήσιμη για τους σημερινούς δημοσιογράφους δεδομένων. Θα τους βοηθούσε να κατανοήσουν, νομίζω, ότι δεν εργάζονται σε μια προκαθορισμένη παράδοση με ένα σεβαστό παρελθόν· αντίθετα, την επινοούν, κυρίως καθώς προχωρούν, με τρόπους που είναι ριζικά περιστασιακοί. Και αυτό θα οδηγούσε σε μια χρήσιμη μορφή κριτικής αυτοαναστοχαστικότητας, η οποία θα μπορούσε να βοηθήσει να μετριαστεί η –κατανοητή και συχνά καλοδεχούμενη– αυτοπεποίθηση των δημοσιογράφων και των ρεπόρτερ που εργάζονται με δεδομένα.

231. ΣτΕ: Κεφάλαιο 4.1, «Δημοσιογραφία Δεδομένων: Πώς Σχετίζεται ο Φεμινισμός με την Τεχνολογία Πληροφοριών»

Επιχείρησα να γράψω έναν τέτοιο γενεαλογικό απολογισμό στο βιβλίο μου, *Apostles of Certainty: Data Journalism and the Politics of Doubt* (Anderson, 2018). Στις σελίδες που ακολουθούν, θέλω να συνοψίσω μερικά από τα κύρια συμπεράσματα του βιβλίου και να συζητήσω τρόπους με τους οποίους τα διδάγματά του θα μπορούσαν να είναι χρήσιμα σήμερα. Θέλω να καταλήξω με το επιχείρημα ότι η δημοσιογραφία, ιδίως το είδος της δημοσιογραφίας που βασίζεται στα δεδομένα, θα μπορούσε και θα έπρεπε να κάνει καλύτερη δουλειά, για να επισημάνει όλα όσα δεν γνωρίζει –και ότι η επίδειξη αβρότητας απέναντι στην αβεβαιότητα θα τιμούσε τις ρίζες της δημοσιογραφίας δεδομένων που βρίσκονται στην κριτική απέναντι στην άνομη εξουσία, και όχι στην αναπαράστασή της.

Η δημοσιογραφία δεδομένων στο πέρασμα του χρόνου: δεκαετίες 1910, 1960 και 2010

Μπορούν οι δημοσιογράφοι να χρησιμοποιούν τα δεδομένα –μαζί με άλλες μορφές ποσοτικοποιημένων πληροφοριών, όπως έντυπα έγγραφα με αριθμούς, οπτικοποιήσεις δεδομένων, διαγράμματα και γραφήματα, για να παράγουν καλύτερη δημοσιογραφία; Και πώς θα μπορούσε αυτή η δημοσιογραφία να βοηθήσει το κοινό να κάνει καλύτερες πολιτικές επιλογές; Αυτά ήταν τα κύρια ερωτήματα που διέπουν το *Apostles of Certainty: Data Journalism and the Politics of Doubt*, όπου προσπάθησα να εξετάσω την ιστορία των ειδήσεων σε βάθος χρόνου. Με στάσεις στη δεκαετία του 1910, στη δεκαετία του 1960 και στο παρόν, το βιβλίο παρακολουθεί τη γενεαλογία της δημοσιογραφίας δεδομένων και τα υλικά και τεχνολογικά της θεμέλια, και υποστηρίζει ότι η χρήση των δεδομένων στην ειδησεογραφία είναι αναπόφευκτα συνυφασμένη με την εθνική πολιτική, την εξέλιξη των υπολογιστικών βάσεων δεδομένων και την ιστορία των επαγγελματικών επιστημονικών πεδίων. Είναι αδύνατο να κατανοήσουμε τις δημοσιογραφικές χρήσεις των δεδομένων, υποστηρίζω στο βιβλίο, χωρίς να κατανοήσουμε τις συχνά συγκρουσιακές σχέσεις μεταξύ των κοινωνικών επιστημών και της δημοσιογραφίας. Είναι, επίσης, αδύνατο να διαχωρίσουμε τις εμπειρικές μορφές δημόσιας αλήθειας, χωρίς να κατανοήσουμε πρώτα την αξιοσημείωτα επίμονη προοδευτική πεποίθηση ότι η δημοσίευση εμπειρικά επαληθεύσιμων πληροφοριών θα οδηγήσει σε έναν πιο δίκαιο και ευημερούντα κόσμο. Το *Apostles of Certainty* καταλήγει στο συμπέρασμα ότι αυτή η διασταύρωση τεχνολογίας και επαγγελματισμού οδήγησε σε καλύτερη δημοσιογραφία αλλά όχι απαραίτητα σε καλύτερη πολιτική. Για να ανταποκριθεί πλήρως στις απαιτήσεις της ψηφιακής εποχής, η δημοσιογραφία θα πρέπει να εκφράζει την εμπειρική αμφιβολία με την ίδια άνεση που εκφράζει και τη βεβαιότητα. Παραδόξως, αυτός ο «εναγκαλισμός της αμφιβολίας» θα μπορούσε να δημιουργήσει μια δημοσιογραφία που μοιάζει περισσότερο, και όχι λιγότερο, με επιστήμη.

Η πρόκληση των κοινωνικών επιστημών

Η αφήγηση του *Apostles of Certainty* εδράζεται σε τρεις διαφορετικές χρονικές περιόδους των ΗΠΑ, οι οποίες παρέχουν τρεις διαφορετικές οπτικές γωνίες για την ανάπτυξη της δημοσιογραφίας δεδομένων. Η πρώτη είναι η λεγόμενη «Προοδευτική Εποχή» («Progressive Era»), μια περίοδος επικράτησης της φιλελεύθερης πολιτικής που συνοδευόταν από την πεποίθηση ότι τόσο το κράτος όσο και οι απλοί πολίτες, που θα είχαν στη διάθεσή τους τις καλύτερες δυνατές στατιστικές για την ενημέρωσή τους, θα μπορούσαν να κάνουν τον κόσμο πιο δίκαιο και ανθρώπινο. Η δεύτερη στιγμή είναι οι δεκαετίες του 1950 και του 1960, όταν κάποιοι μεταρρυθμιστές της δημοσιογραφίας άρχισαν να στρέφουν την προσοχή τους στις ποσοτικές κοινωνικές επιστήμες, ιδίως στις πολιτικές επιστήμες και στην κοινωνιολογία, ως μια πιθανή πηγή νέων ιδεών και μεθόδων για να γίνει η δημοσιογραφία πιο εμπειρική και πιο αντικειμενική. Στην αναζήτησή τους αυτή, τους βοήθησε ένα νέο σύνολο από όλο και πιο προσιτές βάσεις δεδομένων, και όλο και πιο ισχυρούς υπολογιστές. Η τρίτη στιγμή είναι στις αρχές της δεκαετίας του 2010, όταν η αιχμή της τεχνολογίας που είχε στη διάθεσή της η δημοσιογραφία δεδομέ-

νων συμπληρώθηκε από την «υπολογιστική» («computational») ή «δομημένη» («structured») δημοσιογραφία. Στη σημερινή εποχή των μεγάλων δεδομένων («big data») και της «βαθιάς μηχανικής μάθησης» («deep machine learning»), αυτοί οι δημοσιογράφοι υποστηρίζουν ότι η δημοσιογραφική αντικειμενικότητα εξαρτάται λιγότερο από εξωτερικές αναφορές και προέρχεται μάλλον από τη δομή της ίδιας της βάσης δεδομένων.

Σε καθεμία από αυτές τις περιόδους, η δημοσιογραφία με προσανατολισμό στα δεδομένα *ανταποκρινόταν* αλλά και αυτοπροσδιοριζόταν *εν μέρει σε αντίθεση* με ευρύτερα ρεύματα που λειτουργούσαν γενικότερα στις κοινωνικές επιστήμες. Αυτή η σχέση με ευρύτερα πολιτικά και κοινωνικά ρεύματα βοήθησε στην επιλογή των περιπτώσεων στις οποίες επικεντρώθηκα σε αυτό το κεφάλαιο. Με άλλα λόγια, αναζήτησα σημεία καμπής στην ιστορία της δημοσιογραφίας που θα μπορούσαν να βοηθήσουν να φωτιστούν ευρύτερες κοινωνικές και πολιτικές δομές, πέραν της δημοσιογραφίας. Στην Προοδευτική Εποχή[232], τα παραδοσιακά δελτία ειδήσεων απέρριψαν σε μεγάλο βαθμό την αναδυόμενη εστίαση της κοινωνιολογίας στις κοινωνικές δομές και στις αποπροσωποποιημένες πληροφορίες πλαισίου, προτιμώντας να διατηρήσουν την ατομικιστική εστίασή τους σε ισχυρές προσωπικότητες και σημαντικά γεγονότα. Καθώς η δημοσιογραφία και η κοινωνιολογία εξελίχθηκαν σε επαγγελματικές δραστηριότητες, και οι δύο άρχισαν να προβάλλουν δομικούς ισχυρισμούς με ολοένα και μεγαλύτερη άνεση. Όμως, μόλις τη δεκαετία του 1960 ο Φίλιπ Μάγιερ και οι μεταρρυθμιστές που συσπειρώθηκαν γύρω από τη φιλοσοφία της Δημοσιογραφίας Ακριβείας («Precision Journalism») άρχισαν να προβάλλουν την ποσοτική κοινωνιολογία και τις πολιτικές επιστήμες ως πρότυπα για το επίπεδο ακρίβειας και πλαισίου στο οποίο η δημοσιογραφία θα έπρεπε να αποβλέπει. Στις αρχές του 21ου αιώνα, ένα μοντέλο της δημοσιογραφίας δεδομένων, που σε μεγάλο βαθμό είχε πια κανονικοποιηθεί, άρχισε να κατακλύζεται από αμφιβολίες σχετικά με τη δυνατότητα αναπαραγωγιμότητας και την αιτιώδη συνάφεια, που ταλαιπωρούσαν όλο και περισσότερο τις κοινωνικές επιστήμες· όπως και οι κοινωνικές επιστήμες, άρχισε να πειραματίζεται, για να δει αν τα «μεγάλα δεδομένα» και οι μη αιτιώδεις μορφές συσχετιστικής συμπεριφοράς θα μπορούσαν να προσφέρουν πληροφορίες για την κοινωνική δραστηριότητα.

Ως εκ τούτου, το *Apostles of Certainty* υποστηρίζει εμμέσως ότι οι μορφές δημοσιογραφικής εξειδίκευσης και αυθεντίας δεν κατασκευάζονται ποτέ μεμονωμένα ή εντελώς εσωτερικά εντός του ίδιου του δημοσιογραφικού πεδίου. Η δημοσιογραφία δεδομένων δεν έγινε δημοσιογραφία δεδομένων για αμιγώς επαγγελματικούς δημοσιογραφικούς λόγους, ούτε η διαδικασία αυτή μπορεί να αναλυθεί αποκλειστικά μέσω της ανάλυσης του δημοσιογραφικού λόγου ή του «εσωτερικού μονολόγου». Αντίθετα, το είδος της εξειδίκευσης που τη δεκαετία του 1960 άρχισε να αποκαλείται δημοσιογραφία δεδομένων μπορεί να γίνει κατανοητό μόνο *συσχετιστικά*, εξετάζοντας τον τρόπο με τον οποίο οι δημοσιογράφοι δεδομένων ανταποκρίνονταν και αλληλεπιδρούσαν με τους (πιο έγκυρους και ισχυρούς) αδελφούς τους των κοινωνικών επιστημών. Επιπλέον, η διαδικασία αυτή δεν μπορεί να γίνει κατανοητή μόνο από την άποψη των ενεργειών και των αγώνων των ανθρώπων, είτε μεμονωμένα είτε σε ομάδες. Η εξειδίκευση, σύμφωνα με το μοντέλο που πρότεινα στο βιβλίο μου, *Apostles of Certainty*, είναι ένα δικτυωμένο φαινόμενο, σύμφωνα με το οποίο οι επαγγελματικές ομάδες αγωνίζονται να εδραιώσουν τη δικαιοδοσία τους σε μια μεγάλη ποικιλία διαλεκτικών *και* υλικών τεχνουργημάτων. Η δημοσιογραφία δεδομένων, για να το θέσουμε απλά, θα ήταν αδύνατο να υπάρξει χωρίς τις βάσεις δεδομένων, αλλά μιλάμε για τις βάσεις δεδομένων, όπως διαμεσολαβούνται μέσω μιας συγκεκριμένης επαγγελματικής αντίληψης σχετικά με το τι είναι μια βάση δεδομένων και πώς θα μπορούσε να χρησιμοποιηθεί με τρόπους που να είναι δεόντως δημοσιογραφικοί (για μια γενικότερη προσέγγιση αυτού του επιχειρήματος σχετικά με τη δικτυωμένη φύση της εξειδίκευ-

232. Στις Ηνωμένες Πολιτείες η χρονική περίοδος που είναι γνωστή ως «Προοδευτική Εποχή» διήρκεσε από τη δεκαετία του 1880 έως τη δεκαετία του 1920 και συνήθως θεωρείται μια μεγάλη εποχή φιλελεύθερων μεταρρυθμίσεων και μια προσπάθεια ευθυγράμμισης της δημόσιας πολιτικής με τη βιομηχανική εποχή.

σης, βλ. Anderson, 2013). Είναι αδύνατο να κατανοήσουμε τη δημοσιογραφική αυθεντία χωρίς να κατανοήσουμε επίσης την αυθεντία των κοινωνικών επιστημών (και το ίδιο πράγμα θα μπορούσε να ειπωθεί για την επιστήμη των υπολογιστών, την ανθρωπολογία ή την εκτενή, μη μυθοπλαστική αφήγηση). Ο δημοσιογραφικός επαγγελματισμός και η γνώση δεν μπορούν ποτέ να γίνουν κατανοητά αποκλειστικά με την εξέταση του ίδιου του πεδίου της δημοσιογραφίας.

Η επιμονή της πολιτικής

Η δημοσιογραφία δεδομένων πρέπει να γίνει κατανοητή μέσα από μια γενεαλογική οπτική και σε σχέση με παρακείμενα πεδία εξειδίκευσης, όπως η κοινωνιολογία και οι πολιτικές επιστήμες. Όλα αυτά τα πεδία, με τη σειρά τους, πρέπει να αναλυθούν μέσα από τις ευρύτερες αντιλήψεις τους για την πολιτική και τον τρόπο με τον οποίο συμβιβάζονται με το γεγονός ότι τα «γεγονότα» που αποκαλύπτουν είναι «πολιτικά», είτε τους αρέσει είτε όχι. Πράγματι, ακόμη και η επιθυμία για πραγματική γνώση είναι από μόνη της μια πολιτική πράξη. Σε όλη την Ιστορία της δημοσιογραφίας δεδομένων, όπως υποστηρίζω στο βιβλίο μου, *Apostles of Certainty*, έχουμε γίνει μάρτυρες της ξεχωριστής προσπάθειας που έχει καταβάλει, ώστε να στηριχτεί στην ουδετερότητα των κοινωνικών επιστημών, προκειμένου να εφαρμόσει όσα μπορούν να περιγραφούν μόνο ως προοδευτικοί πολιτικοί στόχοι. Ωστόσο, το ευρύτερο πλαίσιο μέσα στο οποίο διαμορφώνεται αυτή η σύνδεση έχει αλλάξει δραματικά με την πάροδο του χρόνου. Αυτές οι ευρύτερες μεταβολές θα πρέπει να μετριάσουν κάθε ενθουσιώδη πεποίθηση ότι αυτό που παρακολουθούμε στη δημοσιογραφία είναι μια τελεολογική εξέλιξη της δημοσιογραφικής βεβαιότητας, όπως αυτή καθίσταται δυνατή από τις ολοένα και πιο εξελιγμένες ψηφιακές συσκευές.

Στην Προοδευτική Εποχή, οι πρώτοι δημοσιογράφοι δεδομένων είδαν τη συλλογή και τη συσσώρευση ποσοτικών στοιχείων ως μια διαδικασία κοινωνικού και πολιτικού διαφωτισμού, μια διαδικασία που ήταν, ωστόσο, απαλλαγμένη από οποιεσδήποτε ευρύτερες πολιτικές δεσμεύσεις. Συγκεντρώνοντας λεπτομερή στοιχεία για τα επίπεδα υγιεινής της πόλης, την κατανομή της φτώχειας στους αστικούς χώρους, στατιστικά στοιχεία για τον εκκλησιασμό και τη θρησκευτική πρακτική, τις συνθήκες εργασίας και ποικιλία άλλων πραγματολογικών στοιχείων (και μεταδίδοντας αυτά τα στοιχεία στο κοινό μέσω του Τύπου), οι κοινωνικοί ερευνητές πίστευαν ότι η κοινωνία ως οργανισμός θα κατανοούσε με μεγαλύτερο σθένος τις συνθήκες της ύπαρξής της. Αποκτώντας μια καλύτερη κατανόηση του εαυτού της, η κοινωνία θα βελτιωνόταν, τόσο από μόνη της όσο και παρακινώντας τους πολιτικούς να θεσπίσουν μεταρρυθμιστικά μέτρα. Σε αυτήν την περίπτωση, οι πραγματικές γνώσεις για τον κόσμο μιλούσαν από μόνες τους· έπρεπε απλώς να συγκεντρωθούν, να οπτικοποιηθούν και να δημοσιοποιηθούν, και η φώτιση θα ακολουθούσε. Θα μπορούσαμε να πούμε ότι αυτή είναι μια «αφελής και διαφανής» αντίληψη για τα γεγονότα –δεν απαιτούν καμία ερμηνεία από μόνα τους και η συγκέντρωσή τους θα οδηγήσει σε θετική κοινωνική αλλαγή. Η δημοσιογραφία δεδομένων, σε αυτή τη φάση, θα μπορούσε να είναι πολιτική, χωρίς να δηλώνει ρητά την πολιτική της τοποθέτηση.

Μέχρι την εποχή του Φίλιπ Μάγιερ και τη δεκαετία του 1960, αυτή η εύκολη ταύτιση μεταξύ διαφανών γεγονότων και πολιτικής είχε καταρρεύσει. Η δημοσιογραφία ήταν ατελής, υποστήριζε ο Μάγιερ και οι οπαδοί του καθόλη τη διάρκεια των δεκαετιών του 1950 και του 1960, επειδή θεωρούσε λανθασμένα ότι η αντικειμενικότητα ήταν απλώς η συλλογή μιας σειράς από καταγραφές των όσων πίστευε η κάθε πλευρά ενός πολιτικού ζητήματος ότι ήταν η αλήθεια, που επέτρεπε στον αναγνώστη να πάρει τις δικές του αποφάσεις σχετικά με το τι πράγματι ήταν η αλήθεια. Σε μια εποχή κοινωνικής αναταραχής και πολιτικής τρικυμίας, η δημοσιογραφική αντικειμενικότητα έπρεπε να αποκτήσει ένα πιο σταθερό βάδισμα και φάνηκε πως μπορούσε να το βρει πατώντας γερά στο έδαφος της αντικειμενικής κοινωνικής επιστήμης. Το σημείο εκκίνησης για το δημοσιογραφικό ρεπορτάζ σχετικά με ένα θέμα δεν θα πρέπει να είναι οι διαλεκτικοί ισχυρισμοί των ιδιοτελών πολιτικών, αλλά μάλλον η ψυχρή, σκληρή αλήθεια που προκύπτει από την ανάλυση των σχετικών δεδομένων με την εφαρμογή μιας κατάλληλης με-

θόδου. Μια τέτοια ανάλυση θα ήταν *επαγγελματική αλλά όχι πολιτική*. Λειτουργώντας ως ένα άκρως επαγγελματικό επιτελείο αφηγητών της αλήθειας, οι δημοσιογράφοι θα μπορούσαν να παρακάμψουν την πολιτική προπαγάνδα και να βοηθήσουν το κοινό να πατήσει στο στέρεο έδαφος της αντικειμενικής αλήθειας. Από την άλλη πλευρά, το μέρος όπου θα οδηγούσε αυτή η αλήθεια, δεν είχε καμία σημασία. Σε αντίθεση με ό,τι συνέβαινε στην προηγούμενη γενιά των μακάριων και αφελώς προοδευτικών δημοσιογράφων δεδομένων, η διαφωτιστική επίδραση των δεδομένων δεν είναι αυτονόητη.

Σήμερα, θα έλεγα ότι μια νέα γενιά υπολογιστικών δημοσιογράφων έχει ασυναίσθητα υιοθετήσει εκ νέου κάποιες από τις πολιτικές και επιστημολογικές πεποιθήσεις των προγόνων της, της Προοδευτικής Εποχής. Από επιστημολογική άποψη, υπάρχει μια αυξανόμενη πεποίθηση μεταξύ των υπολογιστικών δημοσιογράφων ότι τα ψηφιακά γεγονότα κατά κάποιο τρόπο «μιλούν από μόνα τους» –ή, τουλάχιστον, ότι αυτά τα γεγονότα θα το κάνουν, όταν έχουν συλλεχθεί, ταξινομηθεί και καθαριστεί σωστά. Σε μεγάλη κλίμακα, και όταν συνδέονται με μεγαλύτερες και εσωτερικά συνεκτικές εννοιολογικές βάσεις δεδομένων, τα γεγονότα δημιουργούν ένα είδος *συσχετιστικής υπερβολής*, κατά την οποία τα προβλήματα νοήματος ή αιτιότητας παρασύρονται από μια πλημμύρα υπολογιστικών δεδομένων. Επαγγελματικά, οι δημοσιογράφοι δεδομένων αντιλαμβάνονται όλο και περισσότερο την αντικειμενικότητα ως αναδυόμενη μέσα από τη δομή της ίδιας της βάσης δεδομένων –και όχι ως μέρος μιας ευρύτερης επαγγελματικής ερμηνευτικής διαδικασίας. Τέλος, πολιτικά, θα έλεγα ότι έχει επιστρέψει ένα είδος «κρυπτοπροοδευτισμού» μεταξύ πολλών από τους πιο επιμελώς ουδέτερους δημοσιογράφους δεδομένων, με μια βαθιά ριζωμένη πολιτική ελπίδα ότι τα όλο και περισσότερα δεδομένα, που οπτικοποιούνται όμορφα και αναμεταδίδονται μέσω ενός ισχυρού Τύπου, μπορούν να λειτουργήσουν ως ανάχωμα απέναντι στις πιο παράλογες ή παθολογικές πολιτικές τάσεις που εκδηλώνονται ολοένα και περισσότερο στις δυτικές δημοκρατίες. Τουλάχιστον, αυτή η ελπίδα επικρατούσε πριν από το 2016 και το διπλό σοκ του Brexit και του Ντόναλντ Τραμπ.

Βεβαιότητα και αμφιβολία

Η ανάπτυξη της δημοσιογραφίας δεδομένων στις Ηνωμένες Πολιτείες κατά τη διάρκεια του μεγάλου 20ού αιώνα θα πρέπει να θεωρηθεί ως μια εξέλιξη κατά την οποία οι όλο και πιο ακριβείς αξιώσεις για επαγγελματική βεβαιότητα των δημοσιογράφων συνυπήρξαν αμήχανα με την αυγή της συνειδητοποίησης ότι όλα τα γεγονότα, ανεξάρτητα από την προέλευσή τους, ήταν λερωμένα από την πολιτική λάσπη. Αυτές οι συχνά αντιφατικές πεποιθήσεις, φυσικά, είναι εμφανείς σε διάφορους τομείς προσανατολισμένους στα δεδομένα, όχι μόνο στη δημοσιογραφία. Σε ένα άρθρο του 2017 στο *The Atlantic*, για παράδειγμα, ο επιστημονικός συντάκτης Εντ Γιονγκ ασχολήθηκε με τον τρόπο με τον οποίο το κίνημα προς την «ανοικτή επιστήμη» και η αυξανόμενη κρίση αναπαραγωγιμότητας θα μπορούσαν να χρησιμοποιηθούν από ένα αντιεπιστημονικό Κογκρέσο, για να υποβαθμίσουν την επιστημονική έρευνα και να της περιορίσουν τη χρηματοδότηση. Ο Γιονγκ παραθέτει τα λόγια της Κρίστι Άσβαντεν, επιστημονικής ρεπόρτερ στο *FiveThirtyEight*: «Φαίνεται ότι υπάρχουν δύο αντίθετα πράγματα που πιστεύει το κοινό για την επιστήμη», λέει στον Γιονγκ.

> [Είτε] ότι πρόκειται για ένα μαγικό ραβδί που μετατρέπει ό,τι αγγίζει σε αλήθεια, είτε ότι όλα είναι βλακείες, επειδή αυτό που πιστεύαμε έχει αλλάξει... Η αλήθεια βρίσκεται κάπου στη μέση. Η επιστήμη είναι μια διαδικασία μείωσης της αβεβαιότητας. Εάν δεν καταδειχθεί το γεγονός ότι η αβεβαιότητα είναι μέρος της διαδικασίας, επιτρέπεται στους αμφισβητίες να πάρουν τη γνήσια αβεβαιότητα και να τη χρησιμοποιήσουν για να υπονομεύσουν τις καταστάσεις (Yong, 2017).

Οι σκέψεις αυτές ευθυγραμμίζονται με το έργο της ερευνήτριας του επιστημονικού πεδίου Επιστήμης, Τεχνολογίας και Κοινωνίας, Χέλγκα Νοβότνι (2016), η οποία υποστηρίζει στο βιβλίο της, *The Cunning of Uncertainty*, ότι «η αλληλεπίδραση μεταξύ της υπέρβασης της αβεβαιότη-

τας και της αναζήτησης της βεβαιότητας ενισχύει την επιθυμία για γνώση». Η ουσία της σύγχρονης επιστήμης –τουλάχιστον στην ιδανική της μορφή– δεν είναι η επίτευξη της βεβαιότητας, αλλά μάλλον το γεγονός ότι δηλώνει τόσο ανοιχτά την προσωρινότητα της γνώσης της. Τίποτα στην επιστήμη δεν είναι γραμμένο στην πέτρα. Η επιστήμη συχνά παραδέχεται ότι αυτά που γνωρίζει είναι λίγα. Μέσα από αυτό, το πιο σύγχρονο από τα παράδοξα, οι αξιώσεις της για τη γνώση καθίστανται άξιες δημόσιας εμπιστοσύνης.

Μία από τις διαπιστώσεις που προσφέρει αυτή η γενεαλογική θεώρηση της εξέλιξης και της ανάπτυξης της δημοσιογραφίας δεδομένων, θα έλεγα, είναι ότι οι δημοσιογράφοι που προσανατολίζονται στα δεδομένα έχουν αποκτήσει εμμονή με την αυξανόμενη ακρίβεια και βεβαιότητα εις βάρος μιας ταπεινότερης κατανόησης της προσωρινότητας και της αμφιβολίας. Όπως προσπάθησα να καταδείξω, από τα μέσα του 20ού αιώνα οι δημοσιογράφοι έχουν επιδοθεί σε μια ολοένα και πιο επιτυχημένη προσπάθεια να καταστήσουν τους ισχυρισμούς τους περί γνώσης πιο ασφαλείς, συναφείς και επεξηγηματικούς. Αυτό είναι κάτι που έκαναν σε μεγάλο βαθμό με τη χρήση διαφορετικών μορφών αποδεικτικών στοιχείων, ιδίως ποσοτικών στοιχείων. Ωστόσο, θα πρέπει να γίνει σαφές ότι αυτός ο αυξημένος επαγγελματισμός –και η αυξανόμενη αυτοπεποίθηση των δημοσιογράφων ότι είναι ικανοί να προβάλλουν ισχυρισμούς αλήθειας με βάση τα συμφραζόμενα– δεν είχε πάντα τα δημοκρατικά αποτελέσματα που προσδοκούν οι δημοσιογράφοι. Ο σύγχρονος αμερικανικός πολιτικός λόγος προσπάθησε να αντιμετωπίσει την αβεβαιότητα της νεωτερικότητας με μια σειρά από ολοένα και πιο έντονες αξιώσεις βεβαιότητας. Η επαγγελματική δημοσιογραφία δεν έχει λύσει αυτό το δίλημμα· μάλλον το έχει οξύνει. Για να αντιμετωπίσει καλύτερα την πολυπλοκότητα του σύγχρονου κόσμου, καταλήγω, η δημοσιογραφία οφείλει να επανεξετάσει τα μέσα και τους μηχανισμούς με τους οποίους μεταδίδει την προσωρινότητα και την αβεβαιότητά της. Εάν αυτό γίνει σωστά, θα μπορούσε να δημιουργήσει μια δημοσιογραφία που μοιάζει περισσότερο, και όχι λιγότερο, με επιστήμη.

Παραπομπές

Anderson, C. W. (2013) «Towards a sociology of computational and algorithmic journalism», *New Media & Society,* 15(7), σσ. 1005–1021. DOI: https://doi.org/10.1177/1461444812465137

Anderson, C. W. (2018). *Apostles of certainty: Data journalism and the politics of doubt.* Νέα Υόρκη: Oxford University Press.

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. Κέιμπριτζ, Μασαχουσέτη: MIT Press.

Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings,* 1972–1977. Νέα Υόρκη: Vintage.

Nowotny, H. (2016). *The cunning of uncertainty.* Κέιμπριτζ, Ηνωμένο Βασίλειο: Polity Press.

Yong, E. (2017) «How the GOP could use science's reform movement against it», *The Atlantic*, 5 Απριλίου. Διαθέσιμο σε: https://www.theatlantic.com/science/archive/2017/04/reproducibility-science-open-judoflip/521952/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Ο συγγραφέας

Ο Κρις Άντερσον είναι καθηγητής Μέσων και Επικοινωνίας στο Πανεπιστήμιο του Leeds και συγγραφέας των βιβλίων *Rebuilding the News: Metropolitan Journalism in the Digital Age* (Temple University Press, 2013) και *Apostles of Certainty: Data Journalism and the Politics of Doubt* (Oxford University Press, 2018).

9.2

Χρυσά Πρότυπα με Γνώμονα τα Δεδομένα: Τι Αποτιμά ο Κλάδος ως Δημοσιογραφία Δεδομένων που αξίζει Βραβείο & Πώς η Δημοσιογραφία Δεδομένων Εξελίσσεται μαζί με τη Δεδομενοποίηση της Κοινωνίας

Wiebke Loosen

Περίληψη: Αυτό το κεφάλαιο μελετά τη σχέση ανάμεσα στη δεδομενοποίηση («datafication») της κοινωνίας και της δημοσιογραφίας και παρουσιάζει τα βραβεία ως μέσο για τη μελέτη της εξέλιξης της δημοσιογραφίας δεδομένων.

Λέξεις-κλειδιά: Data Journalism Awards, δεδομενοποίηση, δεδομενοποιημένη δημοσιογραφία, κοινωνία των δεδομένων, δημοσιογραφική έρευνα, συνδημιουργία

Εισαγωγή: Η απόκριση της δημοσιογραφίας στη δεδομενοποίηση της κοινωνίας

Ίσως μπορούμε σήμερα να κατανοήσουμε καλύτερα από ό,τι στις πρώτες ημέρες της δημοσιογραφίας δεδομένων την εμφάνιση αυτού του νέου ρεπορταζιακού στυλ, ως μια δημοσιογραφική απόκριση στη δεδομενοποίηση («datafication») της κοινωνίας (Loosen, 2018). Ο όρος αναφέρεται στη συνεχώς αυξανόμενη διαθεσιμότητα δεδομένων που έχει τις ρίζες της στην ψηφιοποίηση του περιβάλλοντος (των μέσων) στο οποίο ζούμε και στα ψηφιακά ίχνη και τα μεγάλα δεδομένα («big data») που προκύπτουν από τη διαβίωση σε ένα τέτοιο περιβάλλον (Dijck, 2014). Η διαδικασία αυτή μετατρέπει πολλές πτυχές της κοινωνικής μας ζωής σε δεδομένα ηλεκτρονικού υπολογιστή –δεδομένα που, για διάφορους σκοπούς, συγκεντρώνονται και υφίστανται αλγοριθμική επεξεργασία. Η δεδομενοποίηση οδηγεί σε ποικίλες συνέπειες και εκδηλώνεται με διαφορετικούς τρόπους στην πολιτική, για παράδειγμα, από ό,τι στον οικονομικό κόσμο ή στον τομέα της εκπαίδευσης. Ωστόσο, αυτό που έχουν κοινό όλα τα κοινωνικά πεδία είναι ότι μπορούμε να υποθέσουμε πως, από εδώ και στο εξής, θα βασίζονται ολοένα και περισσότερο σε ένα όλο και πιο ποικίλο φάσμα και σε μεγαλύτερο όγκο δεδομένων για τις διαδικασίες (αυτο)εκλογίκευσής τους.

Η υπαγωγή της δεδομενοποίησης της δημοσιογραφίας στο πλαίσιο της δεδομενοποίησης της ευρύτερης κοινωνίας μάς βοηθά να δούμε πέρα από τη δημοσιογραφία δεδομένων, να την αναγνωρίσουμε ως «μόνο» μία περίπτωση δεδομενοποίησης στη δημοσιογραφία και να κατανοήσουμε καλύτερα τον μετασχηματισμό της δημοσιογραφίας προς μια όλο και περισσότερο βασισμένη σε δεδομένα, αλγοριθμική, μετρικοποιημένη –ή ακόμη και αυτοματοποιημένη– πρακτική (Loosen, 2018). Συγκεκριμένα, αυτό περιλαμβάνει τα αντικείμενα και τα θέματα που υποτίθεται ότι καλύπτει η δημοσιογραφία ή, διαφορετικά, τη λειτουργία της δημοσιογραφίας ως παρατηρητή της κοινωνίας. Όσο περισσότερο τα πεδία και οι κοινωνικοί τομείς που υποτίθεται ότι καλύπτει η δημοσιογραφία δεδομενοποιούνται οι ίδιοι, τόσο περισσότερο η ίδια η δημοσιογραφία πρέπει να μπορεί να κατανοήσει και να παράγει δεδομένα, για να εκπληρώσει τον κοινωνικό της ρόλο. Είναι αυτή η σχέση που αντικατοπτρίζεται στη σύγχρονη δημοσιογραφία

δεδομένων, η οποία βασίζεται ακριβώς σε αυτήν την αυξημένη διαθεσιμότητα δεδομένων, για να διευρύνει το ρεπερτόριο των πηγών για τη δημοσιογραφική έρευνα και για τον εντοπισμό και την αφήγηση ιστοριών.

Βραβεία: Ένα μέσο για να μελετηθεί αυτό που ορίζεται και αποτιμάται ως δημοσιογραφία δεδομένων

Ένας τρόπος παρακολούθησης της εξέλιξης της δημοσιογραφίας δεδομένων ως ρεπορταζιακό στυλ είναι να εξετάσουμε την παραγωγή της. Ενώ οι πρώτες μελέτες στον τομέα της δημοσιογραφικής έρευνας έτειναν να επικεντρώνονται περισσότερο στους φορείς που εμπλέκονται στην παραγωγή της και βασίζονταν κυρίως σε συνεντεύξεις, όλο και περισσότερες μελέτες χρησιμοποιούν τελευταία την ανάλυση περιεχομένου για την καλύτερη κατανόηση της δημοσιογραφίας δεδομένων με βάση τα προϊόντα της (Ausserhofer κ.α., 2020). Τα δημοσιογραφικά βραβεία αποτελούν ένα καλό εμπειρικό σημείο πρόσβασης για τον σκοπό αυτό για διάφορους λόγους. Πρώτον, οι αιτήσεις απονομής βραβείων έχουν ήδη αποδειχθεί χρήσιμες για την ανάλυση των ειδών και των πτυχών της αφήγησης (Wahl-Jorgensen, 2013). Δεύτερον, η δημοσιογραφία δεδομένων είναι ένα διάχυτο αντικείμενο μελέτης που καθιστά όχι μόνο δύσκολο αλλά, μάλλον, προαπαιτούμενο τον εντοπισμό αντίστοιχων κομματιών για μια ανάλυση περιεχομένου. Η δειγματοληψία υποψηφιοτήτων για βράβευση, με τη σειρά της, μας βοηθά ώστε να μη χρειαστεί να εκκινήσουμε είτε από έναν πολύ ειδικό είτε από έναν πολύ ευρύ ορισμό –αυτή η στρατηγική είναι ουσιαστικά ένα μέσο παρατήρησης της αυτοπαρατήρησης στη δημοσιογραφία, καθώς τα θέματα αυτά αντιπροσωπεύουν εκείνο που ο ίδιος ο τομέας θεωρεί ως δημοσιογραφία δεδομένων και πιστεύει ότι αποτελούν αξιόλογα παραδείγματα αυτού του ρεπορταζιακού στυλ. Τρίτον, οι υποψηφιότητες για βραβεία με διεθνή προσανατολισμό είναι πιθανό να επηρεάσουν την ανάπτυξη του τομέα στο σύνολό του, καθώς αυτά είναι ιδιαίτερα αναγνωρισμένα, θεωρούνται ένα είδος χρυσού προτύπου και, ως εκ τούτου, έχουν επίσης διασυνοριακό αντίκτυπο. Επιπλέον, η εξέταση των διεθνών βραβείων μάς επιτρέπει να ερευνήσουμε ένα δείγμα που καλύπτει ένα ευρύ γεωγραφικό και χρονικό φάσμα.

Ωστόσο, είναι επίσης σημαντικό να έχουμε κατά νου ότι η μελέτη των (δημοσιογραφικών) βραβείων συνεπάγεται διαφορετικές προκαταλήψεις. Η μελέτη στην οποία στηριζόμαστε εδώ βασίζεται σε ανάλυση 225 υποψηφιοτήτων (εκ των οποίων 39 βραβευμένες) για τα Data Journalism Awards (DJA) –βραβεία που απονέμονταν ετησίως από το Global Editors Network– κατά τα έτη 2013 έως 2016 (Loosen κ.α., 2020). Αυτό σημαίνει ότι το δείγμα μας υπόκειται σε διπλή μεροληψία επιλογής: αρχικά, είναι αυτοεπιλεκτικό, αφού οι δημοσιογράφοι πρέπει να υποβάλουν οι ίδιοι τις δημοσιεύσεις τους, για να είναι υποψήφιοι. Σε δεύτερη φάση, μια κατά το μάλλον ή ήττον ετησίως μεταβαλλόμενη κριτική επιτροπή από ειδήμονες θα αποφασίσει ποιες συμμετοχές θα είναι πραγματικά υποψήφιες. Επιπλέον, τα βραβεία και οι διακρίσεις αντιπροσωπεύουν μια ιδιαίτερη μορφή «πολιτιστικού κεφαλαίου», γι' αυτό και τα βραβευμένα έργα μπορούν να έχουν έναν συγκεκριμένο αντίκτυπο για το πεδίο στο σύνολό του και να χρησιμεύσουν ως πρότυπο για τα επόμενα έργα (English, 2002). Αυτό σημαίνει, επίσης, ότι τα βραβεία όχι μόνο αντιπροσωπεύουν τον τομέα (σύμφωνα με ορισμένα πρότυπα), αλλά και τον συγκροτούν. Δηλαδή, στην περίπτωσή μας, χαρακτηρίζοντας το περιεχόμενο ως δημοσιογραφία δεδομένων, τα βραβεία διαδραματίζουν έναν ρόλο στη συγκέντρωση διαφορετικών πρακτικών, φορέων, συμβάσεων και αξιών. Μπορεί, λοιπόν, να θεωρηθεί ότι έχουν όχι μόνο μια λειτουργία απονομής βραβείων αλλά και μια λειτουργία δημιουργίας πεδίου. Αυτό σημαίνει ότι τα έργα που αξίζουν τη βράβευση είναι πάντα το αποτέλεσμα ενός είδους «συν-κατασκευής» των υποψηφίων, των κριτών και των αμοιβαία διαμορφωμένων προσδοκιών τους. Τέτοιες επιδράσεις είναι πιθανό να έχουν ιδιαίτερη επιρροή στην περίπτωση της δημοσιογραφίας δεδομένων, καθώς πρόκειται ακόμη για ένα σχετικά νέο ρεπορταζιακό στυλ με το οποίο πειραματίζονται λίγο-πολύ όλοι οι παράγοντες του χώρου.

Εξελίσσεται, αλλά δεν επαναστατεί: Κάποιες τάσεις στη δημοσιογραφία δεδομένων (που αξίζει να βραβευτεί)

Μελέτες που αναλύουν θέματα με γνώμονα τα δεδομένα καταδεικνύουν γενικά ότι η εξέλιξη της δημοσιογραφίας δεδομένων δεν αποτελεί, σε καμία περίπτωση, επανάσταση στην ειδησεογραφία. Ως αποτέλεσμα, αμφισβητούν την ευρέως διαδεδομένη πεποίθηση ότι η δημοσιογραφία με γνώμονα τα δεδομένα φέρνει μια επανάσταση στη δημοσιογραφία, αντικαθιστώντας τις παραδοσιακές μεθόδους ανακάλυψης και κάλυψης των ειδήσεων. Η δική μας μελέτη συμφωνεί, σε γενικές γραμμές, με αυτό που έχουν διαπιστώσει άλλες εμπειρικές αναλύσεις «καθημερινών» δειγμάτων δημοσιογραφίας δεδομένων (Loosen κ.α., 2020). Οι τελευταίες αντιπροσωπεύουν μόνο αρκετά περιορισμένες συλλογές δεδομένων, αλλά μας επιτρέπουν τουλάχιστον να ανιχνεύσουμε κάποιες εξελίξεις και ίσως, πάνω απ' όλα, κάποιο βαθμό συνέπειας στην παραγωγή της δημοσιογραφίας δεδομένων.

Όσον αφορά το ποιος παράγει δημοσιογραφία με γνώμονα τα δεδομένα σε επίπεδο άξιο για βράβευση, τα αποτελέσματα δείχνουν ότι το «χρυσό πρότυπο» για τη δημοσιογραφία δεδομένων, δηλαδή εκείνη που αξίζει την αναγνώριση από ομότιμους, κυριαρχείται από τις εφημερίδες και τα διαδικτυακά τους τμήματα. Από τα τέσσερα έτη που αναλύσαμε, αυτά αντιπροσωπεύουν μακράν τη μεγαλύτερη ομάδα μεταξύ όλων των υποψηφίων καθώς και μεταξύ των βραβευθέντων (43,1% επί του συνόλου· 37,8% επί των βραβευθέντων με DJA). Η μόνη άλλη εξέχουσα ομάδα περιλαμβάνει οργανισμούς που ασχολούνται με την ερευνητική δημοσιογραφία, όπως η *ProPublica* και η Διεθνής Σύμπραξη Ερευνητών Δημοσιογράφων («International Consortium of Investigative Journalists», ICIJ), που βραβεύονται με σημαντικά μεγαλύτερη συχνότητα. Αυτό μπορεί να αντανακλά την εγγενή μεροληψία των βραβείων προς καθιερωμένους φορείς υψηλού προφίλ, απηχώντας τα ευρήματα άλλων ερευνών, σύμφωνα με τα οποία η δημοσιογραφία δεδομένων πάνω από ένα ορισμένο επίπεδο φαίνεται να είναι εγχείρημα για μεγαλύτερους οργανισμούς που διαθέτουν τους πόρους και την εκδοτική προσήλωση, για να επενδύσουν σε διατομεακές ομάδες οι οποίες αποτελούνται από συντάκτες, προγραμματιστές και γραφίστες (Borges-Rey, 2020· Young κ.α., 2018). Αυτό αντικατοπτρίζεται και στα μεγέθη των ομάδων. Από τα 192 έργα του δείγματός μας, όσα ήταν ενυπόγραφα ανέφεραν, κατά μέσο όρο, περισσότερα από πέντε άτομα ως συγγραφείς ή συνεργάτες και περίπου το ένα τρίτο των έργων ολοκληρώθηκε σε συνεργασία με εξωτερικούς συνεργάτες οι οποίοι είτε συνέβαλαν στην ανάλυση είτε σχεδίασαν οπτικοποιήσεις. Αυτό φαίνεται να ισχύει ιδιαίτερα για τα βραβευμένα έργα, τα οποία, σύμφωνα με την ανάλυσή μας, εκπονήθηκαν από μεγαλύτερες ομάδες (μέση τιμή = 6,31, τυπική απόκλιση = 4,7) σε σχέση με εκείνα που ήταν απλώς υποψήφια (μέση τιμή = 4,75, τυπική απόκλιση = 3,8).

Όσον αφορά στις γεωγραφικές περιοχές της δημοσιογραφίας δεδομένων που έλαβαν αναγνώριση σε αυτόν τον διαγωνισμό, βλέπουμε ότι οι Ηνωμένες Πολιτείες κυριαρχούν: σχεδόν οι μισοί από τους υποψηφίους προέρχονται από τις Ηνωμένες Πολιτείες (47,6%), και ακολουθούν, με διαφορά, η Μεγάλη Βρετανία (12,9%) και η Γερμανία (6,2%). Ωστόσο, η δημοσιογραφία δεδομένων εμφανίζεται ως ένα ολοένα και πιο παγκόσμιο φαινόμενο, καθώς ο αριθμός των χωρών που εκπροσωπούνται από τους υποψηφίους αυξανόταν κάθε χρόνο, φτάνοντας το 2016 σε 33 χώρες και από τις πέντε ηπείρους.

Η εξάρτηση της δημοσιογραφίας δεδομένων από συγκεκριμένες πηγές επηρεάζει τα θέματα που αυτή μπορεί ή δεν μπορεί να καλύψει. Κατά συνέπεια, είναι δυνατό η δημοσιογραφία δεδομένων να παραμελήσει τους κοινωνικούς τομείς για τους οποίους δεν παράγονται τακτικά ή δεν είναι προσβάσιμα τα δεδομένα. Όσον αφορά στα θέματα που καλύπτονται, οι υποψηφιότητες για τα DJA χαρακτηρίζονται από μια σταθερή εστίαση σε πολιτικά, κοινωνικά και οικονομικά ζητήματα, με σχεδόν τα μισά από τα έργα που αναλύθηκαν (48,2%) να καλύπτουν κάποιο πολιτικό θέμα. Το μικρό ποσοστό των ιστοριών για την εκπαίδευση, τον πολιτισμό και τον αθλητισμό –σε συμφωνία και με άλλες μελέτες– μπορεί να μην είναι αντιπροσωπευτικό της δημοσιογραφίας δεδομένων εν γένει και, αντίθετα, να προκύπτει από μια μεροληψία υπέρ των «σοβαρών» θεμάτων που είναι

εγγενής στα βραβεία του κλάδου. Ωστόσο, αυτό μπορεί, επίσης, να αντικατοπτρίζει τη διαθεσιμότητα ή τη μη διαθεσιμότητα πηγών δεδομένων για διαφορετικούς τομείς και θέματα ή, στην περίπτωση του δείγματός μας, τη μεροληψία της αυτοεπιλογής των υποψηφίων, η οποία βασίζεται σε αυτό που οι ίδιοι θεωρούν άξιο υποβολής και σε αυτό που περιμένουν να εκτιμήσουν οι κριτές. Προκειμένου να αποκτηθούν πιο αξιόπιστες γνώσεις σχετικά με αυτό το σημείο ζωτικής σημασίας, θα χρειαζόταν μια διεθνής συγκριτική μελέτη που να συσχετίζει τη διαθεσιμότητα και την προσβασιμότητα των δεδομένων με τα θέματα που καλύπτονται από το ρεπορτάζ δεδομένων στις διάφορες χώρες. Μια τέτοια μελέτη εξακολουθεί να απουσιάζει από τη βιβλιογραφία, αλλά θα μπορούσε να μας διαφωτίσει ως προς τους κοινωνικούς τομείς και τα θέματα που καλύπτονται, τις μεθόδους ανάλυσης και τις πηγές των δεδομένων. Μια τέτοια προσέγγιση θα παρείχε, επίσης, πολύτιμες πληροφορίες για την άλλη πλευρά του νομίσματος: τα τυφλά σημεία στην κάλυψη με γνώμονα τα δεδομένα λόγω της έλλειψης (διαθέσιμων) πηγών δεδομένων.

Μια επαναλαμβανόμενη διαπίστωση στην έρευνα σχετικά με το περιεχόμενο της δημοσιογραφίας δεδομένων είναι ότι παρουσιάζει «εξάρτηση από προεπεξεργασμένα δημόσια δεδομένα» που προέρχονται από στατιστικές υπηρεσίες και άλλους κυβερνητικούς φορείς (Borges-Rey, 2020· Tabary κ.α., 2016· Young κ.α., 2018). Είναι κάτι που ισχύει και για θέματα τα οποία βασίζονται σε δεδομένα και είναι επιπέδου που αξίζει βραβείο: παρατηρήσαμε εξάρτηση από δεδομένα των επίσημων θεσμών (σχεδόν το 70% των πηγών δεδομένων) και άλλων μη εμπορικών οργανισμών, όπως ερευνητικά ινστιτούτα και ΜΚΟ, καθώς και από δεδομένα που είναι διαθέσιμα στο κοινό, τουλάχιστον κατόπιν αιτήματος (σχεδόν το 45%). Αυτό καταδεικνύει, αφενός, ότι η δημοσιογραφία δεδομένων αξιοποιεί την αυξημένη διαθεσιμότητα των πηγών δεδομένων, αλλά, αφετέρου, ότι επίσης βασίζεται σε μεγάλο βαθμό σε αυτή τη διαθεσιμότητα: το ποσοστό των δεδομένων που έχουν συλλεχθεί ιδιοχείρως, με scraping ή που έχουν διαρρεύσει ή έχουν αποκτηθεί ως αποτέλεσμα αίτησης είναι σημαντικά μικρότερο. Από την άλλη, η δημοσιογραφία δεδομένων συνδέεται συνεχώς με το ερευνητικό ρεπορτάζ, γεγονός που έχει «οδηγήσει σε μια αντίληψη ότι η δημοσιογραφία δεδομένων έχει να κάνει με τεράστια σύνολα δεδομένων, τα οποία αποκτώνται μέσω πράξεων δημοσιογραφικής γενναιότητας και παράτολμης δράσης» (Knight, 2015· Parasie, 2015· Royal & Blasingame, 2015). Πρόσφατες περιπτώσεις, όπως τα *Panama Papers*, συνέβαλαν σε αυτή την αντίληψη[233]. Ωστόσο, αυτό που δείχνει, επίσης, η υπόθεση αυτή είναι ότι ορισμένα πολύπλοκα ζητήματα παγκόσμιας σημασίας ενσωματώνονται σε δεδομένα που απαιτούν διασυνοριακή συνεργασία μεταξύ διαφορετικών οργανισμών μέσων ενημέρωσης. Επιπλέον, πιθανόν θα μπορέσουμε να δούμε περισσότερες τέτοιες περιπτώσεις μόλις αναπτυχθούν περαιτέρω οι διαδικασίες για τη συνεχή παρακολούθηση των διεθνών ροών δεδομένων, π.χ. στα χρηματοοικονομικά, όχι απλώς ως υπηρεσία, αλλά και ως βαθύτερες και ερευνητικές ιστορίες παρασκηνίου. Αυτό θα μπορούσε να προωθήσει ένα νέο είδος ερευνητικής δημοσιογραφίας που στηρίζεται σε δεδομένα σε πραγματικό χρόνο, η οποία θα παρακολουθεί συνεχώς ορισμένες ροές οικονομικών δεδομένων, για παράδειγμα, και θα αναζητά αποκλίσεις.

Η διαδραστικότητα μετράει ως κριτήριο ποιότητας στη δημοσιογραφία δεδομένων, αλλά η διαδραστικότητα συνήθως υλοποιείται με ένα σχετικά σαφές σύνολο χαρακτηριστικών –εδώ, τα αποτελέσματά μας συμφωνούν επίσης με άλλες μελέτες και με αυτό που συχνά περιγράφεται ως «έλλειψη πολυπλοκότητας» («lack of sophistication») στη διαδραστικότητα που σχετίζεται με τα δεδομένα (Young κ.α., 2018). Οι χάρτες με δυνατότητα μεγέθυνσης και οι λειτουργίες φιλτραρίσματος χρησιμοποιούνται συχνότερα, ίσως λόγω της τάσης εφαρμογής εύχρηστων ή/και ελεύθερα διαθέσιμων λύσεων λογισμικού, γεγονός που οδηγεί σε λιγότερο εξελιγμένες οπτικοποιήσεις και διαδραστικά χαρακτηριστικά. Ωστόσο, τα βραβευμένα έργα είναι πιθανότερο να διαθέτουν τουλάχιστον μία διαδραστική λειτουργία και να ενσωματώνουν μεγαλύτερο αριθμό διαφορετικών οπτικοποιήσεων. Η τάση προς μάλλον περιορισμένες διαδραστικές επιλογές μπορεί, επίσης, να αντικατοπτρίζει τις εμπειρίες των δημοσιογράφων σχετικά με το

233. https://panamapapers.icij.org/

περιορισμένο ενδιαφέρον του κοινού για εξελιγμένη διαδραστικότητα (όπως οι ευκαιρίες παιχνιδοποιημένης διαδραστικότητας ή τα εργαλεία εξατομίκευσης που καθιστούν δυνατή την προσαρμογή ενός άρθρου με εξατομικευμένα δεδομένα). Ταυτόχρονα, όμως, οι διαδραστικές λειτουργίες και οι οπτικοποιήσεις θα πρέπει να υποστηρίζουν την αφήγηση και την επεξηγηματική λειτουργία ενός άρθρου με τον καλύτερο δυνατό τρόπο –και αυτό απαιτεί λύσεις προσαρμοσμένες σε κάθε ξεχωριστό κομμάτι που βασίζεται σε δεδομένα.

Μια σύνοψη των αναπτυξιακών τάσεων με την πάροδο των ετών δείχνει ένα κάπως μεικτό μοτίβο, καθώς τα ποσοστά και οι μέσοι όροι των υπό μελέτη κατηγοριών έμειναν ως επί το πλείστον σταθερά με την πάροδο του χρόνου ή, αν άλλαξαν, δεν αυξήθηκαν ή μειώθηκαν με γραμμικό τρόπο. Αντιθέτως, διαπιστώσαμε αλλοπρόσαλλες αυξομειώσεις σε μεμονωμένα έτη, γεγονός που υποδηλώνει την εξέλιξη όπως προκύπτει μέσα από δοκιμές και λάθη, καθώς θα περίμενε κανείς σε έναν ακόμη ανερχόμενο τομέα σαν τη δημοσιογραφία δεδομένων. Ως εκ τούτου, διαπιστώσαμε λίγες σταθερές εξελίξεις με την πάροδο των ετών: ένα σημαντικά αυξανόμενο ποσοστό επιχειρηματικών άρθρων, ένας σταθερά και σημαντικά αυξανόμενος μέσος όρος διαφορετικών ειδών οπτικοποιήσεων και ένα –όχι στατιστικά σημαντικό– συνεχώς αυξανόμενο ποσοστό άρθρων που περιλάμβαναν κριτική (π.χ. για τις άδικες μεθόδους κατάσχεσης της αστυνομίας) ή ακόμη και εκκλήσεις για δημόσια παρέμβαση (π.χ. σε σχέση με τις εκπομπές διοξειδίου του άνθρακα). Το ποσοστό αυτό αυξήθηκε σταθερά μέσα στην τετραετία (46,4% το 2013 έναντι 63,0% το 2016) και ήταν σημαντικά υψηλότερο μεταξύ των βραβευμένων (62,2% έναντι 50,0%). Μπορούμε να το ερμηνεύσουμε αυτό ως ένδειξη της μεγάλης εκτίμησης της ερευνητικής και ελεγκτικής δυναμικής της δημοσιογραφίας (δεδομένων) και, ίσως, ως έναν τρόπο νομιμοποίησης αυτού του ανερχόμενου πεδίου.

Από τη δημοσιογραφία δεδομένων στη δεδομενοποιημένη δημοσιογραφία – και ο ρόλος της στην κοινωνία των δεδομένων

Η δημοσιογραφία δεδομένων αντιπροσωπεύει την ανάδυση ενός νέου επιμέρους δημοσιογραφικού πεδίου που εξελίσσεται παράλληλα με τον εκσυγχρονισμό της κοινωνίας –ένα λογικό βήμα στην προσαρμογή της δημοσιογραφίας στην αυξανόμενη διαθεσιμότητα των δεδομένων. Ωστόσο, η δημοσιογραφία δεδομένων δεν είναι πλέον ένα πρωτοεμφανιζόμενο φαινόμενο· έχει, στην πραγματικότητα, καθιερωθεί ως επικρατούσα πρακτική. Μια αξιοσημείωτη ένδειξη αυτού του γεγονότος μπορεί να βρεθεί και πάλι όταν κοιτάξουμε τα Data Journalism Awards. Ο διαγωνισμός του 2018 εισήγαγε μια νέα κατηγορία με την ονομασία «καινοτομία στη δημοσιογραφία δεδομένων», η οποία υποδηλώνει ότι η δημοσιογραφία δεδομένων δεν θεωρείται πλέον καινοτόμος τομέας από μόνη της, αλλά αναζητά νέες προσεγγίσεις στη σύγχρονη πρακτική.

Μπορούμε να περιμένουμε ότι η σημασία και η εξάπλωση της δημοσιογραφίας δεδομένων θα εξελίσσεται παράλληλα με την αυξανόμενη δεδομενοποίηση της κοινωνίας στο σύνολό της –μιας κοινωνίας στην οποία η παραγωγή νοήματος, οι αποφάσεις και όλα τα είδη κοινωνικών δράσεων βασίζονται ολοένα και περισσότερο στα δεδομένα. Σε αυτό το πλαίσιο, δεν είναι πολύ δύσκολο να αντιληφθούμε ότι ο όρος «δημοσιογραφία δεδομένων» θα καταστεί περιττός στο όχι πολύ μακρινό μέλλον, επειδή η δημοσιογραφία στο σύνολό της, καθώς και το περιβάλλον στο οποίο εντάσσεται, δεδομενοποιείται ολοένα και περισσότερο. Είτε αυτή η πρόγνωση επιβεβαιωθεί είτε όχι, ο όρος «δημοσιογραφία δεδομένων», όπως και ο όρος «κοινωνία των δεδομένων», εξακολουθεί να μας προβληματίζει για βασικές διαδικασίες μετασχηματισμού στη δημοσιογραφία και όχι μόνο. Αυτό περιλαμβάνει τον τρόπο και τα μέσα με τα οποία η δημοσιογραφία παρατηρεί και καλύπτει την (δεδομενοποιημένη) κοινωνία, τον τρόπο με τον οποίο αυτοελέγχει τις επιδόσεις της, τον τρόπο με τον οποίο ελέγχει την απήχησή της και τη συμμετοχή του κοινού και τον τρόπο με τον οποίο (αυτόματα) παράγει και διανέμει περιεχόμενο. Με άλλα λόγια, η σύγχρονη δημοσιογραφία χαρακτηρίζεται από τον μετασχηματισμό της προς μια περισσότερο βασισμένη σε δεδομένα, αλγοριθμική, μετρικού τύπου ή ακόμη και αυτοματοποιημένη πρακτική.

Ωστόσο, τα δεδομένα δεν αποτελούν «πρώτη ύλη»· δεν επιτρέπουν άμεση, αντικειμενική ή άλλως προνομιακή πρόσβαση στην κοινωνία (Borgman, 2015). Αυτός ο τρόπος κατανόησης των δεδομένων είναι ακόμη πιο σημαντικός για μια υπεύθυνη δημοσιογραφία δεδομένων, καθώς προχωράει η δεδομενοποίηση της κοινωνίας. Η προώθηση της δεδομενοποίησης και η αυξανόμενη σημασία της δημοσιογραφίας με γνώμονα τα δεδομένα μπορεί, επίσης, να δημιουργήσει κίνητρα για άλλους κοινωνικούς τομείς να παράγουν ή να διαθέσουν περισσότερα δεδομένα (στους δημοσιογράφους). Μάλιστα, είναι πιθανό να δούμε τη συν-εξέλιξη του «PR δεδομένων», δηλαδή των *δημόσιων σχέσεων με γνώμονα τα δεδομένα* που παράγονται και δημοσιεύονται, για να επηρεάσουν τις δημόσιες επικοινωνίες για τους δικούς τους σκοπούς. Αυτό σημαίνει ότι οι διαδικασίες για τον έλεγχο της ποιότητας, της προέλευσης και της σπουδαιότητας των δεδομένων αποκτούν ολοένα και μεγαλύτερη σημασία για τη δημοσιογραφία (δεδομένων) και θέτουν το ερώτημα γιατί μπορεί να μην υπάρχουν διαθέσιμα δεδομένα για ορισμένα γεγονότα ή εξελίξεις.

Συνοψίζοντας, μπορώ να οργανώσω τα ευρήματά μας σύμφωνα με επτά προκλήσεις και ανεκμετάλλευτες ικανότητες της δημοσιογραφίας δεδομένων που μπορεί, επίσης, να φανούν χρήσιμες, για να προτείνουμε τροποποιημένες ή εναλλακτικές πρακτικές στον χώρο.

Συλλογή. Η ερευνητική και κριτική δημοσιογραφία δεδομένων πρέπει να ξεπεράσει την εξάρτησή της από τα δημόσια προσβάσιμα δεδομένα. Πρέπει να καταβληθεί μεγαλύτερη προσπάθεια για την απόκτηση πρόσβασης σε δεδομένα και την ανεξάρτητη συλλογή τους.

Συνεργασία. Ακόμη και αν η παραγωγή «καθημερινών» άρθρων με γνώμονα τα δεδομένα γίνεται όλο και πιο εύκολη, τα πιο απαιτητικά έργα απαιτούν πόρους και προσωπικό και αναμένεται ότι ο αριθμός των θεμάτων που σχετίζονται με την παγκόσμια πραγματικότητα θα αυξηθεί. Αυτά θα απαιτήσουν έρευνες που βασίζονται σε δεδομένα, πέρα από σύνορα και οργανισμούς μέσων ενημέρωσης, και, σε ορισμένες περιπτώσεις, συνεργασία με άλλους κλάδους, όπως η επιστήμη ή ο ακτιβισμός δεδομένων.

Πληθοπορισμός («crowdsourcing»). Η πραγματική διαδραστική δυναμική της δημοσιογραφίας δεδομένων δεν έγκειται σε όλο και πιο εξελιγμένα διαδραστικά χαρακτηριστικά αλλά σε προσεγγίσεις crowdsourcing που εμπλέκουν εμπράκτως τους χρήστες ή τους πολίτες ως συλλέκτες, ταξινομητές και συν-ερευνητές των δεδομένων (Gray, 2018).

Συνδημιουργία. Οι προσεγγίσεις συνδημιουργίας, συνηθισμένες στον τομέα της ανάπτυξης λογισμικού, μπορούν να χρησιμεύσουν ως μοντέλο για μακρόπνοα έργα με γνώμονα τα δεδομένα. Στις περιπτώσεις αυτές, οι χρήστες εμπλέκονται σε ολόκληρη τη διαδικασία, από την εξεύρεση ενός θέματος έως την ανάπτυξή του και τη διαχείρισή του για μεγάλο χρονικό διάστημα.

Ικανότητα. Η ποιοτική δημοσιογραφία δεδομένων απαιτεί ομάδες με ευρύ σύνολο δεξιοτήτων. Ο ρόλος του δημοσιογράφου παραμένει σημαντικός, αλλά οι δημοσιογράφοι χρειάζεται να έχουν όλο και περισσότερο μια πιο εξελιγμένη κατανόηση των δεδομένων, των δομών δεδομένων και των αναλυτικών μεθόδων. Οι οργανισμοί μέσων, με τη σειρά τους, χρειάζονται πόρους για την πρόσληψη αναλυτών δεδομένων, οι οποίοι είναι προοδευτικά όλο και πιο περιζήτητοι σε πολλούς άλλους κλάδους.

Συνδυασμός. Τα ολοένα και πιο σύνθετα δεδομένα απαιτούν ολοένα και πιο περίπλοκη ανάλυση. Μέθοδοι που συνδυάζουν πηγές δεδομένων και εξετάζουν αυτά τα δεδομένα από διάφορες οπτικές γωνίες θα μπορούσαν να βοηθήσουν στη διαμόρφωση ουσιαστικότερων εικόνων των κοινωνικών φαινομένων και να ενισχύσουν την αναλυτική ικανότητα της δημοσιογραφίας δεδομένων.

Πολυπλοκότητα. Η πολυπλοκότητα δεν περιλαμβάνει μόνο τα ίδια τα δεδομένα, αλλά και την αυξανόμενη σημασία τους για διάφορους κοινωνικούς τομείς και για τη λήψη πολιτικών αποφάσεων. Κατά τη διάρκεια αυτών των εξελίξεων, η δημοσιογραφία δεδομένων θα έρχεται όλο και πιο συχνά αντιμέτωπη με το PR δεδομένων και τα «ψευδή δεδομένα» («fake data»).

Τι σημαίνει αυτό; Λαμβάνοντας υπόψη όσα ήδη γνωρίζουμε για τη (βραβευμένη) δημοσιογραφία δεδομένων όσον αφορά στα είδη της που χαίρουν εκτίμησης, λαμβάνουν ευρεία δημόσια προβολή και συμβάλλουν στη γενική αναγνώριση της δημοσιογραφίας, ποια είδη δημοσιογραφίας δεδομένων θέλουμε πραγματικά; Από αυτήν την άποψη, θα υποστήριζα ότι η υπεύθυνη δημοσιογραφία δεδομένων στην κοινωνία των δεδομένων είναι μια δημοσιογραφία που: διερευνά ζητήματα τα οποία σχετίζονται με την κοινωνία και καθιστά την κοινωνία των δεδομένων κατανοητή και αντικείμενο κριτικής από τα ίδια της τα μέσα· έχει επίγνωση των δικών της τυφλών σημείων, ενώ διερωτάται γιατί υπάρχουν ελλείψεις δεδομένων σε ορισμένους τομείς και αν αυτό αποτελεί καλό ή κακό σημάδι· προσπαθεί ενεργά να αποκαλύψει τη χειραγώγηση των δεδομένων και την κατάχρηση δεδομένων· και, τέλος, λαμβάνει υπόψη, εξηγεί και τονίζει τον χαρακτήρα των δεδομένων ως «ανθρώπινα τεχνουργήματα» που δεν αποτελούν, σε καμία περίπτωση, αυταπόδεικτες συλλογές γεγονότων, αλλά συχνά συλλέγονται σε σχέση με πολύ συγκεκριμένες συνθήκες και στόχους (Krippendorff, 2016).

Αυτό, όμως, σημαίνει ότι η ιδιαιτερότητα της δημοσιογραφίας δεδομένων, η εξάρτησή της από τα δεδομένα, είναι ταυτόχρονα και η αδυναμία της. Ο περιορισμός αυτός έχει να κάνει με τη διαθεσιμότητα των δεδομένων, την αξιοπιστία τους, την ποιότητά τους και τη δυνατότητα χειραγώγησής τους. Μια υπεύθυνη δημοσιογραφία δεδομένων θα πρέπει να είναι αναστοχαστική όσον αφορά στην εξάρτησή της από τα δεδομένα –και αυτό θα πρέπει να αποτελεί βασικό θέμα της συζήτησης για τη δεοντολογία στη δημοσιογραφία δεδομένων. Αυτές οι συνθήκες υποδεικνύουν ότι η δημοσιογραφία δεδομένων δεν είναι μόνο ένα νέο ρεπορταζιακό στυλ αλλά και ένα μέσο παρέμβασης που προκαλεί και αμφισβητεί την κοινωνία των δεδομένων –μια κοινωνία φορτωμένη με βασικά επιστημολογικά ερωτήματα που συγκρούονται με τις παραδοχές τής δημοσιογραφίας ως προς το *τι* (μπορούμε να) γνωρίζουμε και το *πώς* γνωρίζουμε (μέσω των δεδομένων).

Τα ερωτήματα αυτά γίνονται όλο και πιο επιτακτικά, καθώς όλο και περισσότερα και όλο και πιο διαφορετικά δεδομένα ενσωματώνονται σε διάφορα σημεία του «κύκλου των ειδήσεων»: ως μέσο δημοσιογραφικής παρατήρησης και έρευνας, ως μέρος των διαδικασιών παραγωγής και διανομής και ως μέσο παρακολούθησης των καταναλωτικών δραστηριοτήτων του κοινού. Με αυτούς τους τρόπους, η δεδομενοποιημένη δημοσιογραφία επηρεάζει: (α) τον τρόπο με τον οποίο η *δημοσιογραφία παρατηρεί τον κόσμο* και δομεί τις ειδήσεις από δεδομένα, (β) τον πυρήνα των επιδόσεων της δημοσιογραφίας, διευκολύνοντας την αυτοματοποίηση της παραγωγής περιεχομένου, (γ) *τη διανομή και την κυκλοφορία της παραγωγής της δημοσιογραφίας* σε ένα περιβάλλον που διαμορφώνεται από αλγορίθμους και την υποκείμενη λογική τους για την επεξεργασία των δεδομένων και (δ) το τι *νοείται ως άξια λόγου ειδησεογραφία* σε ολοένα και πιο αναλυτικά καταμετρημένες κατηγορίες κοινού.

Οι εξελίξεις αυτές επιφορτίζουν τη δημοσιογραφία (δεδομένων) με τρεις βασικές ευθύνες: να παρακολουθεί κριτικά την εξέλιξή μας προς μια δεδομενοποιημένη κοινωνία, να καταστήσει αυτήν την κοινωνία κατανοητή με τα δικά της μέσα και να καταστήσει ορατά τα όρια του τι μπορεί και τι πρέπει να αφηγηθεί και να δει κανείς υπό το πρίσμα των δεδομένων.

Παραπομπές

Ausserhofer, J., Gutounig, R., Oppermann, M., Matiasek, S., & Goldgruber, E. (2020) «The datafication of data journalism scholarship: Focal points, methods, and research propositions for the investigation of data-intensive newswork», Journalism, 21(7), σσ. 950–973. DOI: https://doi.org/10.1177/1464884917700667

Borges-Rey, E. (2020) «Towards an epistemology of data journalism in the devolved nations of the United Kingdom: Changes and continuities in materiality, performativity and reflexivity», *Journalism*, 21(7), σσ. 915–932. DOI: https://doi.org/10.1177/1464884917693864

Borgman, C. L. (2015). *Big data, little data, no data: Scholarship in the networked world.* Κέιμπριτζ, Μασαχουσέτη: The MIT Press.

Dijck, J. van. (2014) «Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology», *Surveillance & Society,* 12(2), σσ. 197–208. DOI: https://doi.org/10.24908/ss.v12i2.4776

English, J. F. (2002) «Winning the culture game: Prizes, awards, and the rules of art», *New Literary History,* 33(1), σσ. 109–135. Διαθέσιμο σε: https://www.jstor.org/stable/20057712 (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Gray, J. (2018) «New project: What can citizen-generated data do? Research collaboration around UN Sustainable Development Goals», *Jonathan Gray,* 8 Αυγούστου. Διαθέσιμο σε: https://jonathangray.org/2018/08/08/what-can-citizen-generated-data-do/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Knight, M. (2015) «Data journalism in the UK: A preliminary analysis of form and content», *Journal of Media Practice,* 16, σσ. 55–72. DOI: https://doi.org/10.1080/14682753.2015.1015801

Krippendorff, K. (2016) «Data», στο K. B. Jensen & R. T. Craig (επιμ.), *The international encyclopedia of communication theory and philosophy, vol. 1: A–D*, σσ. 484–489. Wiley Blackwell. DOI: https://doi.org/10.1002/9781118766804.wbiect104

Loosen, W. (2018) «Four forms of datafied journalism. Journalism's response to the datafication of society», *Communicative Figurations,* 18. Διαθέσιμο σε: https://www.kommunikative-figurationen.de/fileadmin/user_upload/Arbeitspapiere/CoFi_EWP_No-18_Loosen.pdf (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Loosen, W., Reimer, J., & De Silva-Schmidt, F. (2020) «Data-driven reporting: An on-going (r) evolution? An analysis of projects nominated for the Data Journalism Awards 2013–2016», *Journalism*, 21(9), σσ. 1246–1263. DOI: https://doi.org/10.1177/1464884917735691

Parasie, S. (2015) «Data-driven revelation?», *Digital Journalism,* 3(3), σσ. 364–380. DOI: https://doi.org/10.1080/21670811.2014.976408

Royal, C., & Blasingame, D. (2015) *«Data journalism: An explication»,* #ISOJ, 5(1), σσ. 24–46. Διαθέσιμο σε: https://isojjournal.wordpress.com/2015/04/15/data-journalism-an-explication/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Tabary, C., Provost, A.-M., & Trottier, A. (2016) «Data journalism's actors, practices and skills: A case study from Quebec», *Journalism: Theory, Practice & Criticism,* 17(1), σσ. 66–84. DOI: https://doi.org/10.1177/1464884915593245

Wahl-Jorgensen, K. (2013) «The strategic ritual of emotionality: A case study of Pulitzer Prize–winning articles», *Journalism: Theory, Practice & Criticism,* 14(1), σσ. 129–145. DOI: https://doi.org/10.1177/1464884912448918

Young, M. L., Hermida, A., & Fulda, J. (2018) «What makes for great data journalism? A content analysis of Data Journalism Awards finalists 2012–2015», *Journalism Practice,* 12(1), σσ. 115–135. DOI: https://doi.org/10.1080/17512786.2016.1270171

Η συγγραφέας

Η Βίμπκε Λούζεν είναι επικεφαλής ερευνήτρια της δημοσιογραφίας στο Leibniz Institute for Media Research/Hans-Bredow-Institut (HBI) στο Αμβούργο, καθώς και Καθηγήτρια στο Πανεπιστήμιο του Αμβούργου.

9.3

Πέρα από τα Κλικ και τις Κοινοποιήσεις: Πώς και Γιατί Μετράμε τον Αντίκτυπο των Έργων Δημοσιογραφίας Δεδομένων

Lindsay Green-Barber

Περίληψη: Στο κεφάλαιο αυτό υποστηρίζεται ότι η δημοσιογραφία δεδομένων απολαμβάνει μια μοναδική θέση που της επιτρέπει να έχει αντίκτυπο σε άτομα, δίκτυα και θεσμούς, και προτείνονται στρατηγικές για τη μέτρηση του αντικτύπου αυτής της δουλειάς.

Λέξεις-κλειδιά: αντίκτυπος, κοινωνική επιστήμη, μέτρηση αντικτύπου, ανάλυση, δημοσιογραφία δεδομένων, εμπλοκή του κοινού

Δημοσιογραφία και αντίκτυπος

Αν και πολλοί δημοσιογράφοι διαφωνούν με την ιδέα του δημοσιογραφικού αντικτύπου, στην πραγματικότητα η σύγχρονη δημοσιογραφία, ως επάγγελμα, είναι χτισμένη πάνω σε έναν θεμέλιο λίθο του αντικτύπου: στην ενημέρωση του κοινού, έτσι ώστε να έχουμε πολιτική συμμετοχή και να θέτουμε τους ισχυρούς προ των ευθυνών τους. Και ενώ οι δημοσιογράφοι ανησυχούν ότι η σκέψη, η συζήτηση, η χάραξη στρατηγικής και η μέτρηση του θετικού (και αρνητικού) αντικτύπου της δουλειάς τους οδηγεί επικίνδυνα κοντά στην κόκκινη γραμμή που χωρίζει τη δημοσιογραφία από τον ακτιβισμό, τόσο οι επαγγελματίες όσο και οι σχολιαστές έχουν «χύσει πολύ μελάνι», γράφοντας για τις αρνητικές επιπτώσεις που έχουν οι ψευδείς ειδήσεις («fake news»), η παραπληροφόρηση και η κομματική πληροφόρηση στα άτομα, στην κοινωνία μας και στη δημοκρατία. Με άλλα λόγια, ενώ οι δημοσιογράφοι θέλουν να αποφεύγουν να μιλούν για τον αντίκτυπο της δουλειάς τους, αναγνωρίζουν τις σοβαρές κοινωνικές, πολιτικές και πολιτιστικές επιπτώσεις των fake news.

Επιπλέον, πριν να εξελιχθεί η δημοσιογραφία σε ανεξάρτητο επάγγελμα στα τέλη του 19ου και στις αρχές του 20ού αιώνα, ήταν μια πρακτική επιρροής, η οποία υποστηριζόταν από τα πολιτικά κόμματα και παραγόταν με ρητό στόχο την υποστήριξη του κόμματος και την εξασφάλιση της εκλογής των υποψηφίων του (Pitt & Green-Barber, 2017). Έτσι, από ιστορική άποψη, η μετεξέλιξη της δημοσιογραφίας σε επάγγελμα και η υιοθέτηση του μύθου της ουδετερότητας είναι στην πραγματικότητα αρκετά καινούριες ιδέες (Groseclose & Milyo, 2005· Hamilton, 2004). Και η αναζήτηση της «ουδετερότητας» από τη δημοσιογραφία δεν ήταν μια κανονιστική απόφαση αλλά μάλλον μια συνάρτηση των μεταβαλλόμενων οικονομικών μοντέλων και της ανάγκης να απευθυνθεί σε όσο το δυνατόν μεγαλύτερο κοινό, προκειμένου να αποφέρει έσοδα (Hamilton, 2004).

Δεδομένων των ταυτόχρονων και στενά συνδεδεμένων κρίσεων του επιχειρηματικού μοντέλου της ειδησεογραφικής βιομηχανίας και της έλλειψης εμπιστοσύνης του κοινού στα μέσα ενημέρωσης στις Ηνωμένες Πολιτείες και στη Δυτική Ευρώπη, θα μπορούσε κανείς να ισχυριστεί ότι η απομάκρυνση της δημοσιογραφίας από την αναγνώριση του αντικτύπου της ήταν, στην καλύτερη περίπτωση, αποποίηση ευθύνης και, στη χειρότερη, αποτυχία.

Υπάρχουν, όμως, και ελπιδοφόρα σημάδια. Τα τελευταία χρόνια, ορισμένοι ενημερωτικοί οργανισμοί έχουν αρχίσει να αποδέχονται το γεγονός ότι ασκούν επιρροή στην κοινωνία. Ο πολλαπλασιασμός των μη κερδοσκοπικών μέσων ενημέρωσης, που συχνά υποστηρίζονται από φιλανθρωπικά ιδρύματα και ιδιώτες με όραμα, δημιούργησε έναν «δοκιμαστικό σωλήνα» για πειραματισμούς σχετικά με τον αντίκτυπο. Πολλά εμπορικά μέσα ενημέρωσης έχουν, επίσης, καταλήξει στην ιδέα ότι το να ανακοινώνουν τον θετικό αντίκτυπο του έργου τους στο κοινό είναι μια στρατηγική για την οικοδόμηση εμπιστοσύνης και αφοσίωσης, η οποία ελπίζουν ότι θα μεταφραστεί σε αύξηση των εσόδων. Για παράδειγμα, το 2017, η *Washington Post* έβαλε το *Democracy Dies in Darkness* («Η Δημοκρατία Πεθαίνει στο Σκοτάδι») στην επικεφαλίδα της, ενστερνιζόμενη (και διαφημίζοντας) τον ρόλο της στο πολιτικό μας σύστημα. Και το CNN δημιούργησε την ενότητα *Impact Your World* («Επηρέασε τον Κόσμο σου») στον ιστότοπό του, συνδέοντας τα παγκόσμια γεγονότα, το ρεπορτάζ του, τις ιστορίες «αντικτύπου» και τους τρόπους με τους οποίους το κοινό μπορεί να αναλάβει δράση, από καμπάνιες με hashtag μέχρι δωρεές[234].

Οι ενημερωτικοί οργανισμοί έχουν, επίσης, αρχίσει να δοκιμάζουν νέες στρατηγικές, για να μεγιστοποιήσουν τον θετικό αντίκτυπο του έργου τους, καθώς και να χρησιμοποιούν μεθόδους έρευνας και μετρήσεις διαφορετικές από εκείνες που χρησιμοποιούνται στη διαφήμιση, προκειμένου να κατανοήσουν την αποτελεσματικότητα αυτών των στρατηγικών. Ενώ, σε ορισμένες περιπτώσεις, οι ψηφιακές μετρήσεις μπορούν να είναι χρήσιμα αντιπροσωπευτικά στοιχεία για τη μέτρηση του αντικτύπου, οι διαφημιστικές μετρήσεις, όπως οι μοναδικές προβολές σελίδων («page views») ή ακόμη και οι πιο προηγμένες αναλύσεις, όπως ο χρόνος παραμονής σε μια σελίδα, αποσκοπούν στη μέτρηση της απήχησης του περιεχομένου, χωρίς να λαμβάνονται υπόψη οι επιπτώσεις αυτού του περιεχομένου σε ένα άτομο.

Θα ήθελα να προτείνω ένα πλαίσιο για τον αντίκτυπο των μέσων που διαφοροποιείται από την υφιστάμενη κατάσταση ως αποτέλεσμα παρέμβασης και περιλαμβάνει τέσσερις τύπους αντικτύπου: στα άτομα, στα δίκτυα, στους θεσμούς και στον δημόσιο λόγο. Αυτοί οι τύποι αντικτύπου είναι αλληλοσυνδεόμενοι. Για παράδειγμα, όπως συχνά υποθέτει η δημοσιογραφία, το ρεπορτάζ μπορεί να αυξήσει το επίπεδο γνώσης των ατόμων για ένα θέμα, με αποτέλεσμα αυτά να ψηφίσουν με συγκεκριμένο τρόπο και, τελικά, να επηρεάσουν τους θεσμούς. Ή, ένα ρεπορτάζ μπορεί να έχει άμεσες επιπτώσεις στα θεσμικά όργανα, όπως μια απόλυση ή μια αναδιάρθρωση, η οποία στη συνέχεια επηρεάζει τα μεμονωμένα άτομα. Ωστόσο, ο αντίκτυπος που πυροδοτείται από τη δημοσιογραφία συχνά απαιτεί χρόνο και περιλαμβάνει πολύπλοκες κοινωνικές διαδικασίες. Διαφορετικοί τύποι δημοσιογραφίας είναι καλύτερα προσαρμοσμένοι σε διαφορετικούς τύπους αντικτύπου. Για παράδειγμα, ο Τζέιμς Τ. Χάμιλτον αποκαλύπτει ότι το ερευνητικό ρεπορτάζ μπορεί να εξοικονομήσει χρήματα για τους θεσμούς, αποκαλύπτοντας κακοδιαχείριση, διαφθορά ή αδικήματα και ενθαρρύνοντας την αλλαγή. Και το ντοκιμαντέρ ως είδος έχει αποδειχθεί ιδιαίτερα αποτελεσματικό στη δημιουργία νέων ή/και ενισχυμένων δικτύων δράσης για την προώθηση της αλλαγής (Green-Barber, 2014).

Το υπόλοιπο του παρόντος κεφαλαίου διερευνά τη σχέση μεταξύ της δημοσιογραφίας δεδομένων και του αντικτύπου, καταδεικνύοντας τον τρόπο με τον οποίο η δημοσιογραφία δεδομένων μπορεί να συμβάλλει σε διάφορους τύπους κοινωνικής αλλαγής. Στη συνέχεια, προτείνει μεθόδους μέτρησης της αποτελεσματικότητας της δημοσιογραφίας δεδομένων και τι μπορούν να κάνουν οι δημοσιογράφοι και οι ειδησεογραφικοί οργανισμοί με αυτές τις πληροφορίες.

Γιατί δημοσιογραφία δεδομένων;

Ενώ οι δημοσιογράφοι χρησιμοποιούν τη δημοσιογραφία δεδομένων για πολλούς λόγους, υπάρχουν δύο από αυτούς που ξεχωρίζουν: πρώτον, η παροχή αξιόπιστων στοιχείων για την υποστήριξη των ισχυρισμών που διατυπώνονται στην αφήγηση του θέματος και, δεύτερον, η παρουσίαση

234. https://edition.cnn.com/specials/impact-your-world

των πληροφοριών στο κοινό ως δεδομένα, και όχι ως κειμενικές αφηγήσεις. Η πρακτική της δημοσιογραφίας δεδομένων είναι δομημένη πάνω σε μια θεμελιώδη αξιολογική κρίση ότι τα δεδομένα είναι αξιόπιστα και, κατ' επέκταση, ότι ένα δημοσιογραφικό προϊόν που περιλαμβάνει ρεπορτάζ με δεδομένα είναι αξιόπιστο –και ενδεχομένως περισσότερο από όσο θα ήταν χωρίς αυτά.

Το ρεπορτάζ δεδομένων που χρησιμοποιείται για να μεταδώσει πληροφορίες ως στατικούς αριθμούς, δεδομένα, διαγράμματα, γραφήματα ή άλλα οπτικά στοιχεία είναι παρόμοιο με άλλες δημοσιογραφικές μορφές (π.χ. κείμενο, βίντεο, ήχος), καθώς αποτελεί ουσιαστικά μια γραμμική μορφή μετάδοσης επιλεγμένων πληροφοριών σε ένα ακροατήριο. Το ρεπορτάζ δεδομένων που διατίθεται στο κοινό μέσω μιας ειδησεογραφικής διαδραστικής εφαρμογής είναι μια μοναδική μορφή αφήγησης, καθώς προϋποθέτει ότι το κοινό θα αλληλεπιδράσει με τα δεδομένα, θα θέσει τα δικά του ερωτήματα και θα αναζητήσει απαντήσεις μέσα στα δεδομένα που έχει στη διάθεσή του. Έτσι, η «ιστορία» εξαρτάται από τον χρήστη όσο και από τη δημοσιογραφία.

Ακόμα και αυτή η ακατέργαστη εκδοχή της δημοσιογραφίας δεδομένων εμπλέκει και τους τέσσερις τύπους αντικτύπου.

Άτομα

Η δημοσιογραφία δεδομένων τείνει να εστιάζει στα μεμονωμένα μέλη του κοινού ως τη δυνητική μονάδα της αλλαγής, παρέχοντας στο κοινό αξιόπιστες πληροφορίες, ώστε αυτό να αποκτήσει περισσότερες γνώσεις και, κατ' επέκταση, να λάβει πιο συνειδητοποιημένες αποφάσεις. Και ενώ η δημοσιογραφία δεδομένων ως βάση για την παραδοσιακή, γραμμική αφήγηση αυξάνει την εμπιστοσύνη του κοινού στο περιεχόμενο, οι ειδήσεις ή τα διαδραστικά δεδομένα παρέχουν τη μεγαλύτερη εκείνη δυναμική, ώστε να έχει η δημοσιογραφία δεδομένων αντίκτυπο σε ατομικό επίπεδο.

Με τα διαδραστικά δεδομένα, δηλαδή μια «μεγάλη διαδραστική βάση δεδομένων που αφηγείται μια ειδησεογραφική ιστορία», ο χρήστης μπορεί να δημιουργήσει τη δική του ερώτηση και να αναζητήσει απαντήσεις στα δεδομένα (Klein, 2012). Οι εταιρείες των μέσων συχνά υποθέτουν ότι τα διαδραστικά δεδομένα θα επιτρέψουν στο κοινό να εμβαθύνει και να εξερευνήσει δεδομένα, να βρει σχετικές πληροφορίες και να αφηγηθεί ιστορίες. Σε μια ανάλυση των διαδραστικών δεδομένων από έναν ειδησεογραφικό οργανισμό, η συγγραφέας του παρόντος κεφαλαίου διαπίστωσε ότι οι πιο επιτυχημένες εφαρμογές δεδομένων (δηλαδή, εκείνες που είχαν μεγάλη επισκεψιμότητα και διερευνήθηκαν σε βάθος) αποτελούσαν μέρος ενός πλήρους εκδοτικού πακέτου που περιλάμβανε και άλλο περιεχόμενο, προσέφεραν τη δυνατότητα αναζήτησης γεωγραφικών δεδομένων σε τοπικό επίπεδο ή ανάλογα με τον βαθμό συνάφειας, είχαν υψηλό βαθμό διαδραστικότητας, ήταν αισθητικά ευχάριστες και καλά σχεδιασμένες και φορτώνονταν γρήγορα (Green-Barber, 2015b).

Το *Dollars for Docs* της *ProPublica* είναι ένα κλασικό παράδειγμα δημοσιογραφίας δεδομένων, ως προς το ότι εξασφαλίζει πρόσβαση σε σημαντικές ποσότητες δεδομένων (στην προκείμενη περίπτωση, για τις πληρωμές των φαρμακευτικών εταιρειών και των εταιρειών ιατρικών συσκευών προς τους γιατρούς), δομεί τα δεδομένα και τα παρουσιάζει στο κοινό ως μια διαδραστική βάση δεδομένων, με στόχο να εμπνεύσει τα άτομα να διεξάγουν τη δική τους έρευνα και ενδεχομένως να αναλάβουν δράση[235]. Το έργο καθοδηγεί το κοινό να «χρησιμοποιήσει αυτό το εργαλείο», για να αναζητήσει πληρωμές προς τους γιατρούς του και, σε μια πλευρική μπάρα, γράφει: «Ασθενείς, αναλάβετε δράση. Θέλουμε να μάθουμε πώς χρησιμοποιήσατε ή θα μπορούσατε να χρησιμοποιήσετε αυτές τις πληροφορίες στην καθημερινή σας ζωή. Μιλήσατε με τον γιατρό σας; Σκοπεύετε να το κάνετε; Πείτε μας»[236].

235. https://projects.propublica.org/docdollars/
236. https://propublica.forms.fm/was-the-dollars-for-docs-information-helpful-tell-us-how-you-will-use-it/forms/2249<

Δίκτυα

Η δημοσιογραφία δεδομένων παρέχει αξιόπιστες πληροφορίες που μπορούν να χρησιμοποιηθούν από τα δίκτυα (επίσημα ή/και ανεπίσημα), για να ενισχύσουν τις θέσεις και το έργο τους. Για παράδειγμα, οι φορείς δράσης χρησιμοποιούν συχνά το ρεπορτάζ δεδομένων, για να ενισχύσουν τους ισχυρισμούς τους σε δημόσιες προσφυγές ή σε νομικές διαδικασίες, ιδίως σε περιπτώσεις όπου τα δεδομένα δεν είναι διαθέσιμα στο κοινό. Η δημοσιογραφική πρακτική της αίτησης πρόσβασης σε δεδομένα που δεν είναι διαθέσιμα στον δημόσιο χώρο, της ανάλυσης αυτών των δεδομένων και της δημοσίευσης των ευρημάτων απορροφά κόστη που διαφορετικά θα ήταν δυσβάσταχτα για τα άτομα ή τα δίκτυα (Hamilton, 2016).

Θεσμοί

Η δημοσιογραφία δεδομένων μπορεί να παράγει δημοσιεύματα που τα θεσμικά όργανα προσπαθούν σκληρά να κρατήσουν κρυφά, καθώς αποτελούν στοιχεία διαφθοράς, κακοδιαχείρισης, παραβατικότητας ή/και ανικανότητας. Όταν αυτές οι πληροφορίες έρχονται στο φως της δημοσιότητας, ασκείται πίεση στα θεσμικά όργανα για μεταρρυθμίσεις –ως αποτέλεσμα της απειλής που συνδέεται με τις εκλογές για τους πολιτικούς, ή με τις δυνάμεις της αγοράς για τις εισηγμένες στο χρηματιστήριο επιχειρήσεις.

Για παράδειγμα, στο πλαίσιο της συνεργατικής έρευνας της Διεθνούς Σύμπραξης Ερευνητών Δημοσιογράφων («International Consortium of Investigative Journalists», ICIJ) για τα *Panama Papers* αναλύθηκαν περισσότερα από 11,5 εκατομμύρια αρχεία τα οποία αποκάλυψαν «πολιτικούς από περισσότερες από 50 χώρες που συνδέονται με υπεράκτιες εταιρείες σε 21 φορολογικούς παραδείσους»[237]. Η έρευνα αυτή οδήγησε στην παραίτηση πολιτικών, όπως ο πρωθυπουργός της Ισλανδίας, Σίγκμουντουρ Γκουνλάουγκσον, σε έρευνες για άλλους, όπως ο πρώην πρωθυπουργός του Πακιστάν, Ναουάζ Σαρίφ (ο οποίος καταδικάστηκε σε δεκαετή κάθειρξη το 2018), και σε αμέτρητες άλλες αντιδράσεις εκ μέρους των θεσμών.

Δημόσιος λόγος

Επειδή η δημοσιογραφία δεδομένων μπορεί συχνά να αναλυθεί σε μικρότερα μέρη, είτε γεωγραφικά, είτε δημογραφικά, είτε με βάση άλλους παράγοντες, τα δεδομένα μπορούν να χρησιμοποιηθούν για να ειπωθούν διαφορετικές ιστορίες από διαφορετικά μέσα. Με αυτόν τον τρόπο, μπορεί η δημοσιογραφία δεδομένων να γίνει τοπική, ώστε αυτή να φέρει μια αλλαγή στη δημόσια συζήτηση για θέματα που αφορούν γεωγραφικές τοποθεσίες, δημογραφικές ομάδες ή άλλα κοινωνικά στερεότυπα.

Το Center for Investigative Reporting δημοσίευσε εθνικά διαδραστικά σύνολα δεδομένων σχετικά με το Υπουργείο Βετεράνων των ΗΠΑ: ένα με τους μέσους χρόνους αναμονής για τους βετεράνους που επιχειρούν να αποκτήσουν πρόσβαση σε ιατρική περίθαλψη στα νοσοκομεία του Υπουργείου και ένα δεύτερο με τον αριθμό των οπιούχων φαρμάκων που συνταγογραφούνται στους βετεράνους από τα συστήματα του Υπουργείου. Και στις δύο περιπτώσεις, οι τοπικοί δημοσιογραφικοί οργανισμοί χρησιμοποίησαν τα σύνολα δεδομένων ως βάση, για να κάνουν τοπικό ρεπορτάζ σχετικά με τα εν λόγω θέματα.

Πώς μπορούν, λοιπόν, οι δημοσιογράφοι δεδομένων να χαράξουν στρατηγική, για να έχουν αντίκτυπο;

Κάνατε τη σκληρή δουλειά: αποκτήσατε πρόσβαση σε δεδομένα, επεξεργαστήκατε τους αριθμούς, δομήσατε τα δεδομένα και έχετε να πείτε μια σημαντική ιστορία. Τώρα τι γίνεται;

237. https://www.icij.org/investigations/panama-papers/

Μια στρατηγική για μεγάλο αντίκτυπο της δημοσιογραφίας δεδομένων θα μπορούσε να ακολουθήσει τα παρακάτω πέντε βήματα:

Καθορισμός στόχων. Τι μπορεί να συμβεί ως αποτέλεσμα του έργου σας; Ποιος ή ποιοι έχουν τη δύναμη ή/και το κίνητρο να ασχοληθούν με τυχόν παραπτώματα; Ποιος θα πρέπει να έχει πρόσβαση στις πληροφορίες που φέρνετε στο φως; Κάντε στον εαυτό σας αυτές τις ερωτήσεις, προκειμένου να αποφασίσετε ποιος τύπος ή ποιοι τύποι αντικτύπου είναι εύλογοι για το έργο σας.

Περιεχόμενο. Αφού καθορίσετε τους στόχους του έργου σας, προσδιορίστε τις σημαντικές ομάδες κοινού-στόχους για τη δουλειά σας. Ποια πηγή ειδήσεων και πληροφοριών εμπιστεύονται αυτά τα ακροατήρια; Πώς μπορούν να έχουν καλύτερη πρόσβαση στις πληροφορίες; Χρειάζονται μια διαδραστική μορφή ή θα είναι πιο αποτελεσματική μια γραμμική ιστορία;

Αλληλεπίδραση. Πώς θα αλληλεπιδράσετε εσείς και ο ειδησεογραφικός σας οργανισμός με το κοινό και πώς το κοινό θα αλληλεπιδράσει με το έργο σας; Για παράδειγμα, αν έχετε εντοπίσει έναν ειδησεογραφικό οργανισμό διαφορετικό από τον δικό σας ως αξιόπιστη πηγή πληροφοριών για το κοινό-στόχο, συνεργαστείτε. Εάν τα διαδραστικά δεδομένα σας περιέχουν σημαντικές πληροφορίες για την κοινότητα μιας ΜΚΟ, διοργανώστε ένα διαδικτυακό σεμινάριο όπου θα εξηγείται ο τρόπος χρήσης τους.

Στρατηγική έρευνα. Ανάλογα με τους στόχους, το περιεχόμενο και τα σχέδια αλληλεπίδρασης που έχετε οργανώσει, επιλέξτε τις κατάλληλες ερευνητικές μεθόδους ή/και δείκτες, προκειμένου να παρακολουθήσετε την πρόοδο και να κατανοήσετε τι λειτουργεί και τι όχι. Ενώ τα μέσα συχνά αναφέρονται στη «μέτρηση» του αντικτύπου του έργου τους, η ίδια προτιμώ τον όρο «στρατηγική έρευνα», καθώς θα πρέπει να λαμβάνονται υπόψη τόσο οι ποιοτικές όσο και οι ποσοτικές μέθοδοι έρευνας. Όσο πιο γρήγορα μπορείτε να προσδιορίσετε τις μεθόδους έρευνας και τους δείκτες, τόσο καλύτερες θα είναι οι πληροφορίες σας. (Η επόμενη ενότητα εξετάζει τις επιλογές μέτρησης σε μεγαλύτερο βάθος.)

Επανάληψη. Έχετε επενδύσει χρόνο και πόρους στο ρεπορτάζ δεδομένων σας, στο περιεχόμενο, στην αλληλεπίδραση και στη μέτρηση. Τι λειτούργησε; Τι θα αλλάξετε την επόμενη φορά; Ποια ερωτήματα ακόμη εκκρεμούν; Μοιραστείτε αυτά τα συμπεράσματα με την ομάδα σας και τον κλάδο, για να βοηθήσετε το επόμενο έργο να πάει ακόμα πιο μπροστά.

Πώς «μετράμε» τον αντίκτυπο της δουλειάς μας;

Όπως αναφέρθηκε προηγουμένως, η έρευνα για τον αντίκτυπο στα μέσα κυριαρχείται από τις μετρήσεις της διαφήμισης. Ωστόσο, οι διαφημιστικές μετρήσεις, όπως οι προβολές σελίδων, ο χρόνος παραμονής στη σελίδα και το ποσοστό εγκατάλειψης («bounce rate») είναι δυνητικά αντιπροσωπευτικά στοιχεία μόνο για κάποιο βαθμό αντικτύπου. Σκοπός τους είναι να μετρήσουν τη συνολική έκθεση του περιεχομένου σε άτομα, χωρίς να ενδιαφέρονται για τις απόψεις τους σχετικά με τα θέματα, για το αν έμαθαν ή όχι νέες πληροφορίες ή για την πρόθεσή τους να αναλάβουν σχετική δράση, χάρη στο περιεχόμενο. Ωστόσο, όταν εξετάζεται ο αντίκτυπος του περιεχομένου σε άτομα, δίκτυα, θεσμούς και στον δημόσιο διάλογο, υπάρχουν και άλλες καινοτόμες ποιοτικές και ποσοτικές μέθοδοι που μπορούν να χρησιμοποιηθούν για την καλύτερη κατανόηση της επιτυχίας και της αποτυχίας στον τομέα αυτό. Αυτή η ενότητα διερευνά ορισμένες πολλά υποσχόμενες ερευνητικές μεθόδους για την κατανόηση του αντικτύπου της δημοσιογραφίας δεδομένων.

Αναλύσεις δεδομένων. Οι μετρήσεις των μέσων μπορούν να χρησιμοποιηθούν ως αντιπροσωπευτικά στοιχεία για τα επιθυμητά αποτελέσματα, όπως η αυξημένη ευαισθητοποίηση ή η αυξημένη γνώση. Ωστόσο, οι εταιρείες μέσων ενημέρωσης θα πρέπει να είναι συνειδητοποιημένες και προσεκτικές, όταν αποδίδουν αλλαγές στις αναλύσεις. Για παράδειγμα, εάν ένα έργο δημοσιογραφίας δεδομένων έχει ως στόχο να ενθαρρύνει τη θεσμική αλλαγή, οι μοναδικές

προβολές σελίδων δεν αποτελούν κατάλληλο μέτρο επιτυχίας· οι αναφορές των δεδομένων από δημόσιους λειτουργούς σε έγγραφα θα ήταν ένας καλύτερος δείκτης.

Πειραματική έρευνα. Η πειραματική έρευνα δημιουργεί σταθερές συνθήκες υπό τις οποίες μπορούν να δοκιμαστούν τα αποτελέσματα μιας παρέμβασης. Το Center for Media Engagement στο Πανεπιστήμιο του Τέξας στο Όστιν έχει διεξαγάγει συναρπαστική πειραματική έρευνα σχετικά με τις επιδράσεις του «στησίματος» της αρχικής σελίδας ειδήσεων στη μνήμη του κοινού και στην απήχησή τους, καθώς και των ρεπορτάζ που προσανατολίζονται στην εύρεση λύσεων, όσον αφορά στην απήχηση που έχουν στο κοινό των ειδησεογραφικών οργανισμών. Οι εταιρείες τεχνολογίας συνεχώς δοκιμάζουν τις επιδράσεις των διαφόρων διαδραστικών στοιχείων στους χρήστες. Οι δημοσιογραφικοί οργανισμοί μπορούν να κάνουν το ίδιο, για να κατανοήσουν καλύτερα τις επιπτώσεις των διαδραστικών δεδομένων στους χρήστες, είτε σε συνεργασία με πανεπιστήμια είτε συνεργαζόμενοι απευθείας με εσωτερικούς ερευνητές από τομείς όπως το μάρκετινγκ, η επιχειρηματική ανάπτυξη και η αλληλεπίδραση με το κοινό.

Έρευνες. Οι έρευνες, αν και δεν αποτελούν την πιο πρωτοποριακή ερευνητική μέθοδο, είναι ένας αποδεδειγμένος τρόπος συλλογής πληροφοριών από τους ανθρώπους σχετικά με τις αλλαγές στα ενδιαφέροντα, τις γνώσεις, τις απόψεις και τη δράση τους. Οι οργανισμοί μπορούν να είναι δημιουργικοί στον σχεδιασμό ερευνών, κάνοντας χρήση της τεχνολογίας η οποία επιτρέπει χαρακτηριστικά όπως αναδυόμενα παράθυρα που ενεργοποιούνται με την επιστροφή των επισκεπτών ή την καταγραφή των κλικ στα ενημερωτικά δελτία για τη δημιουργία μιας δεξαμενής πιθανών συμμετεχόντων στην έρευνα.

Ανάλυση περιεχομένου. Η ανάλυση περιεχομένου είναι μια ερευνητική μέθοδος που χρησιμοποιείται για τον προσδιορισμό των αλλαγών στον δημόσιο λόγο, με την πάροδο του χρόνου. Η μέθοδος αυτή μπορεί να εφαρμοστεί σε οποιοδήποτε σώμα κειμένων, γεγονός που την καθιστά εξαιρετικά ευέλικτη. Για παράδειγμα, όταν ένας οργανισμός παράγει περιεχόμενο με στόχο τον επηρεασμό του εθνικού δημόσιου λόγου, θα μπορούσε να διεξαγάγει μια εκ των υστέρων ανάλυση περιεχομένου στις δέκα μεγαλύτερες εθνικές εφημερίδες, για να προσδιορίσει την επιρροή των ιστοριών του. Εάν ο στόχος είναι να επηρεαστεί η κρατική νομοθεσία, ένας οργανισμός μπορεί να χρησιμοποιήσει ανάλυση περιεχομένου μετά το έργο σε δημόσια διαθέσιμες νομοθετικές ατζέντες (Green-Barber, 2015a). Ή, αν ο στόχος είναι να διατεθούν δεδομένα σε δίκτυα δράσης, η ανάλυση περιεχομένου μετά το έργο θα μπορούσε να χρησιμοποιηθεί για την ανάλυση των ενημερωτικών δελτίων ενός οργανισμού.

Η ανάλυση περιεχομένου μπορεί να πραγματοποιηθεί με τουλάχιστον τρεις τρόπους. Στο πιο βασικό επίπεδο, ένας ειδησεογραφικός οργανισμός μπορεί να αναζητήσει τις παραπομπές σε ένα έργο, προκειμένου να τεκμηριώσει πού και πότε έχει γίνει παραπομπή σε αυτό. Για παράδειγμα, πολλοί δημοσιογράφοι δημιουργούν ειδοποιήσεις ειδήσεων στο Google, χρησιμοποιώντας μια λέξη-κλειδί από το ρεπορτάζ τους, μαζί με το επώνυμό τους, προκειμένου να προσδιορίσουν από ποια άλλα μέσα προβάλλεται ένα έργο. Αυτό δεν είναι μεθοδολογικά ασφαλές, αλλά παρέχει ενδιαφέρουσες πληροφορίες και μπορεί να χρησιμοποιηθεί για τον έλεγχο του αντικτύπου. Η διαδικασία αυτή μπορεί, επίσης, να δημιουργήσει πρόσθετα ερωτήματα σχετικά με τον αντίκτυπο ενός έργου, τα οποία αξίζουν μια εμβριθέστερη μελέτη. Για τον σκοπό αυτό, πολλοί οργανισμοί χρησιμοποιούν υπηρεσίες συγκέντρωσης αποδελτίωσης ειδήσεων, όπως το Google News Alerts ή το Meltwater.

Η σχολαστική ανάλυση περιεχομένου θα εντοπίσει λέξεις-κλειδιά, δεδομένα ή/και φράσεις σε ένα έργο και, στη συνέχεια, θα αναλύσει τη συχνότητα εμφάνισής τους πριν και μετά τη δημοσίευση σε ένα πεπερασμένο σώμα κειμένων, για να τεκμηριώσει τη μεταβολή. Η υπολογιστική ανάλυση κειμένου πηγαίνει ένα βήμα παραπέρα και αντλεί συμπεράσματα για τις μεταβολές στον δημόσιο λόγο με προηγμένες τεχνικές καταμέτρησης και ανάλυσης. Αυτές οι πιο αυστηρές μέθοδοι ανάλυσης περιεχομένου απαιτούν από έναν ειδησεογραφικό οργανισμό να συνεργαστεί με εκπαιδευμένους ερευνητές.

Κοιτάζοντας μπροστά: Γιατί οι δημοσιογράφοι πρέπει να ενδιαφέρονται για τον αντίκτυπο της δημοσιογραφίας δεδομένων

Για να παραμείνει επίκαιρη, η δημοσιογραφία πρέπει όχι μόνο να αποδεχτεί ότι έχει αντίκτυπο στην κοινωνία, αλλά και να τον αξιοποιήσει. Δουλεύοντας για την κατανόηση του οικοσυστήματος της αλλαγής στο οποίο λειτουργεί η δημοσιογραφία και του συγκεκριμένου ρόλου της στο σύστημα, ο κλάδος μπορεί να μεγιστοποιήσει τον θετικό αντίκτυπό του και να αποδείξει την αξία του στο κοινό.

Οι δημοσιογράφοι δεδομένων, με την κατανόηση της αξίας και της σημασίας τόσο των ποσοτικών όσο και των ποιοτικών δεδομένων, βρίσκονται σε πλεονεκτική θέση απέναντι σε αυτήν την προσπάθεια. Με τη διατύπωση των στόχων των έργων δημοσιογραφίας δεδομένων, την ανάπτυξη δημιουργικών στρατηγικών αλληλεπίδρασης και κατανομής του κοινού και την ενσωμάτωση εξελιγμένων μεθόδων μέτρησης της επιτυχίας σε αυτά τα έργα, οι δημοσιογράφοι μπορούν να ηγηθούν αυτού του κινήματος εκ των έσω.

Παραπομπές

Green-Barber, L. (2014) «Waves of change: The case of rape in the fields», *The Center for Investigative Reporting*. Διαθέσιμο σε: https://www.documentcloud.org/documents/1278731-waves-of-change-the-case-of-rape-in-the-fields.html (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Green-Barber, L. (2015a) «Changing the conversation: The VA backlog», *The Center for Investigative Reporting*. Διαθέσιμο σε: https://s3.amazonaws.com/uploads-cironline-org/uploaded/uploads/VA+Backlog+White+Paper+11.10.14.pdf (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Green-Barber, L. (2015b) «What makes a news interactive successful? Preliminary lessons from the Center for Investigative Reporting», *The Center for Investigative Reporting*. Διαθέσιμο σε: https://s3-us-west-2.amazonaws.com/revealnews.org/uploads/CIR+News+Interactives+White+Paper.pdf (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Groseclose, T., & Milyo, J. (2005) «A measure of media bias», *The Quarterly Journal of Economics*, 120(4), σσ. 1191–1237. DOI: https://doi.org/10.1162/003355305775097542

Hamilton, J. T. (2004). *All the news that's fit to sell: How the market transforms information into news*. Πρίνστον, Νιου Τζέρσεϊ: Princeton University Press.

Hamilton, J. T. (2016). *Democracy's detectives: The economics of investigative journalism*. Κέιμπριτζ, Ηνωμένο Βασίλειο: Harvard University Press.

Klein, S. (2012) «News apps at ProPublica», στο J. Gray, L. Chambers, & L. Bounegru (επιμ.), *The data journalism handbook: How journalists can use data to improve the news*, σσ. 185-186. Σεβαστούπολη, Καλιφόρνια: O'Reilly Media.

Pitt, F., & Green-Barber, L. (2017) «The case for media impact: A case study of ICIJ's radical collaboration strategy», *Tow Center for Digital Journalism*. DOI: https://doi.org/10.7916/D85Q532V

Η συγγραφέας

Η Λίντσι Γκριν-Μπάρμπερ είναι η ιδρύτρια της Impact Architects, μιας συμβουλευτικής εταιρείας αφιερωμένης στην ανάπτυξη στρατηγικών δημοσιογραφίας, μέσων και επικοινωνίας υψηλού αντικτύπου και στη χρήση αυστηρών μεθόδων έρευνας των κοινωνικών επιστημών για την κατανόηση του αντικτύπου αυτού του έργου.

9.4
Δημοσιογραφία Δεδομένων: Προς το Συμφέρον Τίνος;

Mary Lynn Young, Candis Callison

Περίληψη: Σε αυτό το κεφάλαιο διατυπώνεται το ερώτημα τίνος τα συμφέροντα εξυπηρετούνται από τα έργα δημοσιογραφίας δεδομένων και εξετάζονται τα φανταστικά ακροατήρια, ιδίως όσον αφορά στην πρόσφατη δημοσιογραφία δεδομένων που σχετίζεται με το έγκλημα και υποτίθεται ότι εξυπηρετεί το δημόσιο καλό. Το κεφάλαιο αναφέρεται στο έργο αυτοχθόνων μελετητών, οι οποίοι υποστηρίζουν ότι η άρνηση, η παραποίηση, η αποικιοκρατία και η συλλογή δεδομένων αποτελούν μόνιμες προκλήσεις για τη δημοσιογραφία και απαιτούν καλύτερα ηθικά διαγνωστικά κριτήρια.

Λέξεις-κλειδιά: αποικιοκρατία, αυτόχθονες, δημοσιογραφία δεδομένων, περιεχόμενο εγκληματικότητας, ηθική, σπουδές επιστήμης και τεχνολογίας

Μια από τις πρώτες σημαντικές συνεισφορές στη δημοσιογραφία δεδομένων στις Ηνωμένες Πολιτείες ήταν το chicagocrime.org, ένας διαδικτυακός χάρτης του Σικάγο με στατιστικά στοιχεία για την εγκληματικότητα (Anderson, 2018· Holovaty, 2005· Holovaty, 2008). Σύμφωνα με τον ιδρυτή του, Άντριαν Χολοβάτι, το chicagocrime.org, το οποίο ξεκίνησε το 2005:

> ήταν μία από τις πρώτες mashup[238] εφαρμογές χαρτών, που συνδύαζε δεδομένα εγκλημάτων από το αστυνομικό τμήμα του Σικάγο με τους χάρτες της Google. Πρόσφερε μια σελίδα και ροή RSS για κάθε οικοδομικό τετράγωνο στο Σικάγο και μια πληθώρα τρόπων περιήγησης σε δεδομένα εγκλημάτων –ανά τύπο, ανά κατηγορία τοποθεσίας (π.χ. πεζοδρόμιο ή διαμέρισμα), ανά ταχυδρομικό κώδικα, ανά δρόμο/διεύθυνση, ανά ημερομηνία, ακόμη και με βάση μια αυθαίρετη διαδρομή (Holovaty, 2008)[239].

Λίγα χρόνια αργότερα, οι *Los Angeles Times* εγκαινίασαν το δημοσιογραφικό ιστολόγιο *Homicide Report*, το οποίο χρησιμοποιούσε τα δεδομένα της αστυνομίας, για να δημιουργήσει δημοσιεύσεις ανθρωποκτονιών σε ιστολόγια για καθεμία από τις περισσότερες από 900 ανθρωποκτονίες στην κομητεία. Και τα δύο πρότζεκτ χρησιμοποίησαν δεδομένα εγκληματικότητας και γεωγραφικά δεδομένα σε μεγάλα μητροπολιτικά κέντρα των ΗΠΑ. Και τα δύο προσφέρουν πληροφορίες σχετικά με τις επίμονες κριτικές και προκλήσεις που συνδέονται με τους στόχους και τις επιπτώσεις της δημοσιογραφίας με γνώμονα τα δεδομένα –και της δημοσιογραφίας γενικότερα.

Τα κίνητρα του Χολοβάτι για τη δημιουργία του chicagocrime.org ήταν εναρμονισμένα με τους στόχους της δημοσιογραφίας για την παραγωγή τοπικών «χρήσιμων ειδήσεων», σε συνδυασμό με την ολοένα και πιο τεχνική της ταυτότητα και την επικέντρωσή της σε «ενδιαφέροντα τεχνικά πράγματα» (Holovaty, 2005). Οι στόχοι της ιδρύτριας του *Homicide Report*, της δημο-

238. ΣτΕ: Που συνδυάζουν δύο ή περισσότερες πηγές δεδομένων.

239. Ήταν μια πρώιμη επανάληψη του ένδοξου ιστοτόπου δημοσιογραφίας δεδομένων της κοινότητας, *EveryBlock*, ο οποίος δημιουργήθηκε από τον Χολοβάτι το 2008 και εξαγοράστηκε από το MSNBC.com το 2009 [Holovaty, A. (2013) «RIP EveryBlock», Adrian Holovaty, 7 Φεβρουαρίου. Διαθέσιμο σε: http://www.holovaty.com/writing/rip-everyblock/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021].

σιογράφου των *Los Angeles Times*, Τζιλ Λεόβι, ήταν πιο καίριας σημασίας. Η Λεόβι θέλησε να καταγράψει όλες τις ανθρωποκτονίες στην κομητεία του Λος Άντζελες, προκειμένου να αποδομήσει τα παραδοσιακά δημοσιογραφικά πρότυπα και πρακτικές που κάλυπταν μόνο ορισμένες ανθρωποκτονίες (Leovy, 2008· Young & Hermida, 2015)[240]. Σε μια συνέντευξη που έδωσε το 2015 στην εκπομπή *Fresh Air* του National Public Radio, η Λεόβι εξέφρασε τα κίνητρά της για τη δημιουργία του *Homicide Report* ως απάντηση στη δομική προκατάληψη των ειδήσεων και την απογοήτευσή της για το γεγονός ότι οι δημοσιεύσεις των εφημερίδων για το έγκλημα «ωχριούν τόσο πολύ μπροστά στην πραγματικότητα»:

> Η δουλειά της εφημερίδας είναι να καλύπτει ασυνήθιστα γεγονότα, και όταν πρόκειται για ανθρωποκτονίες, αυτό πάντα καταλήγει να σημαίνει ότι καλύπτεις τα πολύ σπάνια περιστατικά. Και δεν ασχολείσαι ποτέ με το «κύριο σώμα», διότι αυτό αποτελεί σιωπηρά τις ανθρωποκτονίες ρουτίνας –αν και, φυσικά, μια ανθρωποκτονία δεν είναι ποτέ ρουτίνα. Αυτές οι ανθρωποκτονίες συνεχίζονται με την ίδια μορφή, με τους ίδιους τρόπους, για τόσο πολύ καιρό στην Αμερική, ιδίως στις αμερικανικές πόλεις που αποτελούν το φόντο της αστικής ζωής. Θεωρούνται δεδομένες και είναι πολύ δύσκολο να τις μετατρέψεις σε μια αφήγηση και μια ιστορία που να λειτουργεί για μια εφημερίδα (Leovy, 2015).

Συνδυάζοντας την εμπειρία της ως δημοσιογράφος αστυνομικού ρεπορτάζ με τον ατελείωτο, λιγότερο ιεραρχικό χώρο της ψηφιακής δημοσιογραφίας (σε σύγκριση με το πρωτοσέλιδο μιας εφημερίδας) και την πρόσβαση σε δημόσια δεδομένα, η Λεόβι (2008) οραματίστηκε ένα ειδησεογραφικό ρεπορτάζ που θα παρουσίαζε πληροφορίες για όλες τις δολοφονίες στην κομητεία, «κυρίως νεαρών Λατίνων και, δυσανάλογα, νεαρών μαύρων ανδρών», με όσο το δυνατόν μεγαλύτερη ισότητα (Leovy, 2015). Σύμφωνα με την ίδια (Leovy, 2008), η ανταπόκριση ήταν μεγάλη: «Οι αναγνώστες ανταποκρίθηκαν δυναμικά. "Θεέ μου", άρχιζε μία από τις πρώτες αναρτήσεις ενός αναγνώστη. "Και μόνο ο αριθμός είναι σοκαριστικός", έγραψε ένας άλλος. "Σχεδόν σαν να είναι αναλώσιμοι άνθρωποι", έγραψε ένας τρίτος».

Ως νέες εκφράσεις μιας αναπτυσσόμενης επιμέρους ειδικότητας, αυτά τα παραδείγματα δημοσιογραφίας δεδομένων έλαβαν τα εύσημα και αναγνωρίστηκαν για την καινοτομία τους. Ο ιστότοπος chicagocrime.org, σύμφωνα με τον Χολοβάτι (2008), αποτέλεσε ακόμη και έκθεμα στο Μουσείο Μοντέρνας Τέχνης της Νέας Υόρκης (MoMA). Όμως, τα ερωτήματα σχετικά με το τίνος τα συμφέροντα εξυπηρετούνται και ποιο ήταν το υποτιθέμενο ακροατήριο αυτών των προσωπικών έργων (και άλλων που ισχυρίζονται ότι μοιράζονται δεδομένα προς όφελος του δημόσιου συμφέροντος) αποσιωπήθηκαν σε μεγάλο βαθμό.

Οι ερευνητές των επιστημονικών και τεχνολογικών σπουδών έχουν επανειλημμένα καταδείξει πώς οι επιβλαβείς σχέσεις μεταξύ ευάλωτων πληθυσμών και ορισμένων ειδών δεδομένων μπορούν και εξακολουθούν να υφίστανται ακόμη και όταν η τεχνολογία εξαγγέλλεται ως νέα και μεταρρυθμιστική (Nelson, 2016· Reardon, 2009· TallBear, 2013).

Ο θετικιστικός προσανατολισμός της δημοσιογραφίας δεδομένων (Coddington, 2019) ενοχοποιείται, επίσης, παρά την εκτεταμένη κριτική σχετικά με την κοινωνική κατασκευή της φυλής και τον ρόλο της τεχνολογίας στην αναπαραγωγή της λευκής υπεροχής (Benjamin, 2019· Noble, 2018). Επιπλέον, μελέτες σχετικά με τις δημοσιογραφικές αναπαραστάσεις, τους κανόνες, τις πρακτικές, τα οικονομικά και τις ειδήσεις για την εγκληματικότητα δείχνουν μια μακρά ιστορία φυλετικοποίησης, κοινωνικού ελέγχου, πρόκλησης δεινών και συνεχιζόμενης αποικιοκρατίας (Anderson & Robertson, 2011· Callison & Young, 2020· Ericson κ.α., 1991· Hamilton, 2000· Schudson, 2005· Stabile, 2006).

Αυτό το κεφάλαιο διερευνά εν συντομία ποιες δομές υποστηρίζονται και τίνος τα δεδομένα είναι πιο πιθανό να συγκεντρωθούν –ή όχι–, ενώ θέτει ερωτήματα σχετικά με την ανάγκη των δημοσιογράφων να μπορούν να ενσωματώσουν μια «δεοντολογία της άρνησης», όταν απο-

240. Αργότερα επανασχεδιάστηκε ως ιστολόγιο αλγοριθμικής δημοσιογραφίας.

φασίζουν εάν και πώς θα χρησιμοποιήσουν τη δημοσιογραφία δεδομένων (Simpson, 2014· TallBear, 2013). Όπως υποστηρίζει η Μπάτλερ, «υπάρχουν τρόποι κατανομής της τρωτότητας, διαφοροποιημένες μορφές κατανομής που καθιστούν ορισμένους πληθυσμούς περισσότερο υποκείμενους σε αυθαίρετη βία από άλλους» (Butler, 2004).

Βασιζόμαστε στον Κόντινγκτον (Coddington, 2014· Coddington, 2019) για τον ορισμό της δημοσιογραφίας δεδομένων που χρησιμοποιούμε, ως ποσοτικής δημοσιογραφίας που έχει τρεις μορφές: δημοσιογραφία δεδομένων, υπολογιστική δημοσιογραφία («computational journalism») και ρεπορτάζ υποβοηθούμενο από υπολογιστή («computer-assisted reporting»). Οι επίμονες κριτικές των επιστημονικών πρακτικών και των κοινωνικών θεσμών που συνδυάζουν τις λογικές της έρευνας και της συλλογής δεδομένων, τις νέες τεχνολογίες και τα κοινωνικά ζητήματα είναι σχετικές και με τις τρεις μορφές δημοσιογραφίας δεδομένων, όπως και τα ερωτήματα σχετικά με την τρωτότητα και τίνος τα συμφέροντα υποστηρίζονται.

Η Τολμπέαρ, ερευνήτρια των επιστημονικών και τεχνολογικών σπουδών και μελετήτρια των αυτοχθόνων, εξέτασε τη γονιδιωματική έρευνα μεταξύ των αυτοχθόνων πληθυσμών στις ΗΠΑ και διαπίστωσε ότι πολλά από τα στερεότυπα και τις αποικιοκρατικές αφηγήσεις που συνδέονται με την έννοια των «Ινδιάνων που εξαφανίζονται» αποτελούσαν μέρος του σκεπτικού για την έρευνα, παράλληλα με τις δηλώσεις σχετικά με την πιθανή ταυτότητα (π.χ. γνώση για τη μετανάστευση και τη συνδετικότητα των προγόνων) και τις ωφέλειες για την υγεία. Επισημαίνει ότι:

> Ενώ η έννοια της γενετικής συνδετικότητας μπορεί να έχει αντικαταστήσει την έννοια της φυλετικής ιεραρχίας στο λεξικό της επικρατούσας επιστήμης, οι σχέσεις εξουσίας, διαχωρισμού και ιεραρχίας παραμένουν αναπόσπαστο μέρος της ευρύτερης κουλτούρας μας, των θεσμών και των δομών μας, καθώς και της κουλτούρας στην οποία διεξάγεται η επιστήμη και την οποία παράγει η επιστήμη (TallBear, 2007, σ. 413).

Αυτό που υποστηρίζει η Τολμπέαρ είναι ότι πίσω από τις επιστημονικές έννοιες της μηχανιστικής ή εργαστηριακής αντικειμενικότητας κρύβονται θεσμικά προνόμια, ιστορικές και συνεχείς σχέσεις με τις κοινότητες και πολιτισμικά πλαίσια που καθορίζουν τόσο τις λογικές της έρευνας όσο και τις διατυπώσεις των επιδιωκόμενων ωφελειών. Τα ερωτήματα «προς τίνος το συμφέρον και γιατί» πρέπει πάντα να τίθενται –και σε ορισμένες περιπτώσεις, οι αναλύσεις έρευνας ή/και εξόρυξης δεδομένων αξίζει να απορρίπτονται, επειδή οι διαδικασίες παραγωγής νοήματος βασίζονται σε παγιωμένες αντιλήψεις για τη φυλή, το φύλο και την τάξη. Αυτό που η Τολμπέαρ αποκαλεί «οι αποικιοκρατικές υποθέσεις και πρακτικές που συνεχίζουν να επηρεάζουν την επιστήμη», θα λέγαμε ότι επηρεάζει παρομοίως και τη δημοσιογραφία και, κατ' επέκταση, τη δημοσιογραφία δεδομένων (TallBear, 2007, σ. 423).

Οι αυτόχθονες πληθυσμοί έχουν, επίσης, γίνει αντικείμενο του περιεχομένου εκτεταμένων ανθρωπολογικών και κυβερνητικών αρχείων και συνεχών διαστρεβλώσεων και στερεοτύπων από τα μέσα ενημέρωσης (Anderson & Robertson, 2011), στην υπηρεσία ποικίλων μορφών και ιστοριών της αποικιοκρατίας των εποίκων (Tuck & Yang, 2012· Wolfe, 2006). Ως εκ τούτου, το διακύβευμα για τη δημοσιογραφία δεδομένων, ειδικά ως επέκταση των εννοιών της αντικειμενικότητας που βασίζεται σε μηχανές, είναι μεγάλο. Στην κριτική για την ανθρωπολογία (Simpson, 2014), η συγγραφέας προτείνει ότι οι κοινότητες των Μοχόκ επιδίδονται τακτικά σε μορφές άρνησης, όταν πρόκειται τόσο για την αντιμετώπιση τέτοιων αρχείων όσο και για τη συμμετοχή σε έρευνες που επικεντρώνονται σε αποικιοκρατικούς θεσμούς και πλαίσια. Η άρνηση στο πλαίσιο της Σίμπσον έχει πολλαπλές κατευθύνσεις: είναι η άρνηση να εξαλειφθείς, η άρνηση να εσωτερικεύσεις τις παραποιήσεις της ταυτότητας, του πολιτισμού και των τόπων σου και η άρνηση να συμμορφωθείς με τις προσδοκίες της διαφορετικότητας, έτσι ώστε να δεχτεί να σε αναγνωρίσει, εσένα ή την κοινότητά σου, το κράτος ή κάποιος άλλος φορέας (σε αυτήν την περίπτωση, προσθέτουμε τα μέσα ενημέρωσης).

Τέτοια επιχειρήματα από αυτόχθονες μελετητές θέτουν άμεσες προκλήσεις ως προς τις προθέσεις, τις λογικές και τις πρακτικές της δημοσιογραφίας δεδομένων, καθώς τοποθετούν στο

επίκεντρο ζητήματα ιστορίας και εξουσίας. Τα ερωτήματα αυτά δεν αφορούν μόνο το κράτος, αλλά και τον ρόλο της δημοσιογραφίας στη διατήρηση κοινωνικών προσταγμάτων που υποστηρίζουν τους κρατικούς σκοπούς, και στόχους, και δομές, και ιδεολογίες, όπως η πατριαρχία, η αποικιοκρατία των εποίκων και η λευκή υπεροχή (Callison & Young, 2020).

Μια περαιτέρω επιπλοκή για τις κοινότητες των αυτοχθόνων είναι ότι τόσο τα δεδομένα όσο και οι ακριβείς αναπαραστάσεις των μέσων ενημέρωσης είναι σχεδόν πάντα δύσκολο να εντοπιστούν –καθώς και το γεγονός ότι τα δεδομένα αντικατοπτρίζουν το θεσμικό πλαίσιο στο οποίο συλλέγονται, αρχειοθετούνται και είναι προσβάσιμα[241]. Η κριτική του Έρικσον για τις αστυνομικές στατιστικές που δεν αντικατοπτρίζουν την κοινωνική πραγματικότητα του εγκλήματος αλλά μάλλον τις «πολιτιστικές, νομικές και κοινωνικές κατασκευές που παράγονται [...] για οργανωτικούς σκοπούς» (Ericson, 1998, σ. 93) είναι σημαντική για τους δημοσιογράφους που επικεντρώνονται αποκλειστικά στη διαμάχη για τα δεδομένα. Για παράδειγμα, η δημοσιογράφος Τζένι Μονέ από το Λαγκούνα Πουέμπλο χαρακτηρίζει τις κοινότητες των αυτοχθόνων στις Ηνωμένες Πολιτείες ως «έθνη με αστερίσκο», τα οποία είναι εκείνα για τα οποία δεν υπάρχουν στοιχεία (Monet, 2020). Ειδικά στην Αλάσκα, πολλοί πίνακες κοινωνικών δεδομένων έχουν έναν αστερίσκο που λέει ότι δεν υπάρχουν δεδομένα για τους αυτόχθονες της Αλάσκας. Τα ψηφιακά μέσα, όπως το Facebook, προσφέρουν ελπιδοφόρες εναλλακτικές πλατφόρμες που θα μπορούσαν να θεωρηθούν εργαλείο της δημοσιογραφίας για την ενασχόληση με το κοινό των αυτοχθόνων και τις ανησυχίες τους, αλλά και για τη δημιουργία ουσιαστικών και ακριβών αναπαραστάσεων που αντιμετωπίζουν τις διαρθρωτικές ανισότητες και τα κενά στα δεδομένα (Monet, 2018). Και πάλι, λοιπόν, το ερώτημα για το αν και πώς θα συμμετάσχουμε περιστρέφεται γύρω από το ποιος επωφελείται, ποιες διαδικασίες χρησιμοποιούνται στη συλλογή δεδομένων και ποιων οι διαδικασίες νοηματοδότησης επικρατούν.

Για τη δημοσιογραφία, σε γενικές γραμμές, οι διαδικασίες δημιουργίας νοήματος συνδέονται συχνά με ζητήματα διαφωνίας, απόκλισης, σύγκρουσης ή «τη συμπεριφορά ενός πράγματος ή προσώπου που ξεφεύγει από τον κανόνα» (Ericson, 1998, σ. 84), στο πλαίσιο ενός θετικιστικού προσανατολισμού. Ο ρόλος της δημοσιογραφίας στην κοινωνική ιεράρχηση είχε, και συνεχίζει να έχει, υλικές επιπτώσεις και επιβλαβείς συνέπειες για τους πληθυσμούς που κατασκευάζονται ως αποκλίνοντες (Callison & Young, 2020· Rhodes, 1993· Stabile, 2006). Η ιστορική μελέτη της Στάμπιλε σχετικά με τις ειδήσεις για την εγκληματικότητα στις Ηνωμένες Πολιτείες, η οποία περιλαμβάνει τις εφημερίδες, την τηλεοπτική κάλυψη της εγκληματικότητας και τα ραδιοφωνικά προγράμματα, καθώς και τη σχέση των ειδήσεων για την εγκληματικότητα με τη φυλή, διατυπώνει τον αντίκτυπο των προτύπων παρέκκλισης σε δομικά ευάλωτους πληθυσμούς μέσα σε ένα ιδεολογικό πλαίσιο λευκής υπεροχής και κερδοσκοπικής δημοσιογραφίας. Επικεντρώνεται στη φυλή και στο φύλο, καθώς «είναι από τους πιο σημαντικούς τομείς των αγώνων για την ιστορική σημασία που αποδίδεται στην παρέκκλιση» (Stabile, 2006, σ. 5), ισχυριζόμενη ότι τα μέσα μαζικής ενημέρωσης υποστηρίζουν τις «διαδικασίες ποινικοποίησης» των μαύρων ανδρών από το κράτος και τους πράκτορές του, όπως η αστυνομία (Stabile, 2006, σ. 3).

Ένα παράδειγμα είναι ο τρόπος με τον οποίο τα μέσα ενημέρωσης εντείνουν και ενισχύουν τις πρακτικές συλλογής δεδομένων της αστυνομίας, εστιάζοντας σε συγκεκριμένα εγκλήματα, όπως οι ληστείες αυτοκινήτων, οι δράστες και τα θύματα. Διαπιστώνει την ύπαρξη μιας «άπληστης και βίαιης λευκής κοινωνίας που άνθισε στις ΗΠΑ, στην οποία οι μυθοπλασίες της λευκής τρομοκράτησης εκτόπισαν σταθερά την υλικότητα της λευκής τρομοκρατίας» (Stabile, 2006, σ. 2). Εδώ η ανάλυση του Κάρεϊ ότι η δημοσιογραφία δημιουργεί εχθρούς και συμμάχους (Carey, 1974) μπορεί να κατανοηθεί ως σχετική και με τις θεσμικές σχέσεις του επαγγέλματος με τον

241. Για περισσότερα σχετικά με αυτά τα ζητήματα, βλ. το κεφάλαιο των Κουκουτάι και Γουόλτερ σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 3.3, «Κυριαρχία των Αυτοχθόνων επί των Δεδομένων τους: Επιπτώσεις για τη Δημοσιογραφία Δεδομένων»).

καπιταλισμό και το κράτος στη Βόρεια Αμερική, οι οποίες περιλαμβάνουν την κρατική γενοκτονία και τις συνεχιζόμενες αποικιοκρατίες. Σε συνδυασμό με την αλλεργία που επιδεικνύει η δημοσιογραφία στην ιδέα ότι τα γεγονότα και η γνώση είναι κοινωνικά κατασκευασμένα, η δημοσιογραφία –και ειδικότερα οι ειδήσεις– καθίσταται το περιτύλιγμα μέσω του οποίου οι συζητήσεις περί κοινωνικής ιεράρχησης έχουν συνδημιουργηθεί και συνδημιουργούνται, αναπαράγονται, αλλά και δυνητικά μετασχηματίζονται (Dumit & O'Connor, 2016).

Σε αυτά τα κρίσιμα σημεία, η βιβλιογραφία από τη δημοσιογραφία, την εγκληματολογία, τις επιστημονικές και τεχνολογικές σπουδές και άλλους κλάδους εγείρει μια σειρά από επείγοντες προβληματισμούς που δεν έχουν εξεταστεί επαρκώς όσον αφορά στη δημοσιογραφία δεδομένων. Οι μελετητές έχουν αφιερώσει περισσότερο χρόνο στις τυπολογίες (Coddington, 2019), στην κατάσταση της δημοσιογραφίας δεδομένων (Heravi, 2017) και στις επιπτώσεις της δημοσιογραφίας δεδομένων στις ευρύτερες δημοσιογραφικές επιστημολογίες, κουλτούρες, πρακτικές και ταυτότητες (Anderson, 2018· Borges-Rey, 2020· Gynnild, 2014· Lewis & Usher, 2014· Young & Hermida, 2015) παρά στις ευρύτερες επιπτώσεις και συνέπειές της. Ελάχιστοι μελετητές έχουν θέσει ζητήματα που σχετίζονται με την εξουσία –με εξαίρεση την έρευνα του Μπόρχες-Ρέι, ο οποίος ενσωματώνει μια ανάλυση πολιτικής οικονομίας της ανάπτυξης της δημοσιογραφίας δεδομένων στο Ηνωμένο Βασίλειο (Borges-Rey, 2016· Borges-Rey, 2020).

Ωστόσο, η δημοσιογραφία δεδομένων μπορεί να υποδείξει κάποιον αντίκτυπο, όπως συμβαίνει σε αυτή τη δήλωση του Χολοβάτι:

> Πολλά καλά βγήκαν από το chicagocrime.org. Σε τοπικό επίπεδο, αμέτρητοι κάτοικοι του Σικάγο επικοινώνησαν μαζί μου, για να εκφράσουν τις ευχαριστίες τους για αυτή τη δημόσια υπηρεσία. Ομάδες της κοινότητας πήραν μαζί τους εκτυπωμένα αντίγραφα του ιστοτόπου στις συνεδριάσεις της αστυνομίας και παθιασμένοι πολίτες πήγαν τα ρεπορτάζ του ιστοτόπου στους δημοτικούς συμβούλους τους, για να επισημάνουν προβληματικούς κόμβους, όπου η πόλη θα μπορούσε να εξετάσει το ενδεχόμενο να εγκαταστήσει φωτεινότερους φανούς στους δρόμους (Holovaty, 2008).

Στην περίπτωση αυτή, οι ομάδες της κοινότητας πήραν τα δεδομένα και έδωσαν το δικό τους νόημα και το δικό τους σκεπτικό για δράση. Όμως, ο τρόπος που αυτό λειτουργεί σε ευρύτερη κλίμακα, σε αγροτικές περιοχές μακριά από τα κέντρα εξουσίας και τα μέσα ενημέρωσης, σε κοινότητες που ήδη παρακολουθούνται δυσανάλογα και σε περιπτώσεις όπου οι κοινότητες δεν εκπροσωπούνται επαρκώς στις αίθουσες σύνταξης (οι οποίες παραμένουν κυρίως λευκές τόσο στον Καναδά όσο και στις Ηνωμένες Πολιτείες) απαιτεί ένα ευρύτερο σύνολο δεοντολογίας (Callison & Young, 2020). Λαμβάνοντας υπόψη αυτά τα παραδείγματα και στοιχεία από κριτικές βιβλιογραφίες εκτός των δημοσιογραφικών σπουδών, η πιθανή ζημία θα μπορούσε και θα έπρεπε να έχει προτεραιότητα απέναντι σε νόρμες όπως «ειδήσεις που μπορείτε να χρησιμοποιήσετε» και σε τεχνολογικά τροφοδοτούμενους πειραματισμούς. Ο τρόπος με τον οποίο οι δημοσιογράφοι καλύπτουν την εγκληματικότητα από την οπτική των δεδομένων απαιτεί βαθιά κατανόηση των συνεπειών, καθώς και των προβλημάτων από την εξέταση προθέσεων, που είναι μόνο εσωτερικές στη δημοσιογραφία –αποδείξεις επιτυχίας και λογικές καινοτομίας[242]. Η δεοντολογία πρέπει να λαμβάνει καλύτερα υπόψη την έννοια της άρνησης, τη μακρά ιστορία της λανθασμένης απεικόνισης και της εξυπηρέτησης της αποικιοκρατίας από τη δημοσιογραφία, καθώς και τις άνισες διαδικασίες με τις οποίες συντελείται η παραγωγή νοήματος και η συλλογή δεδομένων. Τα «προς το συμφέρον τίνος και γιατί» γίνονται τα βασικά ερωτήματα για τους δημοσιογράφους, όταν εξετάζουν πώς, πού και σε ποιον προσφέρει η δημοσιογραφία δεδομένων.

242. Βλ. το κεφάλαιο της Λούζεν σε αυτόν τον τόμο (ΣτΕ: Κεφάλαιο 9.2, «Χρυσά Πρότυπα Με Γνώμονα τα Δεδομένα: Τι Αποτιμά ο Κλάδος ως Δημοσιογραφία Δεδομένων που αξίζει Βραβείο και Πώς η Δημοσιογραφία Δεδομένων Εξελίσσεται μαζί με τη Δεδομενοποίηση της Κοινωνίας»).

Παραπομπές

Anderson, C. W. (2018). Apostles of certainty: Data journalism and the politics of doubt. Νέα Υόρκη: Oxford University Press.

Anderson, M. C., & Robertson, C. L. (2011). Seeing red: A history of natives in Canadian newspapers. Γουίνιπεγκ, Μανιτόμπα, Καναδάς: University of Manitoba Press.

Benjamin, R. (2019). Race after technology: *Abolitionist tools for the New Jim Code*. Κέιμπριτζ, Ηνωμένο Βασίλειο: Polity Press.

Borges-Rey, E. (2016) «Unravelling data journalism: A study of data journalism practice in British newsrooms», *Journalism Practice,* 10(7), σσ. 833–843. DOI: https://doi.org/10.1080/17512786.2016.1159921

Borges-Rey, E. (2020) «Towards an epistemology of data journalism in the de-volved nations of the United Kingdom: Changes and continuities in materiality, performativity and reflexivity», *Journalism*, 21(7), σσ. 915–932. DOI: https://doi.org/10.1177/1464884917693864

Butler, J. (2004). *Precarious life: The powers of mourning and violence.* Νέα Υόρκη: Verso.

Callison, C., & Young, M. L. (2020). *Reckoning: Journalism's limits and possibilities.* Νέα Υόρκη: Oxford University Press.

Carey, J. W. (1974) «The problem of journalism history», *Journalism History, 1*(1), σσ. 3–27. DOI: https://doi.org/10.1080/00947679.1974.12066714

Coddington, M. (2014) «Clarifying journalism's quantitative turn», *Digital Journalism*, 3, σσ. 331–348. DOI: https://doi.org/10.1080/21670811.2014.976400

Coddington, M. (2019) «Defining and mapping data journalism and computational journalism: A review of typologies and themes», στο S. Eldridge & B. Franklin (επιμ.), *The Routledge handbook of developments in digital journalism studies, σσ.* 225–236. Λονδίνο: Routledge. DOI: https://doi.org/10.4324/9781315270449-18

Dumit, J., & O'Connor, K. (2016) «The senses and sciences of fascia: A practice as research investigation», στο L. Hunter, E. Krimmer, & P. Lichtenfels (επιμ.), *Sentient performativities of embodiment: Thinking alongside the human*, σσ. 35–54. Λονδίνο: Rowman & Littlefield.

Ericson, R. (1998) «How journalists visualize fact», *The Annals of the American Academy of Political and Social Science,* 560, σσ. 83–95. DOI: https://doi.org/10.1177/0002716298560001007

Ericson, R., Baranek, P. M., & Chan, J. B. L. (1991). *Representing order: Crime, law, and justice in the news media*. Τορόντο: University of Toronto Press.

Gynnild, A. (2014) «Journalism innovation leads to innovation journalism: The impact of computational exploration on changing mindsets», *Journalism*, 15(6), σσ. 713–730. DOI: https://doi.org/10.1177/1464884913486393

Hamilton, J. T. (2000). *Channeling violence: The economic market for violent television programming.* Πρίνστον, Νιου Τζέρσεϊ: Princeton University Press.

Heravi, B. (2017) *State of data journalism globally: First insights into the global data journalism survey* [Medium], 1 Αυγούστου. Διαθέσιμο σε: https://medium.com/ucd-ischool/state-of-data-journalism-globally-cb2f4696ad3d (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Holovaty, A. (2005) «*Announcing chicagocrime.org*», Adrian Holovaty, 18 Μάϊου. Διαθέσιμο σε: http://www.holovaty.com/writing/chicagocrime.org-launch/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Holovaty, A. (2008) «*In memory of chicagocrime.org*», Adrian Holovaty, 31 Ιανουαρίου. Διαθέσι-

μο σε: http://www.holovaty.com/writing/chicagocrime.org-tribute/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Leovy, J. (2008) «*Unlimited space for untold sorrow*», Los Angeles Times, 4 Φεβρουαρίου. Διαθέσιμο σε: https://www.latimes.com/archives/la-xpm-2008-feb-04-me-homicide4-story.html (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Leovy, J. (2015) «*"Ghettoside" explores why murders are invisible in Los Angeles* [interview with Dave Davies]», Fresh Air, 25 Ιανουαρίου. Διαθέσιμο σε: https://www.npr.org/2015/01/26/381589023/ghettoside-explores-why-murders-are-invisible-in-los-angeles (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Lewis, S. C., & Usher, N. (2014) «Code, collaboration, and the future of journalism: A case study of the hacks/hackers global network», *Digital Journalism,* 2(3), σσ. 383–393. DOI: https://doi.org/10.1080/21670811.2014.895504

Monet, J. (2020) «Native American voters could decide key Senate races while battling intense voter suppression», *Democracy Now!,* 30 Οκτωβρίου. Διαθέσιμο σε: https://www.democracynow.org/2020/10/30/jenni_monet_indigenous_sovereignty_2020 (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Monet, J. (2018b) «*#DeleteFacebook? Not in Indian Country*», Yes! Magazine, 23 Μαρτίου. Διαθέσιμο σε: https://www.yesmagazine.org/social-justice/2018/03/23/deletefacebook-not-in-indian-country (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Nelson, A. (2016). *The social life of DNA: Race, reparations, and reconciliation after the genome.* Βοστώνη: Beacon Press.

Noble, S. (2018). *Algorithms of oppression*. Νέα Υόρκη: NYU Press.

Reardon, J. (2009). *Race to the finish: Identity and governance in an age of genomics.* Πρίνστον, Νιου Τζέρσεϊ: Princeton University Press.

Rhodes, J. (1993) «The visibility of race and media history», *Critical Studies in Mass Communication,* 10(2), σσ. 184–190. DOI: https://doi.org/10.1080/15295039309366859

Schudson, M. (2005) «Autonomy from what?», στο R. Benson & E. Neveu (επιμ.), *Bourdieu and the journalistic field,* σσ. 214–223. Κέιμπριτζ, Ηνωμένο Βασίλειο: Polity Press.

Simpson, A. (2014). Mohawk interruptus: *Political life across the borders of settler states.* Ντάραμ, Ηνωμένο Βασίλειο: Duke University Press.

Simpson, A. (2016) «The state is a man: Theresa Spence, Loretta Saunders and the gender of settler sovereignty», *Theory & Event,* 19(4). Διαθέσιμο σε: https://muse.jhu.edu/article/633280 (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Stabile, C. (2006). *White victims, Black villains: Gender, race and crime news in US culture.* Νέα Υόρκη: Routledge.

TallBear, K. (2007) «Narratives of race and indigeneity in the genographic project», *The Journal of Law, Medicine & Ethics,* 35(3).

TallBear, K. (2013). *Native American DNA: Tribal belonging and the false promise of genetic science. Μινεάπολη:* University of Minnesota Press.

Tuck, E., & Yang, K. (2012) «*Decolonization is not a metaphor*», *Decolonization: Indigeneity, Education & Society,* 1(1), σσ. 1–40.

Wolfe, P. (2006) «*Settler colonialism and the elimination of the native*», *Journal of Genocide Research,* 8(4), σσ. 387–409. DOI: https://doi.org/10.1080/14623520601056240

Young, M. L., & Hermida, A. (2015) «From Mr. and Mrs. Outlier to central tendencies», *Digital Journalism*, 3(3), σσ. 381–397. DOI: https://doi.org/10.1080/21670811.2014.976409

Οι συγγραφείς

Η Μέρι Λιν Γιανγκ είναι Αναπληρώτρια Καθηγήτρια στη Σχολή Δημοσιογραφίας, Συγγραφής και Μέσων στο Πανεπιστήμιο της Βρετανικής Κολομβίας στον Καναδά (University of British Columbia, UBC).
Η Κάντις Κάλισον είναι Αναπληρώτρια Καθηγήτρια στη Σχολή Δημοσιογραφίας, Συγγραφής και Μέσων και στο Ινστιτούτο Κριτικών Αυτοχθόνων Σπουδών Πανεπιστήμιο της Βρετανικής Κολομβίας.

9.5
Δημοσιογραφία Δεδομένων με Αντίκτυπο

Paul Bradshaw

Περίληψη: Δημοσιογραφία δεδομένων με αντίκτυπο: πώς και γιατί μετριέται ο αντίκτυπος, πώς αυτό έχει αλλάξει, και ποιοι παράγοντες διαμορφώνουν τον αντίκτυπο.

Λέξεις-κλειδιά: αντίκτυπος, αλληλεπίδραση, δημοσιογραφία δεδομένων, ανάλυση, ερευνητική δημοσιογραφία, ποιότητα δεδομένων

Εάν δεν έχετε δει την ταινία *Spotlight: Όλα στο Φως* (2015), με θέμα την έρευνα της εφημερίδας *Boston Globe* για τη σιωπή των θεσμικών οργάνων σχετικά με την παιδική κακοποίηση, τότε πρέπει να τη δείτε αμέσως. Συγκεκριμένα, θα πρέπει να την παρακολουθήσετε μέχρι και τους τίτλους του τέλους[243].

Μια λίστα ξετυλίγεται στην οθόνη. Περιγράφει λεπτομερώς τα δεκάδες μέρη όπου έχουν αποκαλυφθεί σκάνδαλα κακοποίησης αφότου έλαβαν χώρα τα γεγονότα της ταινίας, από το Ακούτε της Νιγηρίας μέχρι το Γουόλονγκονγκ της Αυστραλίας. Βέβαια, οι τίτλοι μάς αναγκάζουν, επίσης, να διακόψουμε τους πανηγυρισμούς μας: ένα από τα πρόσωπα-κλειδιά που εμπλέκονται στο σκάνδαλο, σημειώνεται, πήρε μετάθεση για «μία από τις πιο σημαντικές ρωμαιοκαθολικές εκκλησίες στον κόσμο».

Αυτή είναι η πρόκληση του αντικτύπου στη δημοσιογραφία δεδομένων: η γνωστοποίηση ενός προβλήματος είναι «αντίκτυπος»; Πρέπει η ιστορία να οδηγήσει σε ποινή ή ανταμοιβή; Σε ορατή αλλαγή πολιτικής; Πόσο σημαντικός είναι ο αντίκτυπος; Και για ποιον;

Αυτά τα δύο τελευταία ερωτήματα αξίζει να απαντηθούν πρώτα. Παραδοσιακά, ο αντίκτυπος ήταν σημαντικός για δύο βασικούς λόγους: εμπορικά και πολιτισμικά. Από εμπορική άποψη, οι μετρήσεις του αντικτύπου, όπως η αναγνωρισιμότητα της επωνυμίας και τα υψηλά ποσοστά κοινού, μπορούν να συμβάλλουν άμεσα στο περιθώριο κέρδους μιας έκδοσης μέσω της διαφήμισης (αυξάνοντας τόσο την τιμή όσο και τον όγκο) και των πωλήσεων συνδρομών/αντιτύπων (Rusbridger, 2018). Σε πολιτισμικό επίπεδο, ωστόσο, οι ιστορίες με αντίκτυπο έχουν, επίσης, εξασφαλίσει σε ειδησεογραφικούς οργανισμούς και μεμονωμένους δημοσιογράφους το «δικαίωμα να καυχώνται» μεταξύ των συναδέλφων τους. Και τα δύο, όπως θα δούμε, έχουν γίνει πιο περίπλοκα στις μέρες μας.

Οι μετρήσεις του αντικτύπου στη δημοσιογραφία ήταν, ιστορικά, περιορισμένες: οι εκδότες μπορούσαν να βασιστούν μόνο σε συγκεντρωτικά στοιχεία πωλήσεων και κοινού, σε μια περιορισμένη λίστα βραβείων του κλάδου και σε περιστασιακές έρευνες κοινού. Τώρα, βέβαια, η πρόκληση δεν έγκειται μόνο στον πολλαπλασιασμό των μετρήσεων αλλά και στον πολλαπλασιασμό των επιχειρηματικών μοντέλων, με την επέκταση της μη κερδοσκοπικής παροχής ειδήσεων ειδικότερα να οδηγεί σε μια αυξανόμενη έμφαση στον αντίκτυπο και στη συζήτηση για

243. https://www.imdb.com/title/tt1895587/

τη μέτρησή του (Schlemmer, 2016). Επιπλέον, η δυνατότητα μέτρησης του αντικτύπου για κάθε ιστορία ξεχωριστά σημαίνει ότι δεν είναι πλέον μόνο οι εκδότες υπεύθυνοι για τον αντίκτυπο στο κοινό αλλά και οι δημοσιογράφοι.

Η μέτρηση του αντικτύπου σε αριθμούς

Ίσως η πιο απλή μέτρηση του αντικτύπου να είναι η καθαρή απήχηση: διαδραστικά θέματα που βασίζονται σε δεδομένα, όπως το *7 Billion People and You: What's Your Number?* («7 δισεκατομμύρια άνθρωποι κι εσύ: Ποιος είναι ο αριθμός σου;») του BBC[244], έχουν απασχολήσει εκατομμύρια αναγνώστες με μια επίκαιρη ιστορία, ενώ κάποια στιγμή το 2012 η δημοσιογραφία δεδομένων του Νέιτ Σίλβερ έφτανε σε έναν στους πέντε επισκέπτες των *New York Times* (Byers, 2012).

Κάποιοι θα χλευάσουν τέτοιου είδους χονδροειδείς μετρήσεις, αλλά είναι σημαντικές. Εάν κάποτε οι δημοσιογράφοι επικρίνονταν ότι προσπαθούσαν να εντυπωσιάσουν τους συναδέλφους τους εις βάρος του κοινού τους, η σύγχρονη δημοσιογραφία αναμένεται τουλάχιστον να αποδείξει ότι μπορεί να συνδεθεί με αυτό το κοινό. Στις περισσότερες περιπτώσεις αυτή η απόδειξη είναι απαραίτητη για τους διαφημιστές, αλλά ακόμη και οι δημόσια χρηματοδοτούμενοι πάροχοι οικουμενικών ειδήσεων, όπως το BBC, τη χρειάζονται επίσης, για να αποδείξουν ότι πληρούν τις προϋποθέσεις χρηματοδότησης.

Η *αλληλεπίδραση* είναι η πιο εξελιγμένη σχέση της απήχησης και εδώ η δημοσιογραφία δεδομένων τα καταφέρνει επίσης καλά: σε ένα συνέδριο συντακτών για την εκδοτική εταιρεία εφημερίδων Reach, για παράδειγμα, αποκαλύφθηκε ότι η απλή προσθήκη μιας οπτικοποίησης δεδομένων σε μια σελίδα μπορεί να αυξήσει τον χρόνο παραμονής (τον χρόνο που περνάει ένα άτομο σε μια ιστοσελίδα) κατά το ένα τρίτο. Η βασισμένη στα δεδομένα διαδραστικότητα μπορεί να μεταμορφώσει και τα πιο βαρετά θέματα: το 2015, ο Ντέιβιντ Χίγκερσον της ίδιας εταιρείας παρατήρησε ότι περισσότεροι από 200.000 άνθρωποι εισήγαγαν τον ταχυδρομικό τους κώδικα σε ένα διαδραστικό widget της ομάδας δεδομένων της που βασίστηκε σε στατιστικά στοιχεία φτώχειας –ένας πολύ μεγαλύτερος αριθμός, όπως τόνισε, «από ό,τι θα φανταζόμουν [για] μια απλή ιστορία για την οποία "τα δεδομένα λένε κάτι"» (Higgerson, 2015).

Η αλληλεπίδραση είναι ιδιαίτερα σημαντική για τους οργανισμούς που βασίζονται στη διαφήμιση (τα ποσοστά μπορούν να αυξηθούν, όταν η αλληλεπίδραση είναι υψηλή) αλλά και για εκείνους για τους οποίους οι συνδρομές, οι δωρεές και οι εκδηλώσεις είναι σημαντικές: και αυτές τείνουν να συνδέονται με την αλληλεπίδραση.

Η επέκταση της μη κερδοσκοπικής χρηματοδότησης και των επιχορηγήσεων συχνά συνοδεύεται από τη ρητή απαίτηση παρακολούθησης ή απόδειξης του αντικτύπου, ο οποίος δεν αφορά μόνο την απήχηση. Συχνά γίνεται αναφορά στην αλλαγή και στη δράση, ιδίως την πολιτική ή τη νομική. Η Διεθνής Σύμπραξη Ερευνητών Δημοσιογράφων («International Consortium of Investigative Journalists, ICIJ»), για παράδειγμα, επισημαίνει τον αντίκτυπο της έρευνάς της για τα *Panama Papers* αναφερόμενη στο γεγονός ότι οδήγησε σε «τουλάχιστον 150 εξεταστικές διαδικασίες, ελέγχους ή έρευνες [...] σε 79 χώρες», παράλληλα με την πιο παραδοσιακή μέτρηση των σχεδόν 20 βραβείων, συμπεριλαμβανομένου του βραβείου Πούλιτζερ (Fitzgibbon & Díaz-Struck, 2016· ICIJ, άνευ ημερομηνίας). Στο Ηνωμένο Βασίλειο, ιδιαίτερη θέση στην ιστορία της δημοσιογραφίας δεδομένων κατέχει το σκάνδαλο των εξόδων των βουλευτών. Αυτό όχι μόνο οδήγησε την εφημερίδα *Telegraph* στην κορυφή της ειδησεογραφικής ατζέντας για εβδομάδες, αλλά οδήγησε, επίσης, στη δημιουργία ενός νέου φορέα: του Independent Parliamentary Standards Authority («Ανεξάρτητη Αρχή Κοινοβουλευτικών Προτύπων» –IPSA).

244. https://www.bbc.com/news/world-15391515

Ο φορέας αυτός πλέον δημοσιεύει ανοιχτά δεδομένα σχετικά με τις δηλώσεις δαπανών των πολιτικών, γεγονός που οδηγεί σε μεγαλύτερη λογοδοσία εκ μέρους τους και σε επιπλέον έργα δημοσιογραφίας δεδομένων.

Η χάραξη πολιτικής μπορεί, όμως, να είναι πολύ ευρύτερη από την πολιτικολογία. Οι δανειοδοτικές πολιτικές των τραπεζών επηρεάζουν εκατομμύρια ανθρώπους και έγιναν γνωστές στα τέλη της δεκαετίας του 1980 στις ΗΠΑ από τον Μπιλ Ντέντμαν στη σειρά άρθρων του *The Color of Money* («Το Χρώμα του Χρήματος»), που κέρδισε το βραβείο Πούλιτζερ. Εντοπίζοντας φυλετικά διαχωρισμένες πρακτικές δανειοδότησης[245], η έρευνα με γνώμονα τα δεδομένα οδήγησε, επίσης, σε πολιτικές, οικονομικές και νομικές αλλαγές, με έρευνες, νέες χρηματοδοτήσεις, αγωγές και την ψήφιση νέων νόμων ανάμεσα στις επακόλουθες ενέργειες[246].

Κάνοντας ένα άλμα 30 χρόνων προς τα εμπρός, βλέπουμε μια πολύ σύγχρονη εκδοχή αυτής της προσέγγισης: Η σειρά *Machine Bias* της *ProPublica* ρίχνει φως στην αλγοριθμική λογοδοσία, ενώ το *Bureau Local* αξιοποίησε το δίκτυό του, για να αντλήσει, μέσω πληθοπορισμού («crowdsourcing»), πληροφορίες σχετικά με τις αλγοριθμικά στοχευμένες «σκοτεινές διαφημίσεις» στα μέσα κοινωνικής δικτύωσης (McClenaghan, 2017).

Και οι δύο συνέβαλαν στην αλλαγή αρκετών πολιτικών του Facebook, ενώ οι μέθοδοι της *ProPublica* υιοθετήθηκαν από μια ομάδα για τη δίκαιη στέγαση, προκειμένου να δημιουργηθεί η βάση για μια αγωγή κατά του μέσου κοινωνικής δικτύωσης (Angwin & Tobin, 2018· About Facebook, 2017· Jalonick, 2017). Καθώς οι πολιτικές των αλγορίθμων γίνονται όλο και πιο επιδραστικές στη ζωή μας (από τον επηρεασμό της τοποθέτησης αστυνομικών δυνάμεων μέχρι την τιμολόγηση της Uber «σε μη λευκές περιοχές»), η απόδοση ευθυνών σε αυτές καθίσταται εξίσου σημαντική με την απόδοση ευθυνών σε πιο παραδοσιακές πολιτικές μορφές εξουσίας (Chammah, 2016· Stark, 2016).

Το αξιοσημείωτο σε ορισμένα από αυτά τα παραδείγματα είναι ότι ο αντίκτυπός τους εξαρτάται –και εν μέρει καταδεικνύεται– από τη συνεργασία με άλλους. Όταν το *Bureau Local* μιλάει για τον αντίκτυπο, για παράδειγμα, αναφέρεται στον αριθμό των ιστοριών που παράγονται από τα μέλη του δικτύου βάσης του, εμπνέοντας άλλους να αναλάβουν δράση, ενώ το ICIJ αναφέρει την αυξανόμενη κλίμακα των δικτύων του: «Στο *LuxLeaks* (2014) συμμετείχαν περισσότεροι από 80 δημοσιογράφοι σε 26 χώρες. Στο *Swiss Leaks* (2015), περισσότεροι από 140 δημοσιογράφοι σε 45 χώρες» (Cabra, 2017). Ο αριθμός αυτός αυξάνεται σε περισσότερους από 370 δημοσιογράφους σε σχεδόν 80 χώρες για την έρευνα των *Panama Papers*: 100 ενημερωτικοί οργανισμοί δημοσίευσαν 4.700 άρθρα (Blau, 2016).

Επιπλέον, τα δεδομένα που συγκεντρώνονται και δημοσιεύονται ως αποτέλεσμα ερευνών μπορούν να αποτελέσουν από μόνα τους μια πηγή αντικτύπου: η βάση δεδομένων των *Offshore Leaks*, επισημαίνει το ICIJ, «χρησιμοποιείται τακτικά από ακαδημαϊκούς, ΜΚΟ και φορολογικές υπηρεσίες» (Cabra, 2017).

Υπάρχει κάτι αξιοσημείωτο σε αυτή τη μετατόπιση από το καμάρι της δημοσίευσης στο να κερδίζει κανείς επαίνους για τη δράση του ως διαμεσολαβητής, διοργανωτής και διαχειριστής βάσεων δεδομένων. Ως αποτέλεσμα, η συνεργασία έχει καταστεί μια δεξιότητα από μόνη της: πολλοί μη κερδοσκοπικοί οργανισμοί έχουν αναλάβει ρόλους διαχείρισης κοινότητας ή έργου αφιερωμένους στην οικοδόμηση και διατήρηση σχέσεων με συνεργάτες και εταίρους και η εκπαίδευση στη δημοσιογραφία αντικατοπτρίζει, επίσης, όλο και περισσότερο αυτή τη μεταστροφή.

Μέρος αυτού μπορεί να αποδοθεί στην επιρροή της πρώιμης κουλτούρας της δημοσιογραφίας δεδομένων: γράφοντας για την πρακτική αυτή στον Καναδά το 2016, οι Άλφρεντ Ερμίντα και

245. ΣτΜ: «Redlining», η πρακτική της «κόκκινης γραμμής» που ευνοεί τη δημιουργία οικιστικών περιοχών που κατοικούνται αποκλειστικά (ή σχεδόν αποκλειστικά) από λευκούς.
246. http://powerreporting.com/color/

Μέρι Λιν Γιανγκ σημείωσαν «έναν εξελισσόμενο καταμερισμό εργασίας που δίνει προτεραιότητα στις δια-οργανωτικές δικτυωμένες δημοσιογραφικές σχέσεις» (Hermida & Young, 2017). Και η επιρροή αναγνωρίστηκε περαιτέρω το 2018, όταν το Ινστιτούτο Reuters δημοσίευσε ένα βιβλίο σχετικά με την άνοδο της συνεργατικής δημοσιογραφίας, σημειώνοντας ότι «η συνεργασία μπορεί να γίνει θέμα από μόνη της, αυξάνοντας περαιτέρω τον αντίκτυπο της δημοσιογραφίας» (Sambrook, 2018).

Αλλάζοντας αυτό που μετράμε, πώς το μετράμε και αν το κάνουμε σωστά

Οι προηγμένες τεχνικές δεξιότητες δεν είναι απαραίτητες για τη δημιουργία μιας ιστορίας με αντίκτυπο. Ένα από τα μακροβιότερα έργα δημοσιογραφίας δεδομένων, το *Drone Warfare* του *Bureau of Investigative Journalism*, παρακολουθεί τα χτυπήματα των αμερικανικών μη επανδρωμένων αεροσκαφών εδώ και περισσότερα από πέντε χρόνια[247]. Ο πυρήνας της μεθοδολογίας του συνοψίζεται σε μια λέξη: επιμονή[248]. Σε εβδομαδιαία βάση, οι δημοσιογράφοι του *Bureau* μετατρέπουν τις εκθέσεις «ελεύθερου κειμένου» σε ένα σύνολο δομημένων δεδομένων που μπορούν να αναλυθούν, να αναζητηθούν και να διερευνηθούν. Τα δεδομένα αυτά, συμπληρωμένα από συνεντεύξεις με πηγές, χρησιμοποιήθηκαν από ΜΚΟ και το *Bureau* υπέβαλε γραπτά στοιχεία στην Επιτροπή Άμυνας του Βρετανικού Κοινοβουλίου[249].

Η καταμέτρηση των μη μετρήσιμων είναι ένας ιδιαίτερα σημαντικός τρόπος με τον οποίο η δημοσιογραφία δεδομένων μπορεί να έχει αντίκτυπο –και μάλιστα, ίσως είναι δίκαιο να πούμε ότι είναι το ισοδύναμο της δημοσιογραφίας δεδομένων που «δίνει φωνή σε όσους δεν έχουν φωνή». Το έργο The *Migrants' Files*, στο οποίο συμμετέχουν δημοσιογράφοι από περισσότερες από 15 χώρες, ξεκίνησε όταν οι δημοσιογράφοι παρατήρησαν ότι δεν υπήρχε «καμία αξιοποιήσιμη βάση δεδομένων για τους ανθρώπους που πέθαναν στην προσπάθειά τους να φτάσουν ή να παραμείνουν στην Ευρώπη» (The Migrants' Files, άνευ ημερομηνίας). Ο αντίκτυπός του ήταν να αναγκάσει άλλους οργανισμούς να αναλάβουν δράση: ο Διεθνής Οργανισμός Μετανάστευσης και άλλοι πλέον συλλέγουν τα δικά τους δεδομένα.

Ακόμα και όταν μια κυβέρνηση εμφανίζεται να μετράει κάτι, μπορεί να αξίζει να ερευνηθεί. Ενώ συνεργαζόμουν με το τμήμα δεδομένων του BBC Αγγλίας σε μια έρευνα σχετικά με το μέγεθος των περικοπών στις βιβλιοθήκες, για παράδειγμα, βίωσα μια στιγμή πανικού, όταν είδα ότι στο Κοινοβούλιο γινόταν μια ερώτηση σχετικά με τα δεδομένα για το θέμα (BBC News, 2016b). Θα μπορούσε η απάντηση να τινάξει στον αέρα την πολύμηνη εργασία μας; Στην πραγματικότητα, δεν το έκανε –αντιθέτως, επιβεβαίωσε ότι η ίδια η κυβέρνηση γνώριζε λιγότερα από όσα εμείς για την πραγματική κλίμακα αυτών των περικοπών, επειδή δεν είχε διεξαγάγει μια τόσο λεπτομερή έρευνα όσο εμείς.

Και μερικές φορές ο αντίκτυπος δεν έγκειται στην απλή ύπαρξη δεδομένων, αλλά στην αναπαράστασή τους: Ένα έργο της μεξικανικής εφημερίδας *El Universal,* το *Ausencias Ignoradas* («Αγνοημένες απουσίες»), δίνει πρόσωπο σε περισσότερες από 4.500 γυναίκες που έχουν εξαφανιστεί στη χώρα μέσα σε μια δεκαετία (Crosas Batista, 2016). Τα δεδομένα υπήρχαν, αλλά δεν είχαν αναλυθεί σε «ανθρώπινο» επίπεδο. Το *Meurtres conjugaux, des vies derrière les chiffres* («Ενδοοικογενειακές δολοφονίες, οι ζωές πίσω από τους αριθμούς») της *Libération* κάνει το ίδιο πράγμα για τις ενδοοικογενειακές δολοφονίες γυναικών, και το πρότζεκτ *Kadin Cinayetleri* («Γυναικοκτονίες») της Τζέιντα Ουλούκαγια έχει χαρτογραφήσει τις γυναικοκτονίες στην Τουρκία[250].

247. https://www.thebureauinvestigates.com/projects/drone-war
248. https://www.thebureauinvestigates.com/explainers/our-methodology
249. https://publications.parliament.uk/pa/cm201314/cmselect/cmdfence/772/772vw08.htm
250. https://www.liberation.fr/apps/2018/02/meurtres-conjugaux-derriere-les-chiffres/ (στα γαλλικά), http://kadincinayetleri.org/ (στα τουρκικά)

Όταν τα δεδομένα είναι κακά: Επιπτώσεις στην ποιότητα των δεδομένων

Μερικά από τα αγαπημένα μου έργα ως δημοσιογράφος δεδομένων ήταν εκείνα που επισήμαναν ελαττωματικά ή ελλιπή δεδομένα ή οδήγησαν στον εντοπισμό τους. Το 2016, το τμήμα δεδομένων του BBC Αγγλίας εξέτασε πόσες ακαδημαϊκές σχολές ακολουθούσαν τους κανόνες διαφάνειας: επιλέξαμε ένα τυχαίο δείγμα 100 ακαδημιών και ελέγξαμε αν δημοσίευαν μητρώο με τα οικονομικά συμφέροντα όλων των διοικητών τους, όπως απαιτείται από τους επίσημους κανόνες. Μία στις πέντε ακαδημίες δεν το έκανε –και, ως αποτέλεσμα, η ρυθμιστική αρχή Ofcom έλαβε μέτρα εναντίον εκείνων που είχαμε εντοπίσει (BBC News, 2016a). Θα μεριμνούσαν, όμως, στα σοβαρά να διασφαλίσουν ότι αυτό θα συνεχιστεί; Η επανεξέταση της ιστορίας σε μεταγενέστερο χρόνο θα ήταν σημαντική, για να διαπιστωθεί αν ο αντίκτυπος ήταν απλώς βραχυπρόθεσμος ή πιο συστημικός.

Ορισμένες φορές ο αντίκτυπος ενός έργου δημοσιογραφίας δεδομένων είναι ένα υποπροϊόν που εντοπίζεται μόνο όταν η ιστορία είναι έτοιμη και ζητούνται απαντήσεις. Όταν το *Bureau Local* φάνηκε να διαπιστώνει ότι 18 συμβούλια στην Αγγλία δεν είχαν τίποτα στα αποθεματικά τους, για να προστατευτούν από την οικονομική αβεβαιότητα, και ζήτησε μια απάντηση, αποδείχθηκε ότι τα στοιχεία ήταν λάθος. Κανείς δεν είχε παρατηρήσει τα εσφαλμένα δεδομένα, ανέφεραν. «Ούτε τα συμβούλια που συγκέντρωσαν τα στοιχεία, ούτε το Υπουργείο Στέγασης, Κοινοτήτων και Τοπικής Αυτοδιοίκησης, το οποίο τα έλεγξε και κατόπιν τα δημοσίευσε» (Davies, 2018). Η έρευνά τους ήρθε να προστεθεί σε μια κλιμακούμενη εκστρατεία που ζητούσε από τους τοπικούς φορείς να δημοσιεύουν τα δεδομένα με μεγαλύτερη συνέπεια, διαφάνεια και ακρίβεια.

Ο αντίκτυπος πέρα από την καινοτομία

Καθώς η δημοσιογραφία δεδομένων γίνεται όλο και πιο συνηθισμένη και ενσωματώνεται σε όλο και πιο σύνθετα επιχειρηματικά μοντέλα, ο αντίκτυπός της έχει μετατοπιστεί από τη σφαίρα της καινοτομίας στη σφαίρα του παραδοτέου. Όπως έγραψε ο συντάκτης δεδομένων Ντέιβιντ Ότεγουελ για τη διάκριση το 2018:

> Η καινοτομία φέρνει τη δημοσιογραφία δεδομένων στην πρώτη σελίδα. Το παραδοτέο τη φέρνει στην πρώτη σελίδα κάθε μέρα. Καινοτομία είναι η δημιουργία ενός εντυπωσιακού διαδραστικού υλικού που επιτρέπει στους αναγνώστες να εξερευνήσουν και να κατανοήσουν ένα σημαντικό θέμα. Το παραδοτέο συνεπάγεται να το κάνεις αυτό και να καταφέρεις να το χρησιμοποιήσει ένας μεγάλος αριθμός ανθρώπων και, στη συνέχεια, να φτιάξεις άλλο ένα την επόμενη μέρα και άλλο ένα τη μεθεπόμενη (Ottewell, 2018).

Επίσης, το παραδοτέο αφορά, φυσικά, τον αντίκτυπο πέρα από τους συναδέλφους μας, πέρα από τον παράγοντα εντυπωσιασμού μιας συναρπαστικής οπτικοποίησης ή ενός διαδραστικού χάρτη –στον πραγματικό κόσμο. Ο αντίκτυπος μπορεί να είναι άμεσος, προφανής και μετρήσιμος ή μπορεί να είναι αργός, αφανής και αόριστος. Μερικές φορές μπορεί να νιώθουμε ότι δεν πετύχαμε την αλλαγή –όπως στην περίπτωση του καθολικού ιερέα της *Boston Globe*–, αλλά η αλλαγή χρειάζεται χρόνο: το ρεπορτάζ μπορεί να σπείρει τους σπόρους της αλλαγής, με τα αποτελέσματα να έρχονται χρόνια ή δεκαετίες αργότερα. Το *Bureau Local* και το BBC δεν γνωρίζουν εάν τα δεδομένα των συμβουλίων ή των σχολείων θα είναι πιο αξιόπιστα στο μέλλον, αλλά γνωρίζουν ότι τα φώτα της δημοσιότητας είναι στραμμένα και στα δύο, προκειμένου αυτά να βελτιωθούν.

Μερικές φορές η δημοσιογραφία δεν μπορεί να κάνει τίποτε άλλο πέρα από το να στρέφει τα φώτα της δημοσιότητας προς μια κατεύθυνση και να αποδέχεται ότι είναι ευθύνη άλλων να αναλάβουν δράση· μερικές φορές αναλαμβάνει η ίδια δράση και κάνει εκστρατείες για μεγαλύτερη διαφάνεια. Σε αυτό το πλαίσιο, η δημοσιογραφία δεδομένων προσθέτει τη δυνατότητα να επιβληθεί μεγαλύτερη διαφάνεια ή να δημιουργηθούν τα εργαλεία που επιτρέπουν σε άλλους να αναλάβουν δράση.

Τελικά, η δημοσιογραφία δεδομένων που έχει αντίκτυπο μπορεί να καθορίσει την ατζέντα. Απευθύνεται σε ακροατήρια τα οποία άλλα είδη δημοσιογραφίας δεν προσεγγίζουν και τα εμπλέκει με τρόπους που άλλα είδη δημοσιογραφίας δεν το κάνουν. Δίνει φωνή σε όσους δεν έχουν φωνή και ρίχνει φως σε πληροφορίες που διαφορετικά θα παρέμεναν σκοτεινές. Θέτει τα δεδομένα προ των ευθυνών τους και λέει την αλήθεια στην εξουσία τους.

Μέρος του αντικτύπου αυτού είναι μετρήσιμο, ενώ ένα άλλο μέρος του είναι πιο δύσκολο να μετρηθεί –και αυτό θα πρέπει να ληφθεί υπόψη σε κάθε προσπάθεια παρακολούθησής του, αλλά αυτό δεν σημαίνει ότι δεν πρέπει να προσπαθήσουμε.

Παραπομπές

Angwin, J., & Tobin, A. (2018) «Fair housing groups sue Facebook for allowing discrimination in housing ads», *ProPublica*, 27 Μαρτίου. Διαθέσιμο σε: https://www.propublica.org/article/facebook-fair-housing-lawsuit-ad-discrimination (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

BBC News (2016a) «*Academy schools breach transparency rules*», 18 Νοεμβρίου. Διαθέσιμο σε: https://www.bbc.com/news/uk-england-37620007 (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

BBC News (2016b) «*Libraries lose a quarter of staff as hundreds close*», 29 Μαρτίου. Διαθέσιμο σε: https://www.bbc.com/news/uk-england-35707956 (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Blau, U. (2016) «How some 370 journalists in 80 countries made the Panama Papers happen», *Nieman Reports*, 6 Απριλίου. Διαθέσιμο σε: https://niemanreports.org/articles/how-some-370-journalists-in-80-countries-made-the-panama-papers-happen/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Byers, D. (2012) «*20% of NYT visitors read 538*», Politico, 6 Νοεμβρίου. Διαθέσιμο σε: https://www.politico.com/blogs/media/2012/11/nate-silver-draws-of-nyt-traffic-148670.html (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Cabra, M. (2017) «How ICIJ went from having no data team to being a tech-driven media organization», *ICIJ*, 29 Νοεμβρίου. Διαθέσιμο σε: https://www.icij.org/blog/2017/11/icij-went-no-data-team-tech-driven-media-organization/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Chammah, M. (2016) «*Policing the future*», *The Marshall Project,* 3 Φεβρουαρίου. Διαθέσιμο σε: https://www.themarshallproject.org/2016/02/03/policing-the-future (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Crosas Batista, M. (2016) «How one Mexican data team uncovered the story of 4,000 missing women», *Online Journalism Blog,* 22 Ιουνίου. Διαθέσιμο σε: https://onlinejournalismblog.com/2016/06/22/mexico-data-journalism-ausencias-ignoradas/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Davies, G. (2018) «Inaccurate and unchecked: Problems with local council spending data», The *Bureau of Investigative Journalism,* 2 Μαΐου. Διαθέσιμο σε: https://www.thebureauinvestigates.com/blog/2018-05-02/inaccurate-and-unchecked-problems-with-local-council-spending-data (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Fitzgibbon, W. & Díaz-Struck, E. (2016) «Panama Papers have had historic global effects—And the impacts keep coming», *ICIJ*, 1 Δεκεμβρίου. Διαθέσιμο σε: https://www.icij.org/investigations/panama-papers/20161201-global-impact/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Hermida, A. & Young, M. L. (2017) «Finding the data unicorn», *Digital Journalism*, 5(2), σσ. 159–176. DOI: https://doi.org/10.1080/21670811.2016.1162663

Higgerson, D. (2015) «*How audience metrics dispel the myth that readers don't want to get involved*

with serious stories», David Higgerson, 14 Οκτωβρίου. Διαθέσιμο σε: https://davidhiggerson.wordpress.com/2015/10/14/how-audience-metrics-dispel-the-myth-that-readers-dont-want-to-get-involved-with-serious-stories/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

ICIJ (άνευ ημερομηνίας) «*ICIJ's awards*». Διαθέσιμο σε: https://www.icij.org/about/awards/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

About Facebook (2017) «*Improving enforcement and promoting diversity: Updates to ads policies and tools»*, 8 Φεβρουαρίου. Διαθέσιμο σε: https://about.fb.com/news/2017/02/improving-enforcement-and-promoting-diversity-updates-to-ads-policies-and-tools/ (Προσπελάστηκε: 16 Δεκεμβρίου 2021).

Jalonick, M. C. (2017) «Facebook vows more transparency over political ads», *The Seattle Times,* 27 Οκτωβρίου. Διαθέσιμο σε: https://www.seattletimes.com/business/facebook-vows-more-transparency-over-political-ads/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

McClenaghan, M. (2017) «Campaigners target voters with Brexit "dark ads"», *The Bureau of Investigative Journalism,* 18 Μαΐου. Διαθέσιμο σε: https://www.thebureauinvestigates.com/stories/2017-05-18/campaigners-target-voters-brexit-dark-ads (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

The Migrants' Files (άνευ ημερομηνίας). Διαθέσιμο σε: http://www.themigrantsfiles.com/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Ottewell, D. (2018) *The evolution of data journalism* [Medium], 28 Μαρτίου. Διαθέσιμο σε: https://towardsdatascience.com/the-evolution-of-data-journalism-1e4c2802bc3d (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Rusbridger, A. (2018) «Alan Rusbridger: Who broke the news?», *The Guardian,* 31 Αυγούστου. Διαθέσιμο σε: https://www.theguardian.com/news/2018/aug/31/alan-rusbridger-who-broke-the-news (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Sambrook, R. (επιμ.) (2018). *Global teamwork: The rise of collaboration in investigative journalism.* Οξφόρδη: Reuters Institute for the Study of Journalism.

Schlemmer, C. (2016). Speed is not everything: How news agencies use audience metrics. Οξφόρδη: *Reuters Institute for the Study of Journalism.* Διαθέσιμο σε: https://reutersinstitute.politics.ox.ac.uk/our-research/speed-not-everything-how-news-agencies-use-audience-metrics (Προσπελάστηκε: 19 Δεκεμβρίου 2021)

Stark, J. (2016) «Investigating Uber surge pricing: A data journalism case study», *Global Investigative Journalism Network*, 2 Μαΐου. Διαθέσιμο σε: https://gijn.org/2016/05/02/investigating-uber-surge-pricing-a-data-journalism-case-study/ (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Ο συγγραφέας

Ο Πολ Μπράντσο διευθύνει τα Προγράμματα Μεταπτυχιακών Σπουδών στη Δημοσιογραφία Δεδομένων και στη Δημοσιογραφία Πολλαπλών Πλατφορμών και Κινητών Συσκευών στο Πανεπιστήμιο Birmingham City, ενώ παράλληλα με τη συγγραφή βιβλίων και την εκπαίδευση δημοσιογράφων σε διάφορους ειδησεογραφικούς οργανισμούς, εργάζεται ως σύμβουλος δημοσιογραφίας δεδομένων στο τμήμα δεδομένων του BBC Αγγλίας.

9.6

Για Ποιον Λόγο Υπάρχει η Δημοσιογραφία Δεδομένων; Μετρητά, Κλικ, Δοκιμές και Σφάλματα

Nikki Usher

Περίληψη: Εξετάζονται τα οικονομικά κίνητρα και οι ανεπιθύμητες συνέπειες της εμπορικής δημοσιογραφίας δεδομένων.

Λέξεις-κλειδιά: δημοσιογραφία, δεδομενοποίηση, παραπληροφόρηση, πολιτική οικονομία, εκλογές, πειραματισμός

Η καθημερινή ανανέωση των διαδραστικών προβλέψεων των εκλογικών χαρτών του *FiveThirtyEight* για τις εκλογές του 2016 αποτέλεσε σχεδόν ιεροτελεστία μεταξύ των συμπολιτών μου στην Ουάσινγκτον, από πολιτικούς μέχρι δημοσιογράφους, φοιτητές, κυβερνητικούς υπαλλήλους και όχι μόνο. Κάποιοι από αυτούς προτίμησαν την υπηρεσία συγκέντρωσης δημοσκοπήσεων του Upshot των *New York Times*, αυτοί που σκέφτονταν περισσότερο με βάση τις πιθανότητες, το *Real Clear Politics*, και όσοι είχαν πιο εξωτικά γούστα, στράφηκαν στην κάλυψη των αμερικανικών εκλογών από τον *Guardian*. Για αυτούς που ανανέωναν τις σελίδες με μανία, όλα ήταν και θα παρέμεναν εντάξει στον κόσμο, όσο οι πιθανότητες ήταν υπέρ της Χίλαρι Κλίντον στους Αγώνες Πείνας στους οποίους εξελίχθηκαν οι εκλογές στις ΗΠΑ, και όσο μεγαλύτερη ήταν η διαφορά, τόσο το καλύτερο.

Ξέρουμε πώς τελείωσε αυτή η ιστορία. Ο χάρτης του Νέιτ Σίλβερ, ακόμη και την ημέρα των εκλογών, έδειχνε ότι η Χίλαρι Κλίντον θα κέρδιζε με ποσοστό 71,4%. Ίσως έχει έρθει η ώρα να ξεπεράσουμε τις αμερικανικές εκλογές του 2016. Άλλωστε, η εμμονή με τους εκλογικούς χάρτες ίσως είναι ένα ιδιαίτερα αμερικανικό χόμπι, λόγω του σταθερού κύκλου των εθνικών εκλογών –αν και αυτό δεν σημαίνει ότι το παγκόσμιο κοινό δεν παρακολουθεί επίσης (Lewis & Waters, 2018). Βέβαια, μέχρι η καταστροφή των συνδέσμων να εξαφανίσει τον χάρτη, αυτός υπάρχει, εξακολουθεί να στοιχειώνει τους δημοσιογράφους και τους υποστηρικτές της Κλίντον, προσφέροντας τροφή στους Ρεπουμπλικάνους, για να υπενθυμίζουν στους εχθρούς τους ότι τα «λαμόγια τα media» προβάλλουν «ψευδείς ειδήσεις». Αφήνοντας κατά μέρος την πολιτική, οι δημοσιογράφοι δεδομένων δεν πρέπει να ξεχάσουν τις προεδρικές εκλογές του 2016 στις ΗΠΑ: ακόμα και αν η ποσοτικοποίηση ήταν σωστή κατά γενική ομολογία, οι αστοχίες στη χαρτογράφηση και την οπτικοποίηση έχουν γίνει ακόμα ένα εργαλείο μέσω του οποίου μπορεί να αποδομηθεί η αξίωση των δημοσιογράφων για γνωσιολογική αυθεντία (ή πιο απλά, η αξίωσή τους να θεωρούνται «εξουσιοδοτημένοι ειδήμονες»).

Ναι, είναι άδικο να συγχέουμε τη δημοσιογραφία δεδομένων με την εκλογική πρόβλεψη –σίγουρα είναι κάτι πολύ περισσότερο από αυτό, ιδίως από μια παγκόσμια σκοπιά, αλλά μερικές φορές φαίνεται ότι αυτή είναι η τελική συμβολή της δημοσιογραφίας δεδομένων: ατελείωτοι χάρτες, διαγράμματα που κλικάρονται, αριθμομηχανές επιρρεπείς σε σφάλματα χρηστών, υπεραπλούστευση και περιθωριοποίηση, ανεξάρτητα από την αυστηρότητα των υπολογισμών και τη στατιστική δεινότητα που τα παρήγαγε. Με τη δεύτερη έκδοση αυτού του εγχειριδίου που

κρατάτε τώρα στα χέρια σας, μπορούμε να δηλώσουμε ότι η δημοσιογραφία δεδομένων έχει φτάσει σε ένα σημείο ωρίμανσης και αυτοαναστοχασμού και, ως εκ τούτου, είναι σημαντικό να αναρωτηθούμε: «Για ποιον λόγο υπάρχει η δημοσιογραφία δεδομένων;».

Η δημοσιογραφία δεδομένων, όπως είναι σήμερα, εξακολουθεί να υπονοεί μόνο τις δυνατότητες που προσφέρει για την αναδιαμόρφωση και αναζωπύρωση της δημοσιογραφίας. Η πρώτη έκδοση αυτού του εγχειριδίου ξεκίνησε ως συνεργατικό έργο, σε ένα μεγάλο ομαδικό πλαίσιο το 2011 σε ένα Mozilla Festival, μια προσπάθεια που παρακολούθησα, αλλά γρήγορα αμφέβαλλα ότι θα υλοποιηθεί ποτέ σε ένα απτό αποτέλεσμα (έκανα λάθος). Αυτή η δεύτερη απόπειρα εκδίδεται τώρα από τις Εκδόσεις του Πανεπιστημίου του Άμστερνταμ[251] και διανέμεται στις Ηνωμένες Πολιτείες από τις Εκδόσεις του Πανεπιστημίου του Σικάγου με τη συνδρομή συνεργατών, γεγονός που υποδηλώνει ότι ο ελεύθερος χαρακτήρας της δημοσιογραφίας δεδομένων έχει αντικατασταθεί κατά κάποιον τρόπο με τον επαγγελματισμό, την τάξη και τη νομιμότητα. Και πράγματι, αυτό συμβαίνει: η δημοσιογραφία δεδομένων είναι κυρίαρχη, διδάσκεται στις σχολές δημοσιογραφίας και έχει καθιερωθεί στις αίθουσες σύνταξης (Usher, 2016). Η δημοσιογραφία δεδομένων έχει, επίσης, τυποποιηθεί και, ως τέτοια, έχει αλλάξει ελάχιστα τα τελευταία πέντε έως επτά χρόνια· οι ανασκοπήσεις των διεθνών διαγωνισμών δημοσιογραφίας δεδομένων αποκαλύπτουν περιορισμένη καινοτομία όσον αφορά στη μορφή και το θέμα (συχνότερα: πολιτική), με τους χάρτες και τα διαγράμματα να εξακολουθούν να είναι η πρώτη επιλογή (Loosen κ.α., 2020). Η διαδραστικότητα περιορίζεται σε αυτό που θεωρείται «τεχνική εισαγωγικού επιπέδου» από όσους ασχολούνται με την οπτικοποίηση πληροφοριών (Young κ.α., 2018)· επιπλέον, η δημοσιογραφία δεδομένων δεν έχει προχωρήσει αρκετά μακριά για την οπτικοποίηση «δυναμικών, κατευθυνόμενων και σταθμισμένων γραφημάτων» (Stoiber κ.α., 2015). Οι δημοσιογράφοι δεδομένων εξακολουθούν να ασχολούνται με προεπεξεργασμένα δεδομένα και όχι με πρωτότυπα «μεγάλα δεδομένα» («big data»). Και αυτά τα δεδομένα με τα οποία ασχολούνται είναι, στην καλύτερη περίπτωση, «μεγαλούτσικα» κυβερνητικά δεδομένα –και όχι πολυεπίπεδα δεδομένα με βάθος και μέγεθος του είδους που μπορεί να συλλέγει ένας πάροχος υπηρεσιών Διαδικτύου.

Αυτή η κριτική που ασκώ προέρχεται, σε μεγάλο βαθμό, από μια δυτικοκεντρική προοπτική (αν όχι από μια αμερικανοκεντρική θέση), αλλά αυτό δεν υπονομεύει την ουσιαστική έκκληση για δράση που διατυπώνω: οι δημοσιογράφοι δεδομένων εξακολουθούν να κρατούν στα χέρια τους μια δυνητικά επαναστατική εργαλειοθήκη για τη δημοσιογραφία που ακόμη δεν έχει αποκαλυφθεί. Η επανάσταση, ωστόσο, αν πραγματοποιηθεί ανεπαρκώς, θα υπονομεύσει περαιτέρω τόσο την εμπειρία του χρήστη όσο και τις προσπάθειες των καταναλωτών ειδήσεων για την αναζήτηση γνώσης και, στη χειρότερη περίπτωση, θα προκαλέσει περαιτέρω δυσπιστία απέναντι στις ειδήσεις. Εάν η δημοσιογραφία των δεδομένων συνεχίσει να λειτουργεί όπως τα τελευταία πέντε με δέκα χρόνια, τότε δεν συμβάλλει και πολύ στην προώθηση της δημοσιογραφίας στην ψηφιακή εποχή και στην εποχή των πλατφορμών. Έτσι, για να αρχίσουμε να θέτουμε αυτό το υπαρξιακό ερώτημα («για ποιον λόγο υπάρχει η δημοσιογραφία δεδομένων;»), προτείνω ότι οι δημοσιογράφοι δεδομένων (μαζί με τους δημοσιογράφους που ασχολούνται λιγότερο με τα δεδομένα, αλλά εντρυφούν στο Διαδίκτυο, οι οποίοι εργάζονται με βίντεο, ήχο και κώδικα), καθώς και οι μελετητές που τους πιέζουν και τους ωθούν, πρέπει να επανεξετάσουν την ιστορία προέλευσης της δημοσιογραφίας δεδομένων, την παρούσα λογική της και το μέλλον της.

Δημοσιογραφία δεδομένων στις Ηνωμένες Πολιτείες: η ιστορία προέλευσής της

Η ιστορία προέλευσης είναι η ιστορία που λέμε στους εαυτούς μας για το πώς και γιατί δημιουργηθήκαμε, και τις περισσότερες φορές τη βλέπουμε μέσα από ροζ γυαλιά και είναι περισσότερο κομπορρημοσύνη παρά πραγματικότητα. Η ιστορία προέλευσης της δημοσιογραφίας δεδομένων

251. ΣτΕ: Αναφορά στο πρωτότυπο έντυπο έργο, που εκδόθηκε και κυκλοφόρησε το 2021.

στις Ηνωμένες Πολιτείες πάει κάπως έτσι: στον αρχέγονο κόσμο πριν από τη δημοσιογραφία δεδομένων, εκείνη υπήρχε σε μια προηγούμενη μορφή, ως ρεπορτάζ υποβοηθούμενο από υπολογιστή (ή έτσι, «computer-assisted reporting», ονομαζόταν στις Ηνωμένες Πολιτείες), το οποίο προσέφερε την ευκαιρία να προσδοθεί η αυστηρότητα των κοινωνικών επιστημών στη δημοσιογραφία.

Στον μύθο της εισόδου της δημοσιογραφίας δεδομένων στον Ιστό, οι δημοσιογράφοι δεδομένων γίνονται ερευνητές δημοσιογράφοι με την ανώτερη υπολογιστική ικανότητα του 21ου αιώνα, οι οποίοι απελευθερώνουν τα δεδομένα (ή τα έγγραφα), προκειμένου να βοηθήσουν να ακουστούν ιστορίες που διαφορετικά δεν θα λέγονταν. Πέρα όμως από τη διερεύνηση ιστοριών, οι δημοσιογράφοι δεδομένων επρόκειτο, επίσης, να σώσουν τη δημοσιογραφία με τις νέες διαδικτυακές τους δεξιότητες, προσφέροντας ένα επίπεδο διαφάνειας, εξατομίκευσης και διαδραστικότητας στις ειδήσεις που οι καταναλωτές θα εκτιμούσαν, θα μάθαιναν από αυτές και, φυσικά, θα τους έκαναν κλικ. Οι διαδικτυακές ιστορίες του παρελθόντος, τρόπον τινά, δεν θα ήταν ποτέ οι ίδιες.

Η δημοσιογραφία δεδομένων θα διόρθωνε τα κακώς κείμενα και θα παρείχε την απαραίτητη αντικειμενική βάση που έλειπε από τις ποιοτικές αξιολογήσεις της δημοσιογραφίας, κάνοντάς το αυτό σε μεγάλη κλίμακα και με μια μαεστρία αδιανόητη πριν από το σημερινό διαδραστικό ψηφιακό περιβάλλον μας του πραγματικού χρόνου, που είναι γεμάτο με ισχυρούς διακομιστές που εδρεύουν στο cloud και απαλλάσσουν τους υπολογιστές από την πίεση σε κάθε ειδησεογραφικό οργανισμό. Τα πρώτα σημάδια επιτυχίας θα χάραζαν την πορεία προς το μέλλον και θα μετέτρεπαν ακόμη και απλούς αναγνώστες σε ερευνητές συνεργάτες ή επιστήμονες-πολίτες, όπως με την κάλυψη του σκανδάλου των βουλευτικών εξόδων από τον *Guardian* ή με το έργο *Cicada Tracker* («Ιχνηλατώντας τα Τζιτζίκια») του δημόσιου ραδιοφωνικού σταθμού WNYC της Νέας Υόρκης, το οποίο έβαλε έναν μικρό στρατό κατοίκων της περιοχής να κατασκευάσει θερμόμετρα εδάφους, για να βοηθήσει στην καταγραφή της άφιξης των φοβερών καλοκαιρινών εντόμων. Και αυτή η εμπνευσμένη ενορχήστρωση της δημοσιογραφίας, του υπολογισμού, του πλήθους, των δεδομένων και της τεχνολογίας θα συνεχιστεί, οδηγώντας την αλήθεια στη δικαιοσύνη.

Το Παρόν: Ο δημοσιογράφος-«χάκερ» ως ακόμη ένας (βαρετός) εργαζόμενος στο newsroom

Το παρόν δεν έχει απομακρυνθεί πολύ από την ιστορία προέλευσης που οι σημερινοί δημοσιογράφοι δεδομένων έχουν διηγηθεί στον εαυτό τους, ούτε ως προς το όραμα ούτε ως προς την πραγματικότητα. Αυτό που προέκυψε είναι δύο διαφορετικοί τύποι δημοσιογραφίας δεδομένων: η «ερευνητική» δημοσιογραφία δεδομένων, που συνεχίζει να κουβαλά τον ευγενή μανδύα των προσπαθειών της δημοσιογραφίας, και η καθημερινή δημοσιογραφία δεδομένων, η οποία μπορεί να αξιοποιηθεί με γνώμονα το τελευταίο ενδιαφέρον viral –το οποίο μπορεί να σημαίνει οτιδήποτε, από μια επείγουσα προσπάθεια δημοσιογραφικής χαρτογράφησης μέχρι τη μετατροπή μιας δημοσκόπησης κοινής γνώμης ή μιας ερευνητικής μελέτης σε ένα εύκολα κοινοποιήσιμο meme με το επίχρισμα της δημοσιογραφίας. Η δημοσιογραφία δεδομένων, στην καλύτερη περίπτωση, έχει γίνει βαρετή και υπερβολικά επαγγελματική και, στη χειρότερη, έχει γίνει άλλη μια στρατηγική για τη δημιουργία ψηφιακών εσόδων.

Δεν είναι υπερβολή να πούμε ότι η δημοσιογραφία δεδομένων θα μπορούσε να έχει μεταμορφώσει τη δημοσιογραφία όπως την ξέρουμε, αλλά μέχρι στιγμής δεν το έχει κάνει. Στο Mozilla Festival του 2011, ένας από τους κορυφαίους χάκερ του φεστιβάλ ήταν ένα plug-in που επέτρεπε στο πρόσωπο οποιουδήποτε να γίνει η κεντρική φωτογραφία μιας μακέτας της αρχικής σελίδας της *Boston Globe*. Αυτό ήταν διασκεδαστικό, αλλά η *Boston Globe* σίγουρα δεν επρόκειτο να επιτρέψει σε περιεχόμενο που δημιουργείται από χρήστες να χρησιμοποιηθεί στην αρχική της σελίδα χωρίς κανενός είδους φιλτράρισμα. Ομοίως, κατά τη γέννηση του πρώτου *Data Journalism Handbook*, ο δημοσιογράφος δεδομένων ήταν ο δημοσιογράφος-«χάκερ», ο οποίος απεικονιζόταν σαν να ερχόταν από την τεχνολογία στη δημοσιογραφία ή, τουλάχιστον,

να χρησιμοποιούσε το πνεύμα του ανοιχτού κώδικα και του «χάκινγκ», για να εμπνεύσει έργα που αντιστέκονταν στις συμβατικές διαδικασίες της επίσημης δημοσιογραφίας και παρείχαν χώρο για πειραματισμό, ατέλεια και παιχνίδι –μερεμέτια που είχαν σκοπό να οδηγήσουν σε κάτι το οποίο μπορεί να μην είναι σπουδαίο σε μορφή ή σε περιεχόμενο, αλλά θα μπορούσε να υπονομεύσει τη δημοσιογραφία (Lewis & Usher, 2013). Το 2011, η ιστορία ήταν η μετακίνηση των outsider στη δημοσιογραφία, αλλά το 2018, η ιστορία είναι η επαγγελματοποίηση του προγραμματισμού στη δημοσιογραφία από τους insider. Επομένως, μέσα σε πέντε χρόνια το πνεύμα της καινοτομίας και της εφευρετικότητας έχει γίνει σαφώς εταιρικό, σαφώς υπαλληλικό και σαφώς λιγότερο διασκεδαστικό (Lewis & Usher, 2016).

Το βαρετό είναι εντάξει και παίζει τον ρόλο του. Μέρος της επαγγελματοποίησης της δημοσιογραφίας δεδομένων έχει δικαιολογηθεί με την αυτοαντίληψη του «δημοσιογράφου δεδομένων ως ήρωα» –οι δημοσιογράφοι δεδομένων ως εκείνοι που, χάρη σε ένα διαφορετικό σύνολο αξιών (π.χ. συνεργασία, διαφάνεια) και δεξιοτήτων (οπτικοποίηση, διάφορες υπολογιστικές δεξιότητες) θα μπορούσαν να φέρουν την αλήθεια στην εξουσία με νέους τρόπους. Τα *Panama* και *Paradise Papers* είναι ίσως μία από τις καλύτερες εκφράσεις αυτού του οράματος. Όμως, η ερευνητική δημοσιογραφία δεδομένων απαιτεί χρόνο, προσπάθεια και εξειδίκευση που υπερβαίνει κατά πολύ την απλή επεξεργασία δεδομένων και περιλαμβάνει πολλές άλλες πηγές πιο παραδοσιακών δεδομένων, κυρίως συνεντεύξεις, επιτόπια ρεπορτάζ και έγγραφα. Μια τακτικά εμφανιζόμενη, πρωτοποριακή ερευνητική δημοσιογραφία είναι σχήμα οξύμωρο, αν και όχι λόγω έλλειψης προσπάθειας: Το European Data Journalism Network, το Institute for NonProfit News των Ηνωμένων Πολιτειών και το Global Investigative Journalism Network αναδεικνύουν το τεράστιο δίκτυο των επίδοξων ερευνητικών προσπαθειών. Η αλήθεια είναι ότι δεν είναι εύκολο να γίνει μια έρευνα που αλλάζει τα πάντα, γι' αυτό και μπορούμε γενικά να μετρήσουμε αυτές τις επιτυχίες υψηλού επιπέδου σε περίπου δέκα δάχτυλα των χεριών μας, ενώ η επιτυχία της έρευνας μέσω του πληθοπορισμού («crowdsourcing») στο παράδειγμα των εξόδων των βουλευτών του *Guardian*, από το 2010, ακόμη δεν έχει αντικατασταθεί από κάτι νεότερο.

Το παρελθόν είναι παρελθόν όσον αφορά στη δημοσιογραφία δεδομένων. Το *Snow Fall*, το επαναστατικό πρότζεκτ με την καθηλωτική αφήγηση των *New York Times*, που κέρδισε βραβείο Πούλιτζερ το 2012, εμφανίστηκε τον Δεκέμβριο του 2017 ως *Deliverance from 27,000 Fee*t ή *Everest*. Πέντε χρόνια αργότερα, οι *New York Times* παρουσίασαν άλλο ένα μακροσκελές θέμα για μια καταστροφή σε ένα χιονισμένο βουνό, απλώς λίγο διαφορετικά (αλλά από τον ίδιο συγγραφέα, τον Τζον Μπραντς). Μέσα σε αυτά τα πέντε χρόνια, το «Snowfall» ή «Snowfalled» καθιερώθηκε ως συντομογραφία, εντός και εκτός των *New York Times*, για την προσθήκη μιας διαδραστικής πινελιάς σε μια ιστορία. Μετά το 2012, μια αντιπαράθεση μαίνεται όχι μόνο στους *Times*, αλλά και σε άλλες αίθουσες σύνταξης των ΗΠΑ και του Ηνωμένου Βασιλείου, σχετικά με το αν οι δημοσιογράφοι δεδομένων θα έπρεπε να ξοδεύουν τον χρόνο τους στην κατασκευή εργαλείων που μπορούν αυτομάτως να μετατρέψουν σε θέμα «*Snowfall*» οποιαδήποτε ιστορία ή αν πρέπει να εργάζονται σε καινοτόμα έργα εφάπαξ (Usher, 2016). Στο μεταξύ, το Snow Fall, ελάχιστα διαδραστικό στην καλύτερη περίπτωση το 2012, παρέμεινε ελάχιστα διαδραστικό στην καλύτερη περίπτωση στο τέλος του 2017.

«Μα, για μια στιγμή», μπορεί να διακηρύξει ο πρώην δημοσιογράφος δεδομένων, «το *Snow Fall* δεν είναι δημοσιογραφία δεδομένων –ίσως να είναι ένα φανταχτερό κόλπο ορισμένων προγραμματιστών ενημερωτικών εφαρμογών, αλλά δεν υπάρχουν δεδομένα στο *Snow Fall*!». Εδώ έγκειται το πρόβλημα: ίσως οι δημοσιογράφοι δεδομένων να μην πιστεύουν ότι το *Snow Fall* είναι δημοσιογραφία δεδομένων, αλλά γιατί όχι; *Σε τι χρησιμεύει η δημοσιογραφία δεδομένων, αν όχι στην αφήγηση ιστοριών με νέους τρόπους και νέες δεξιότητες που αξιοποιούν τα καλύτερα στοιχεία του Διαδικτύου;*

Επίσης, η δημοσιογραφία δεδομένων δεν μπορεί να είναι μόνο για χάρτες ή διαγράμματα, ούτε η χαρτογράφηση ή η αναπαράσταση δεδομένων σε διαγράμματα προσδίδει στη δημοσιογρα-

φία δεδομένων πνευματική υπεροχή έναντι των προσπαθειών εμβυθιστικής ψηφιακής δημοσιογραφίας («immersive digital journalism»). Ό,τι μπορεί να χαρτογραφηθεί, χαρτογραφείται. Πέρα από τη χαρτογράφηση των εκλογών στις Ηνωμένες Πολιτείες, οι δεοντολογικές συνέπειες της ποσοτικοποίησης και της οπτικοποίησης των τελευταίων διαθέσιμων δεδομένων με μια συνεκτικότητα διαθέσιμη σε ένα κλικ χρήζουν κριτικής. Στην πιο συνηθισμένη της μορφή, η δημοσιογραφία δεδομένων γίνεται ο μαϊντανός της οπτικοποίησης. Αυτό είναι ιδιαίτερα αληθές, δεδομένης της μετάβασης προς την καθημερινή και τακτική ζήτηση για έργα δημοσιογραφίας δεδομένων. Μπορεί να είναι νέα στατιστικά στοιχεία για την απασχόληση, δεδομένα για την ποδηλασία στην πόλη, ποσοστά ανακύκλωσης, τα αποτελέσματα μιας ακαδημαϊκής μελέτης –οποιαδήποτε οπτικοποίηση γίνεται, επειδή μπορεί να γίνει (και ίσως λάβει περισσότερα κλικ). Στη χειρότερη περίπτωση, η δημοσιογραφία δεδομένων μπορεί να υπεραπλουστεύσει σε σημείο που να στερεί το ανθρώπινο στοιχείο, το οποίο υποτίθεται πως φωτίζει το έργο της, από τα δεδομένα. Οι χάρτες των μεταναστών και των ροών τους στην Ευρώπη έχουν τη μορφή διαδραστικών βελών ή εικονιδίων ατόμων χωρίς φύλο. Όπως υποστηρίζει ο ανθρωπογεωγράφος Πολ Άνταμς, η ψηφιακή χαρτογράφηση των ειδήσεων έχει μετατρέψει την προσφυγική κρίση σε μια ασώματη σειρά ενεργειών με δυνατότητα κλικ, το ακριβώς αντίθετο από ό,τι θα μπορούσε να κάνει η δημοσιογραφία, για να μετατρέψει τους άγνωστους «πρόσφυγες» σε κάτι ζωντανό, κάτι περισσότερο από έναν αριθμό (Adams, 2017). Πριν από τη χαρτογράφηση ενός ακόμη κοινωνικού προβλήματος ή μιας ακόμη ακαδημαϊκής μελέτης, οι δημοσιογράφοι δεδομένων πρέπει να αναρωτηθούν: για ποιον σκοπό χαρτογραφούμε και δημιουργούμε διαγράμματα (ή διαγραμματογραφούμε, για την ακρίβεια);

Και κάπου μεταξύ του *Snow Fall* και των χαρτών μετανάστευσης έγκειται το πρόβλημα: για ποιον λόγο υπάρχει η δημοσιογραφία δεδομένων; Στο παρόν παρουσιάζονται κυρίως στοιχεία επαγγελματισμού και ισομορφισμού, με μια δόση εταιρικού κινήτρου, σύμφωνα με το οποίο η δημοσιογραφία δεδομένων δεν προορίζεται απλώς για να βοηθά τους καταναλωτές ειδήσεων να κατανοήσουν τον κόσμο, αλλά απώτερος σκοπός της είναι να γεμίζει τα ταμεία των ειδησεογραφικών οργανισμών. Σίγουρα αυτό δεν είναι το μόνο που μπορεί να προσφέρει η δημοσιογραφία δεδομένων.

Το Μέλλον: Πώς μπορεί η δημοσιογραφία δεδομένων να ανακτήσει την αξία της (και να είναι και διασκεδαστική)

Για ποιον λόγο υπάρχει η δημοσιογραφία δεδομένων; Η δημοσιογραφία δεδομένων πρέπει να επιστρέψει στις ρίζες της αλλαγής και της επανάστασης, του εμπνευσμένου «χάκινγκ» και του πειραματισμού, του αυτοπροσδιοριζόμενου οράματος των αποστατών που κινούνται μέσα σε μια κουρασμένη και ανέμπνευστη βιομηχανία, για να αναγκάσουν τους δημοσιογράφους να αναμετρηθούν με την υποτιθέμενη εξουσία τους πάνω στη γνώση, την αφήγηση και τη διανομή. Οι δημοσιογράφοι δεδομένων πρέπει να παραδεχτούν ότι έχουν εμπνευστεί από τους χάκερ και να «χακάρουν» την αίθουσα σύνταξης, όπως είχαν υποσχεθεί κάποτε ότι θα κάνουν· πρέπει να ξεπεράσουν την εστίαση στο κέρδος και τον επαγγελματισμό μέσα στις αίθουσες σύνταξης. Η ανάκτηση της ιδιότητας του περιθωριακού θα μας φέρει πιο κοντά στην ουσιαστική προσφορά που υποσχέθηκε η δημοσιογραφία δεδομένων: ένας τρόπος να σκεφτούμε διαφορετικά για τη δημοσιογραφία, να παρουσιάσουμε διαφορετικά τη δημοσιογραφία, να φέρουμε νέα είδη επαγγελματιών, και σκεπτόμενων και αποτελεσματικών, στην αίθουσα σύνταξης και, πέρα από αυτό, να αναζωογονήσουμε τη δημοσιογραφία.

Στο μέλλον, φαντάζομαι ότι η δημοσιογραφία δεδομένων θα απεγκλωβιστεί από τον όρο «δεδομένα» και θα επικεντρωθεί στη λέξη «δημοσιογραφία». Οι δημοσιογράφοι δεδομένων έχουν κατά πάσα πιθανότητα δεξιότητες που οι υπόλοιποι μέσα στην αίθουσα σύνταξης ή άλλοι δημοσιογράφοι δεν έχουν: την ικανότητα να κατανοούν πολύπλοκα δεδομένα ή να χρησιμοποιούν έναν υπολογιστή για να το κάνει αυτό για λογαριασμό τους, την ικανότητα να οπτικοποιούν

αυτά τα δεδομένα με έναν πιθανώς ουσιαστικό τρόπο και την ικανότητα να προγραμματίζουν. Η δημοσιογραφία δεδομένων, ωστόσο, πρέπει να γίνει αυτό που έχω αποκαλέσει διαδραστική δημοσιογραφία –η δημοσιογραφία δεδομένων πρέπει να αποβάλει την «κομπιουτερίστικη» παρόρμησή της για τη δημιουργία χαρτών και διαγραμμάτων, καθώς και την περιφρόνησή της για τεχνολογίες και δεξιότητες που δεν εστιάζουν στα δεδομένα, όπως το βίντεο 360 μοιρών, η επαυξημένη πραγματικότητα και τα κινούμενα σχέδια. Στο δικό μου όραμα για το μέλλον, θα υπάρξουν πολύ περισσότερες διαδραστικές εφαρμογές σαν το έργο *Secret Life of the Cat* του BBC και τα *Dialect Quizzes* των *New York Times*· θα υπάρξουν περισσότερα έργα που θα συνδυάζουν τα βίντεο 360 μοιρών ή την εικονική πραγματικότητα με δεδομένα, όπως η προσπάθεια του Dataverse που χρηματοδοτείται από την πρωτοβουλία για εμβυθιστικές ειδήσεις Journalism 360. Θα έχουμε πολύ λιγότερη χαρτογράφηση των εκλογών και χαρτογραφία που θα απεικονίζει τις ειδήσεις της ημέρας, ανάγοντας τα μακρινά θύματα σε γραμμές και ροές που βλέπεις με ένα κλικ. Ας ελπίσουμε ότι θα μπει ένα τέλος στη νέα τάση προς τα διαδραστικά μέσα που παρουσιάζουν σε πραγματικό χρόνο τα αποτελέσματα των δημοσκοπήσεων, ένα νέο φετίχ των κορυφαίων ειδησεογραφικών πρακτορείων στις Ηνωμένες Πολιτείες. Αντίθετα, θα υπάρχει πολύ περισσότερη πρωτοτυπία, διασκέδαση και εμπνευσμένη υπέρβαση σε αυτό που υποτίθεται ότι πρέπει να μοιάζει και πρέπει να κάνει η δημοσιογραφία. Η δημοσιογραφία δεδομένων αποσκοπεί στη λογοδοσία αλλά και στη διασκέδαση και στη φαντασία· αποκτά τη δύναμή της όχι μόνο επειδή ένας βουλευτής μπορεί να παραιτηθεί ή μια γραμμή τάσης να γίνει σαφέστερη, αλλά και επειδή οι απλοί άνθρωποι βλέπουν την αξία της επιστροφής στους ειδησεογραφικούς οργανισμούς και στους δημοσιογράφους, διότι οι δημοσιογράφοι καλύπτουν ποικίλες ανθρώπινες ανάγκες πληροφόρησης –για προσανατολισμό, για ψυχαγωγία, για την κοινότητα και όχι μόνο.

Και για να διεκδικήσουν πραγματικά ανώτερη γνώση σχετικά με τα δεδομένα, οι δημοσιογράφοι δεδομένων που σκοπεύουν να καταστήσουν τα δεδομένα γνωστά και κατανοητά πρέπει να συλλέξουν αυτά τα δεδομένα μόνοι τους –η δημοσιογραφία δεδομένων δεν είναι μόνο για την παραγωγή νέων οπτικοποιήσεων δεδομένων που έχουν συλλεχθεί από κάποιον άλλον. Στην καλύτερη περίπτωση, η επεξεργασία των δεδομένων κάποιου άλλου καθιστά ορατές τις υποθέσεις των παρόχων δεδομένων· στη χειρότερη, η δημοσιογραφία δεδομένων γίνεται τόσο στενογραφική όσο ένα δελτίο Τύπου για τον πάροχο δεδομένων. Ωστόσο, πολλοί δημοσιογράφοι δεδομένων δεν δείχνουν ιδιαίτερο ενδιαφέρον για τη συλλογή των δικών τους δεδομένων και τη θεωρούν εκτός των ορίων του ρόλου τους· όπως εξήγησε ο αρχισυντάκτης δεδομένων της *Washington Post*, Στίβεν Ριτς, σε ένα tweet, η *Post* «και άλλοι δεν πρέπει να χρειάζεται να συλλέγουν και να διατηρούν βάσεις δεδομένων που είναι πανεύκολο να συλλέξει η κυβέρνηση. Δεν θα έπρεπε να είναι αυτή η γαμημένη η δουλειά μας» (Rich, 2018). Ταυτόχρονα, ωστόσο, οι στατιστικές για την ένοπλη βία που ο Ριτς εκνευρίστηκε, επειδή έπρεπε να συντηρήσει, είναι πιο ενδυναμωτικές από όσο συνειδητοποιεί: ενσωματωμένες στα κυβερνητικά δεδομένα, είναι υποθέσεις και αποφάσεις σχετικά με το τι πρέπει να συλλεχθεί, οι οποίες χρειάζονται την κατάλληλη έρευνα και εξέταση. Τα δεδομένα δεν είναι αδρανή, αλλά γεμάτα με τεκμήρια σχετικά με τα γεγονότα που έχουν σημασία. Οι δημοσιογράφοι που επιδιώκουν να αναλάβουν τον έλεγχο του πεδίου των γεγονότων πρέπει να είναι σε θέση να εξηγήσουν γιατί τα γεγονότα είναι αυτά που είναι, και, στην πραγματικότητα, η συστηματική παραγωγή γεγονότων είναι ο τρόπος με τον οποίο οι δημοσιογράφοι αξιώνουν τη γνωσιολογική τους αυθεντία στο μεγαλύτερο μέρος της σύγχρονης δημοσιογραφίας.

Η δημοσιογραφία των δεδομένων, λοιπόν, μπορεί να προσφέρει πολλά περισσότερα από όσα προσφέρει ως τώρα –μπορεί να προσφέρει διασκέδαση, παιχνίδι και πειραματισμό. Μπορεί να αλλάξει τον τρόπο με τον οποίο λέγονται οι ιστορίες και μπορεί να ενθαρρύνει νέους τρόπους σκέψης σχετικά με αυτές. Όμως, μπορεί, επίσης, να διαδραματίσει ζωτικό ρόλο στην αποκατάσταση της δημοσιογραφίας ως αφηγητή της αλήθειας και των γεγονότων· στη δημιουργία και τη γνώση των δεδομένων και στην ικανότητα να εξηγεί τη διαδικασία παρατήρησης και συλ-

λογής δεδομένων που οδήγησαν σε ένα γεγονός. Η δημοσιογραφία δεδομένων θα μπορούσε κάλλιστα να γίνει μια βασική γραμμή υπεράσπισης για το πώς οι επαγγελματίες δημοσιογράφοι μπορούν και συλλέγουν γεγονότα καλύτερα από ό,τι θα μπορούσε ποτέ οποιοδήποτε άλλο επάγγελμα, θεσμός ή απλό άτομο.

Παραπομπές

Adams, P. (2017) «Migration maps with the news: Guidelines for ethical visualization of mobile populations», *Journalism Studies*, 19(1), σσ. 1–21. DOI: https://doi.org/10.1080/1461670X.2017.1375387

Lewis, N. P. & Waters, S. (2018) «Data journalism and the challenge of shoe-leather epistemologies», *Digital Journalism*, 6(6), σσ. 719–736. DOI: https://doi.org/10.1080/21670811.2017.1377093

Lewis, S. C. & Usher, N. (2013) «Open source and journalism: Toward new frameworks for imagining news innovation», *Media, Culture & Society,* 35(5), σσ. 602–619. DOI: https://doi.org/10.1177/0163443713485494

Lewis, S. C. & Usher, N. (2016) «Trading zones, boundary objects, and the pursuit of news innovation: A case study of journalists and programmers», *Convergence*, 22(5), σσ. 543–560. DOI: https://doi.org/10.1177/1354856515623865

Loosen, W., Reimer, J., & De Silva-Schmidt, F. (2020) «Data-driven reporting: An on-going (r) evolution? An analysis of projects nominated for the Data Journalism Awards 2013–2016», *Journalism*, 21(9), σσ. 1246–1263. DOI: https://doi.org/10.1177/1464884917735691

Rich, S. (2018) *The @washingtonpost and others should not have to collect and maintain databases that are no-brainers for the government to collect. This should not be our fucking job.* [Twitter], 15 Φεβρουαρίου. Διαθέσιμο σε: https://twitter.com/dataeditor/status/964160884754059264 (Προσπελάστηκε: 19 Δεκεμβρίου 2021).

Stoiber, C., Aigner, W. & Rind, A. (2015) «Survey on visualizing dynamic, weighted, and directed graphs in the context of data-driven journalism», *Proceedings of the International Summer School on Visual Computing (ICVSS 2015)*, 12-18 Ιουλίου, Σικελία, Ιταλία, σσ. 49–58.

Usher, N. (2016). *Interactive journalism: Hackers, data, and code.* Ουρμπάνα, Ιλινόις: University of Illinois Press.

Young, M. L., Hermida, A., & Fulda, J. (2018) «What makes for great data journalism? A content analysis of Data Journalism Awards finalists 2012–2015», *Journalism Practice,* 12(1), σσ. 115–135. DOI: https://doi.org/10.1080/17512786.2016.1270171

Η συγγραφέας

Η Νίκι Άσερ είναι Αναπληρώτρια Καθηγήτρια στο University of Illinois Urbana-Champaign και συγγραφέας του βιβλίου *Interactive Journalism: Hackers, Data, and Code* (University of Illinois Press, 2016).

9.7
Δημοσιογραφία Δεδομένων και Ψηφιακός Φιλελευθερισμός

Dominic Boyer

Περίληψη: Πώς η άνοδος της δημοσιογραφίας δεδομένων συνδιαλέγεται με τον πολιτικό φιλελευθερισμό.

Λέξεις-κλειδιά: φιλελευθερισμός, καθιστική δημοσιογραφία, εργασία στην οθόνη, οριζόντια ανταλλαγή μηνυμάτων, αυτολογική ατομικότητα, δημοσιογραφία δεδομένων

Τα τελευταία 30 χρόνια έχει συντελεστεί ένας μαζικός μετασχηματισμός του δημοσιογραφικού επαγγέλματος και της οργανωτικής κουλτούρας της παραγωγής ειδήσεων. Τα αίτια και τα αποτελέσματα αυτού του μετασχηματισμού είναι πολύ περίπλοκα για να αναλυθούν εδώ. Αρκεί να πούμε ότι το μοντέλο των έντυπων και επίγειων ραδιοτηλεοπτικών μέσων, που φαινόταν ακόμη αρκετά ισχυρό μέχρι και τη δεκαετία του 1990, έχει σχεδόν πλήρως αντικατασταθεί από ένα ψηφιακό μοντέλο ειδησεογραφικών μέσων που δημιουργήθηκε από την άνοδο του Διαδικτύου, των μηχανών αναζήτησης και των μέσων κοινωνικής δικτύωσης ως κυρίαρχων συστημάτων επικοινωνίας και πληροφόρησης, καθώς και από την εκτεταμένη χρηματιστικοποίηση και ιδιωτικοποίηση των ειδησεογραφικών μέσων που προωθείται από την οικονομική φιλοσοφία του «νεοφιλελευθερισμού». Όπως αναφέρεται στον παρόντα τόμο, οι πολλαπλασιαζόμενες ψηφιακές ροές και αναπαραστάσεις δεδομένων αποτελούν πλέον την προκαθορισμένη συνθήκη της δημοσιογραφικής πρακτικής. Το σύνολο της δημοσιογραφίας είναι πλέον, σε κάποιο βαθμό, «δημοσιογραφία δεδομένων». Το 2013, στο βιβλίο μου, *The Life Informatic*, περιέγραψα αυτή τη διαδικασία ως «οριζόντια επανάσταση» («lateral revolution»). Η θεωρία που προτείνω είναι ότι γίναμε μάρτυρες μιας μετατόπισης της οικολογίας της ενημέρωσης, από την κυριαρχία των ακτινικών (π.χ., σε μεγάλο βαθμό μονοκατευθυντικών, με τη μορφή ακτινωτού τροχού) υποδομών των μέσων ενημέρωσης σε οριζόντιες (π.χ., σε μεγάλο βαθμό πολυκατευθυντικές, από σημείο σε σημείο) υποδομές[252] (Boyer, 2013).

Όπως παρατήρησε ο Ρέιμοντ Γουίλιαμς στη λαμπρή ιστορική μελέτη του για την ανάπτυξη της τηλεόρασης (Williams, 1974), τα ηλεκτρονικά μέσα παρουσιάζουν τόσο ακτινικές όσο και οριζόντιες δυνατότητες από τον 18ο αιώνα. Ο τρόπος με τον οποίο αυτές οι δυνατότητες απελευθερώθηκαν και θεσμοθετήθηκαν καθοδηγείται πάντα από κοινωνικές και πολιτικές συνθήκες πέρα από τις ίδιες τις τεχνολογίες. Υπήρχε ένα πρωτότυπο μηχάνημα φαξ πάνω από έναν αιώνα πριν να υπάρξει μια προφανής κοινωνική ανάγκη για μια τέτοια τεχνολογία και, έτσι, η επίσημη «εφεύρεσή» του καθυστέρησε ανάλογα. Τα ραδιοτηλεοπτικά συστήματα, από το ραδι-

252. ΣτΕ: Αναφέρεται στην αλλαγή του παραδείγματος της ενημέρωσης και της μαζικής επικοινωνίας από εκείνη όπου το μήνυμα μεταδίδεται «από έναν προς πολλούς» σε εκείνη όπου το μήνυμα μεταδίδεται «από πολλούς προς πολλούς» –και όπου, μάλιστα, υπάρχει η δυνατότητα της ασύγχρονης επικοινωνίας. Καταλυτικό ρόλο στην αλλαγή παραδείγματος διαδραμάτισαν τα μέσα κοινωνικής δικτύωσης, τα οποία αρχικά αποτέλεσαν και την κατεξοχήν περίπτωση διάδοσης του μηνύματος «από πολλούς προς πολλούς», χωρίς ιεραρχίες και χωρίς κεντρικό έλεγχο στη μετάδοση της πληροφορίας στον οποίο ο συγγραφέας αναφέρεται μιλώντας για «ακτινική μετάδοση» του μηνύματος.

όφωνο μέχρι την τηλεόραση, έγιναν κοινωνικά αναγκαία, υποστηρίζει ο Γουίλιαμς (1974), όταν αυτό που αποκαλεί «κινητή ιδιωτικοποίηση» της δυτικής κοινωνίας είχε προχωρήσει σε σημείο που ήταν δύσκολο για την κυβέρνηση και τη βιομηχανία να εντοπίσουν και να επικοινωνήσουν με τους πολίτες-καταναλωτές με άλλον τρόπο πέρα από την «οριζόντια» ακτινική μετάδοση μηνυμάτων[253] σε ολόκληρη την επικράτεια. Το συμπέρασμα για τη σημερινή μας κατάσταση είναι απλώς ότι δεν πρέπει να υποθέσουμε πως η πρόσφατη επανάσταση των δεδομένων στην ειδησεογραφία οφείλεται αποκλειστικά, ή ακόμη και κατά κύριο λόγο στις νέες τεχνολογίες και υποδομές, όπως το Διαδίκτυο. Θα πρέπει μάλλον να εξετάσουμε με προσοχή τον τρόπο με τον οποίο τα ειδησεογραφικά μέσα έχουν εξελιχθεί (και συνεχίζουν να εξελίσσονται) μέσα σε μια πιο σύνθετη οικολογία κοινωνικών δυνάμεων.

Η προσέγγιση του Γουίλιαμς διαμόρφωσε την έννοια του «ψηφιακού φιλελευθερισμού» που ανέπτυξα στο *Life Informatic*, για να αποτυπώσω ένα προαίσθημα που προέκυψε κατά τη διάρκεια της επιτόπιας έρευνάς μου με ειδησεογράφους στα τέλη της δεκαετίας του 2000 ότι υπήρχε μια συμβιωτική σχέση μεταξύ των ψηφιακών πρακτικών πληροφόρησης των δημοσιογράφων και της ευρύτερης νεοφιλελευθεροποίησης της κοινωνίας και της οικονομίας από τη δεκαετία του 1980. Με ενδιέφερε, για παράδειγμα, η αυξανόμενη σημασία της εργασίας μπροστά από μια οθόνη μεταξύ των ειδησεογράφων. Κάποιος θα μπορούσε να θεωρήσει την εργασία μπροστά σε μια οθόνη ηλεκτρονικού υπολογιστή ως προϋπόθεση υποδομής για την ανάπτυξη της δημοσιογραφίας δεδομένων. Η εργασία μπροστά σε οθόνη εμφανίστηκε για πρώτη φορά ως πτυχή της παραγωγής ειδήσεων στις δεκαετίες του 1970 και 1980, λόγω οργανωτικών πρωτοβουλιών στα ειδησεογραφικά μέσα και αλλού στη δυτική εταιρική κουλτούρα για την αξιοποίηση των προσωπικών υπολογιστών και των ψηφιακών πληροφοριακών συστημάτων γραφείου, προκειμένου να επιτευχθεί μια νέα αποδοτικότητα της παραγωγής. Στη βιομηχανία ειδήσεων, η αυτοματοποίηση αρχικά αντιμετωπίστηκε σαν ένα μέσο βελτίωσης της ταχύτητας επεξεργασίας κειμένου και μείωσης του εργατικού κόστους με τον αυτοματισμό στη σύνταξη και σε ορισμένες εργασίες επιμέλειας κειμένου. Όμως, με τη διαδικασία της θεσμοθέτησης, οι υπολογιστές ενσωματώθηκαν ταχύτατα σε κάθε πτυχή της παραγωγής ειδήσεων, από το μάρκετινγκ έως τη σελιδοποίηση και την αρχειοθέτηση, δημιουργώντας νέες ευκαιρίες για τον αυτοματισμό εργασιών που προηγουμένως εκτελούνταν απευθείας από ανθρώπους και για τη συγκέντρωση των υπόλοιπων εργασιών παραγωγής στα χέρια λιγότερων εργαζομένων στις ειδήσεις. Βετεράνοι δημοσιογράφοι που θυμούνται την εργασιακή ζωή πριν από την αυτοματοποίηση συχνά μου έλεγαν πόσο περισσότερο ήταν το βοηθητικό προσωπικό, πόσο περισσότερο χρόνο περνούν τώρα πια στα γραφεία τους και πόσο αυξήθηκε ο ατομικός φόρτος εργασίας τους.

Είναι αλήθεια ότι η ειδησεογραφία είχε πάντα την καθιστική της πλευρά. Η δακτυλογράφηση, για παράδειγμα, συνεπαγόταν επίσης καθιστική παραγωγή, όπως και η χρήση του τηλεφώνου πριν από τα κυψελοειδή συστήματα. Η ειδοποιός διαφορά ανάμεσα στις προηγούμενες μορφές καθιστικής δημοσιογραφίας και στη σύγχρονη εκδοχή της έγκειται στον τρόπο που η εργασία στην οθόνη οργανώνει σήμερα έναν άνευ προηγουμένου αριθμό βασικών δημοσιογραφικών εργασιών (π.χ. επεξεργασία κειμένου, επιμέλεια κειμένου, αρχειακή έρευνα, παρακολούθηση των έκτακτων ειδήσεων, παρακολούθηση του ανταγωνισμού, επικοινωνία και συντονισμός εντός του γραφείου) μέσω μιας ενιαίας διεπαφής με μια συνήθως σταθερή θέση. Η συνδυασμένη σημασία της χρήσης των έξυπνων κινητών τηλεφώνων και των γρήγορων μέσων κοινωνικής δικτύωσης, όπως το Twitter, για την ειδησεογραφία έχει καταστήσει την εργασία σε φορητές οθόνες τουλάχιστον εξίσου σημαντική με την εργασία σε επιτραπέζιες οθόνες,

253. ΣτΕ: Πρόκειται για το μοντέλο μετάδοσης της πληροφορίας «από έναν προς πολλούς», χαρακτηριστικό γνώρισμα των λεγόμενων «παραδοσιακών» μέσων μαζικής ενημέρωσης, όπως η τηλεόραση και το ραδιόφωνο: μάλιστα, στις περιπτώσεις αυτές, η κεντρική μετάδοση της πληροφορίας και η μαζική λήψη του εκπεμπόμενου μηνύματος ήταν, διαχρονικά, σύγχρονες διαδικασίες.

αλλά έχει κάνει ελάχιστα για να αλλάξει το φαινόμενο των δημοσιογράφων που είναι «κολλημένοι στις οθόνες τους». Λίγοι θα αμφισβητούσαν σήμερα ότι η οθόνη ως διεπαφή έχει γίνει μια κεντρική πτυχή της δημοσιογραφικής πρακτικής. Σχεδόν όλα όσα κάνουν οι δημοσιογράφοι, σχεδόν κάθε πηγή πληροφοριών, σχεδόν κάθε επαγγελματική παραγωγή περιλαμβάνει την ενασχόλησή τους με μία ή περισσότερες οθόνες.

Αυτή η συστέγαση κρίσιμων εργασιών δημιουργεί ευκολία, αλλά συνάμα αποσπά την προσοχή. Πολλοί δημοσιογράφοι αναφέρουν ότι νιώθουν εξουθενωμένοι από τον τεράστιο αριθμό και την ταχύτητα των ροών δεδομένων που πρέπει να διαχειριστούν. Είναι σημαντικό να αναγνωρίσουμε ότι η άσκηση της δημοσιογραφίας δεδομένων είναι συχνά αγχωτική. Στο πλαίσιο της έρευνας πεδίου που διεξήγαγα, οι δημοσιογράφοι που εργάζονταν σε οθόνη συχνά ανέφεραν ότι έπρεπε να βασίζονται σε άλλες αξιόπιστες πηγές ειδήσεων για να κρίνουν (για παράδειγμα, ως προς την αξία των ειδήσεων), επειδή οι ίδιοι ένιωθαν υπερβολικά εξαντλημένοι. Είναι εύκολο να καταλάβει κανείς πώς η εργασία σε οθόνη συμβάλλει στην «νοοτροπία της αγέλης» των σύγχρονων ειδήσεων που τόσο έχει κατηγορηθεί, με τους αφηρημένους, καταβεβλημένους δημοσιογράφους να βασίζονται συχνά ο ένας στον άλλον για ενημέρωση σε πραγματικό χρόνο, ενώ οι ροές δεδομένων κινούνται με ιλιγγιώδεις ρυθμούς.

Γνωρίζοντας ότι η κυριαρχία της οθόνης δεν προέκυψε εν κενώ, στο σημείο αυτό αποδεικνύεται γόνιμη μια παράλληλη διερεύνηση του νεοφιλελευθερισμού. Ο κλασικός φιλελευθερισμός δημιουργήθηκε τον 17ο και 18ο αιώνα, καθώς ο ευρωπαϊκός πνευματικός πολιτισμός προσαρμόστηκε στην πραγματικότητα του σχηματισμού αποικιακών αυτοκρατοριών σε ολόκληρο τον κόσμο. Η πολιτιστική κυριαρχία του μεσαιωνικού χριστιανικού συντηρητισμού, ακόμη και του ουμανισμού της Αναγέννησης, εκτοπίζονταν όλο και περισσότερο από κοινωνικές φιλοσοφίες που έδιναν έμφαση στην εργασία, στην ελευθερία, στην ατομική ιδιοκτησία και στην παραγωγικότητα. Ένα κρίσιμο πρόβλημα για τον πρώιμο φιλελευθερισμό ήταν πώς θα μπορούσε να μετατραπεί η επιδίωξη και η κατοχή της ατομικής ιδιοκτησίας σε μια ενάρετη διαδικασία, καθώς έμοιαζε να απειλεί να στερήσει από τους φτωχούς το μερίδιό τους από τα δώρα του Θεού προς την ανθρωπότητα. Η λύση ήταν να δοθεί έμφαση στο γεγονός ότι η ικανότητα της ανθρώπινης επιστήμης και της βιομηχανίας να βελτιώνουν την παραγωγική χρήση των πόρων, σε συνδυασμό με την απόλυτη αφθονία των νέων αποικιακών μεθορίων, σήμαινε πως η απόκτηση ατομικής ιδιοκτησίας δεν ήταν απαραίτητα αντίθετη προς τις χριστιανικές αξίες.

Μια ίσως ακούσια συνέπεια αυτού του νέου ηθικού μορφώματος ήταν η επικέντρωση της προσοχής στο άτομο ως υποκείμενο της λογικής, της δράσης, της ελευθερίας και της αρετής. Καθώς ο φιλελευθερισμός αναπτύχθηκε σε συνδυασμό με την άνοδο του καπιταλισμού και των σύγχρονων τρόπων ζωής του, το άτομο αναδείχθηκε σε μια όλο και πιο σημαντική φιγούρα του δυτικού πολιτισμού. Αρχικά, το άτομο προσπαθούσε να εξισορροπήσει αρμονικά τις περιοριστικές δυνάμεις της «κοινωνίας», αλλά ολοένα και περισσότερο η ατομικότητα εμφανιζόταν ως αυτοσκοπός, όπου όλες οι κοινωνικές και οικονομικές σχέσεις, τελικά, εξυπηρετούσαν την ανάπτυξη και την ανάδειξη εύρωστων, παραγωγικών, αυτοσυντηρούμενων ατόμων. Τα άτομα αυτά φάνταζαν ως ιδανικά απαλλαγμένα από τους κοινωνικούς προσδιορισμούς και, κατά συνέπεια, ελεύθερα να σκέφτονται και να ενεργούν όπως επιθυμούσαν. Περιγράφω αυτό το μοντέλο ατομικότητας, σύμφωνα με την ανθρωπολόγο Ελίζαμπεθ Ποβινέλι, ως «αυτολογικό» στην ιδεολογική παραδοχή του ότι τα άτομα είναι σε μεγάλο βαθμό ικανά να «αυτοδημιουργηθούν», μια έννοια που παραμένει αγαπητή σε όλο το φιλελεύθερο πολιτικό φάσμα σήμερα.

«Τι σχέση έχουν όλα αυτά με τις οθόνες των υπολογιστών;», θα με ρωτήσετε. Είναι αλήθεια ότι τα πρώτα μεγάλα εγχειρήματα στους αναλογικούς και ψηφιακούς υπολογισμούς πραγματοποιήθηκαν τη δεκαετία του 1930, καθώς οι κεϋνσιανές σοσιαλδημοκρατίες των μέσων του 20ού αιώνα προετοιμάζονταν για τον πόλεμο. Όμως, η ανάπτυξη του προσωπικού υπολογιστή, που ήταν ο πιο άμεσος πρόδρομος της σύγχρονης οθόνης, έγινε κατά τη διάρκεια των δεκαετιών του 1970 και 1980, την ίδια στιγμή που ο νεοφιλελευθερισμός κατέκτησε την πολιτική και φιλοσοφική κυ-

ριαρχία –καθώς η κεϋνσιανή σοσιαλδημοκρατία φαινόταν να καταρρέει υπό το βάρος των πολλαπλών γεωπολιτικών κρίσεων στα τέλη της δεκαετίας του 1960 και του 1970 (ο πόλεμος του Βιετνάμ, οι αραβοϊσραηλινές συγκρούσεις, ο σχηματισμός του ΟΠΕΚ[254] μεταξύ άλλων). Εκεί που ο φιλελευθερισμός πίστευε επί μακρόν ότι ο καλύτερος τρόπος για την επιδίωξη των δημόσιων συμφερόντων ήταν η ενδυνάμωση των ιδιωτικών συμφερόντων, θα μπορούσε κανείς να περιγράψει τον νεοφιλελευθερισμό ως ανελέητα αυταρχικό στην ενδυνάμωση των ιδιωτικών συμφερόντων εις βάρος των δημόσιων επενδύσεων και θεσμών.

Η νεοφιλελεύθερη στροφή στην πολιτική και τις πολιτικές είχε βαθύτατα αρνητικό αντίκτυπο στο είδος της ειδησεογραφίας δημοσίου συμφέροντος που συνόδευε τα κεϋνσιανά πρότυπα των μέσων του 20ού αιώνα, ακόμη και όταν προωθούσε μαζικές νέες επενδύσεις σε επικοινωνιακές και πληροφοριακές υποδομές, όπως το Διαδίκτυο, η δορυφορική μετάδοση και η κινητή τηλεφωνία σε ολόκληρο τον κόσμο. Η φαντασία και η δημιουργία αυτών των υποδομών είχε αρχικά ελάχιστη σχέση με τα ειδησεογραφικά μέσα. Το Διαδίκτυο, όπως είναι ευρέως γνωστό, δημιουργήθηκε χάρη στα κοινά συμφέροντα της στρατιωτικής άμυνας και της ερευνητικής επιστήμης. Λιγότερο γνωστή αλλά εξίσου σημαντική, ήταν η χρησιμότητα των άμεσων διακρατικών επικοινωνιών για χρηματοοικονομικές πρακτικές όπως το αρμπιτράζ. Ωστόσο, οι νέες υποδομές πληροφόρησης και επικοινωνίας επηρέασαν όλους τους τομείς της κοινωνικής επικοινωνίας, συμπεριλαμβανομένης, φυσικά, της παραγωγής ειδήσεων. Το καθαρό αποτέλεσμά τους ήταν η ριζική ενίσχυση της οριζόντιας ανταλλαγής μηνυμάτων από σημείο σε σημείο, καθώς και ο πλουραλισμός και ο εκ νέου εκσυγχρονισμός της κεντρικής μετάδοσης του μηνύματος «από έναν προς πολλούς», έτσι ώστε, μολονότι η ακτινική εκπομπή μηνύματος εξακολουθεί να υφίσταται, να γίνεται όλο και περισσότερο διακρατική και ασύγχρονη. Το μοντέλο του έθνους που κάθεται για να ακούσει συλλογικά τις βραδινές ειδήσεις απλά δεν υφίσταται με καμία πρακτική έννοια στα περισσότερα μέρη του κόσμου, ακόμη και στην Ευρώπη και στην Ασία, όπου διατηρούνται ισχυρότερες παραδόσεις δημόσιας ραδιοτηλεόρασης.

Η δική μας σύγχρονη οικολογία των ειδήσεων, στην πραγματικότητα, δεν εγγυάται έναν ισχυρό ατομικισμό, παρόλο που έχει καταστήσει την εύρεση αξιόπιστων πληροφοριών για την κοινότητα πιο επισφαλές εγχείρημα. Εκεί που τα πράγματα, όμως, σοβαρεύουν για τον ψηφιακό φιλελευθερισμό, θα λέγαμε, είναι στην εξατομικευμένη εμπειρία της εργασίας μπροστά στην οθόνη (και του παιχνιδιού μπροστά στην οθόνη, αν θέλετε). Η εξέλιξη των προσωπικών υπολογιστών, του Διαδικτύου και των μέσων κοινωνικής δικτύωσης διαμορφώθηκε βαθιά από την κοινωνική σημασία των (νεο)φιλελεύθερων αρχών της μεγιστοποίησης των ατομικών δυνατοτήτων για δράση, επικοινωνία και ιδεολογία.

Κατά την τελευταία δεκαετία, ένα αυξανόμενο ποσοστό του πληθυσμού (πάνω από το 70% στις Ηνωμένες Πολιτείες, για παράδειγμα) έχει μαζί του μια φορητή συσκευή πολυμέσων για όλες τις χρήσεις, η οποία λειτουργεί σαν πέμπτο μέλος του σώματος. Αυτό το επιπλέον μέλος επιτρέπει την πρόσβαση σε πολλαπλές ροές πληροφοριών, τη δυνατότητα διαμόρφωσης αυτών των ροών, ώστε να αντικατοπτρίζουν τα προσωπικά ενδιαφέροντα και τις επιθυμίες, καθώς και μυριάδες τρόπους, για να μεταδίδονται προσωπικές απόψεις και σκέψεις και να δημιουργούνται εγωκεντρικές μικροκοινότητες. Είναι τόσο η κληρονομιά αιώνων φιλελεύθερης επιστημολογίας όσο και το κρίσιμο μέσο για την αναπαραγωγή και την εντατικοποίηση αυτής της επιστημολογίας σε αυτό που έχουμε καταλήξει να αποκαλούμε «ψηφιακή εποχή». Έχετε δει τις εικόνες των αγνώστων σε ένα μπαρ ή σε ένα τρένο, που είναι όλοι κολλημένοι στις οθόνες τους. Το smartphone δεν εφηύρε, φυσικά, την κοινωνική αποξένωση. Αυτό που εφηύρε είναι μια επικοινωνιακή διεπαφή που μας επιτρέπει να βιώνουμε την ενεργό, παραγωγική ατομικότητα, ελαχιστοποιώντας παράλληλα την κοινωνική σύνδεση και την υπευθυνότητα, ακόμη και όταν συνωστιζόμαστε μεταξύ αγνώστων σε οποιοδήποτε μέρος του κόσμου. Σε άλλες εποχές, αυτοί

254. ΣτΜ: Οργανισμός Εξαγωγών Πετρελαιοπαραγωγών Χωρών (ΟΠΕΚ).

οι άγνωστοι θα μπορούσαν να αντιμετωπίσουν τη συνύπαρξή τους ως αφορμή και ευκαιρία, για να γίνουν αγενείς μεταξύ τους.

Εν ολίγοις, παραμένω πεπεισμένος ότι η αυτολογική ατομικότητα ενισχύεται από τον πολλαπλασιασμό και την εντατικοποίηση των διεπαφών οθόνης, ακόμη και αν το γεγονός ότι αυτές οι διεπαφές υπάρχουν εξαρχής έχει να κάνει σε μεγάλο βαθμό με τις τεχνολογίες που αναπτύχθηκαν, για να υλοποιήσουν φιλελεύθερες κοσμοθεωρίες και προτεραιότητες κατά τη διάρκεια των τελευταίων αιώνων. Παραφράζοντας τον Μάρσαλ Μακλούαν, νομίζουμε ότι δουλεύουμε με τις οθόνες μας, αλλά θα πρέπει να συνειδητοποιήσουμε ότι οι οθόνες μας, επίσης, δουλεύουν εμάς. Αυτό το σημείο συνάντησης των κινητών, φορητών μέσων που βασίζονται σε οθόνες και των φιλελεύθερων αντιλήψεων για την αυτολογική ατομικότητα είναι αυτό που ονομάζω «ψηφιακό φιλελευθερισμό» –και θα είναι ενδιαφέρον να δούμε πώς αυτός ο φιλελευθερισμός θα εξελιχθεί περαιτέρω στο μέλλον. Τι θα γινόταν, αν όλοι οι άγνωστοι σε εκείνο το βαγόνι του τρένου φορούσαν σετ εικονικής πραγματικότητας που τους επέτρεπαν μια εμβυθιστική πρόσβαση σε εικονικούς κόσμους; Πώς θα μπορούσαν αυτές οι νέες διεπαφές των μέσων ενημέρωσης να δημιουργήσουν νέους τύπους ατομικότητας και κοινωνικότητας; Αν και η δημοσιογραφία δεδομένων θεωρείται συχνά ύποπτη για σύνδεση με τις τεχνολογίες επιτήρησης και τον αλγοριθμικό αυταρχισμό, θα έλεγα ότι η εξέλιξη του ψηφιακού φιλελευθερισμού, στην πραγματικότητα, είναι η βαθύτερη ιστορία της δημοσιογραφίας δεδομένων.

Παραπομπές

Boyer, D. (2013). *The life informatic: Newsmaking in the digital era.* Ίθακα, Νέα Υόρκη: Cornell University Press.

Williams, R. (1974). *Television*. Μιντλτάουν, Κονέκτικατ: Wesleyan University Press.

Ο συγγραφέας

Ο Ντόμινικ Μπόγιερ είναι ανθρωπολόγος, κινηματογραφιστής και podcaster, ιδρυτής του Center for Energy and Environmental Research in the Human Sciences (CENHS) στο Rice University.

Ευρετήριο Όρων

#ddj	7.1
αδόμητα δεδομένα (unstructured data)	4.3
ακαδημαϊκή κοινότητα, συνεργασία με (academia, collaboration with)	3.5, 3.7, 8.2
ακρίβεια στην παρουσίαση δεδομένων (precision in data presentation)	5.8
αλγόριθμοι (alorithms)	
αριθμοί και (numbers and)	3.2
δημοσιογραφία και (journalism and)	4.7
διαχειριστικός έλεγχος (auditing)	6.1
δυσφήμιση (defamation)	6.1
κατάταξη (ranking)	6.4
λάθη σε (mistakes in)	6.1
λήψη αποφάσεων (decision making)	6.1, 6.5
λογοδοσία (accountability)	5.1, 6.1, 6.5, 8.1, 9.5
φυλετική προκατάληψη (racial bias)	6.1
αριθμοί και (numbers and)	3.2
δημοσιογραφία και (journalism and)	4.7
διαχειριστικός έλεγχος (auditing)	6.1
δυσφήμιση (defamation)	6.1
κατάταξη (ranking)	6.4
λάθη σε (mistakes in)	6.1
λήψη αποφάσεων (decision making)	6.1, 6.5
λογοδοσία (accountability)	5.1, 6.1, 6.5, 8.1, 9.5
φυλετική προκατάληψη (racial bias)	6.1
αναλογική συλλογή δεδομένων (analogue data collection)	8.5
ανάλυση κειμένου (text analysis)	4.3
ανάλυση περιεχομένου (content analysis)	7.2
ανεπεξέργαστων δεδομένων, δημοσίευση (raw data, publication of)	2.6, 4.5, 5.2, 5.6, 8.2
ανθρώπινες ιστορίες στα δεδομένα (human stories in data)	2.1, 2.3, 2.4, 5.8, 6.5
ανισότητα (inequality)	3.3, 4.1
ανοιχτά δεδομένα, κίνημα (open data, movement)	3.4, 7.3
ανοιχτή δημοσιογραφία (open journalism)	4.6, 4.8, 5.2, 7.5
ανοιχτού κώδικα, λογισμικό (open-source software)	2.5, 4.2, 5.4, 7.5, 8.1
αντίκτυπος (impact)	
ατομικός (individual)	9.3
εκπαιδευτικός (educational)	1.5, 3.5
εμπορικός (commercial)	3.4, 9.5
θεσμικός (institutional)	9.3, 9.5
κοινότητα, για την (community)	1.5, 3.4, 3.6, 8.3, 9.4
πολιτικός (political)	1.5, 3.4, 3.5, 6.5, 8.3, 9.1, 9.3, 9.4, 9.5
πολιτιστικός (cultural)	9.5

σχεδιασμός (planning for) 9.3
αξία οικοσυστημάτων (ecosystems value) 3.2
αποικιοκρατία (colonialism) 2.1, 3.3, 8.1, 9.4
αρχειοθέτηση δημοσιογραφίας δεδομένων (archiving data journalism) 7.2, 8.2
ασφάλεια και δεδομένα (security and data) 4.2
ατμοσφαιρική ρύπανση (air pollution) 3.4
αυτόχθονες πληθυσμοί και δεδομένα (Indigenous peoples and data) 3.3, 8.5, 9.4
βάσεις δεδομένων (databases) 3.1, 3.5, 4.1, 4.2, 4.4, 4.8, 5.6, 7.4, 9.1
βελτιστοποίηση μηχανής αναζήτησης (search engine optimization) 5.6
γενεαλογική ανάλυση (genealogical analysis) 9.1
Γενικός Κανονισμός για την Προστασία Δεδομένων της Ευρωπαϊκής Ένωσης (General Data Protection Regulation, GDPR) 3.3, 5.6
δεδομένα σε πραγματικό χρόνο (live data) 5.4
δεδομενοποίηση της κοινωνίας (datafication of society) 3.3, 4.1, 6.2, 7.3, 8.2, 9.2
Δείκτης Εντροπίας Πολλαπλών Ομάδων (Multigroup Entropy Index) 2.2
δημιουργία συνόλου δεδομένων (building a dataset) 2.4, 3.1, 3.5, 3.8, 4.2, 4.7
δημοσιεύσεις με τα δεδομένα στο επίκεντρο (data-centric publications) 5.5
διαδραστικότητα (interactivity) 2.2, 3.4, 5.2, 5.4, 5.5, 5.6, 5.7, 9.2, 9.3, 9.5, 9.6
διαμάχη για τη δημοσίευση δεδομένων (conflict over data publication) 2.3, 2.6
διατήρηση της δημοσιογραφίας δεδομένων (preservation of data journalism) 7.2, 7.3, 8.2
διατήρηση του περιβάλλοντος (environmental conservation) 3.2, 3.4
διδάσκοντας δημοσιογραφία δεδομένων (teaching data journalism) 3.6, 4.1, 8.2, 8.4
εγγράμματος προγραμματισμός (literate programming) 4.5
έγκλημα με μαχαίρια (knife crime) 3.1
εικονική πραγματικότητα (virtual reality, VR) 5.5, 7.4, 9.6
ελευθερία της πληροφόρησης (freedom of information, FOI) 2.3, 2.6, 3.1, 3.4, 4.7, 5.8, 7.4, 8.2
ελευθεροτυπία (freedom of press) 3.4, 3.8
εμπειρική δημοσιογραφία (empirical journalism) 5.1, 9.1
εμπιστοσύνη στη δημοσιογραφία (trust in journalism) 4.5, 4.6, 5.2, 5.3, 5.6, 7.3, 8.2
ενημερωτικά γραφήματα (news graphics) 5.5
ενημερωτικές εφαρμογές (news applications) 5.1
εξόρυξη δεδομένων (data mining) 4.3
επαυξημένη πραγματικότητα (augmented reality, AR) 5.4, 5.5, 7.3
επεξεργασία φυσικής γλώσσας (natural language processing, NLP) 4.3, 4.7
επισκεψιμότητα, διαδικτυακή (traffic, web) 5.6, 7.7
εργαλεία σχεδιασμού διαγραμμάτων (charting tools) 5.5
εργασία μπορστά στην οθόνη και ειδησεογράφοι (screenwork and news journalists) 9.7
έρευνες κοινού (surveys) 3.8, 9.3
εφαρμογές γνωριμιών (dating apps) 6.4
εφαρμογές και δεδομένα (apps and data) 6.4
ηθικοί προβληματισμοί (ethical considerations) 3.1, 5.6
ιστορία της δημοσιογραφίας δεδομένων (history of data journalism) 5.1, 7.3, 8.2, 9.1, 9.2, 9.6, 9.7
καθαρισμός δεδομένων (data cleaning) 2.5, 4.4, 4.5, 4.7
κάλυψη εκλογών (election coverage) 9.6

κατανεμημένη νόηση (distributed cognition) 4.8
Κατάσταση της Ένωσης, ομιλία (State of the Union, addresses) 4.3
κατάσταση του τομέα (state of the field) 1.2, 7.1, 7.3, 9.2, 9.6
κρυπτογράφηση (encryption) 4.2, 8.2
κυβερνητική διαφάνεια (government transparency) 3.4, 3.5, 3.8
κυριαρχία των αυτοχθόνων επί των δεδομένων τους (Indigenous data sovereignty) 3.3
λάθος των Ρέινχαρτ και Ρόγκοφ (Reinhart–Rogoff error) 4.5
μεθοδολογία δημοσίευσης (methodology, publication of) 3.7, 4.4, 4.6, 5.2, 5.3, 5.8, 8.2
μέλλον της δημοσιογραφίας δεδομένων (future of data journalism) 8.2, 9.2, 9.6
μέσα κοινωνικής δικτύωσης (social media)
ακτιβισμός (activism) 8.1
δεδομένα (data) 6.2, 6.4, 8.1
μετα-αλήθεια, στιγμή της (post-truth moment) 1.2, 4.6, 5.2
μηχανική μάθηση (machine learning) 4.3, 4.7, 5.1, 5.4, 7.3, 8.2
νεοφιλελευθερισμός (neoliberalism) 3.2, 9.7
οδική ασφάλεια (road safety) 2.5
οπτικοποίηση δεδομένων (data visualization) 3.2, 3.4, 4.1, 5.1, 5.2, 5.3, 5.4, 5.7, 6.2, 6.3, 6.4, 7.3, 8.2, 9.2
αίθουσες σύνταξεις και (newsrooms and) 5.2
αλληλεπίδραση με (engagement with) 5.2, 9.5
αναλογική (analogue) 8.5
αρχειοθέτηση (archiving of) 7.2
Ιστός και (web, the, and) 5.4
τρισδιάστατα γραφικά (3D graphics) 5.4
ορισμός της δημοσιογραφίας δεδομένων (definition of data journalism) 1.2, 5.1, 5.2, 5.3, 7.1, 9.2, 9.3
ουδετερότητα των δεδομένων (neutrality of data) 1.2, 5.1, 5.4, 5.6, 9.1
παιχνίδια (games) 5.4
παραγωγή φυσικής γλώσσας (natural language generation, NLG) 4.8
παρακολούθηση (surveillance)
από εταιρείες τεχνολογίας (by technology companies) 1.2, 7.5
από κράτη (by states) 1.2, 3.3
πηγές (sources) 4.5, 8.2
πλατφορμοποίηση της δημοσιογραφίας δεδομένων (platformization of data journalism) 7.6
πληθοπορισμός (crowdsourcing) 2.3, 3.7, 5.2, 5.6, 6.1, 6.5, 7.3, 8.2, 9.2
προεδρικές εκλογές ΗΠΑ, 2016 (US presidential election 2016) 6.3, 8.1, 9.6
ντιμπέιτ/αντιλογία δημοκρατικών (Democratic Party debate) 4.3
προστασία δημοσιογραφικών δεδομένων (journalism data protection) 3.3, 6.5
πρώτα για κινητές συσκευές (mobile first) 3.4, 5.2, 5.4, 5.5, 7.3, 7.5, 8.2
ρεπορτάζ για εγκλήματα (crime reporting) 3.1, 4.1, 4.7, 5.1, 7.2, 9.4
Ρωσικές εκστρατείες παραπληροφόρησης (Russian disinformation campaigns) 6.3
σεξουαλικές επιθέσεις σε πανεπιστημιουπόλεις (sexual assault on college campuses) 5.1
σκιτσογραφία με δεδομένα (data sketching) 5.3
συγρούσεις για τη γη (land conflicts) 3.5

συναίσθημα και δεδομένα (emotion and data) 4.1, 5.2, 7.7
συνεργατική δημοσιογραφία (collaborative journalism) 4.2, 5.6, 6.2, 6.5, 7.3, 7.5, 8.2, 9.2, 9.5
σύστημα διαχείρισης περιεχομένου (content management system) 5.6, 7.2
σχεδιάζοντας ένα έργο (storyboarding a project) 2.4, 3.1
τεχνητή νοημοσύνη (artificial intelligence) 3.6, 4.8, 5.1, 7.3
τεχνολογία των πολιτών και δημοσιογραφία δεδομένων (civic tech and data journalism) 7.4, 7.5
τηλεόραση, δημοσιογραφία δεδομένων και (television, data journalism and) 5.8
τρολάρισμα (trolling) 6.2, 7.7
υπολογιστικά φύλλα (spreadsheets) 4.4, 4.5, 5.6, 8.4
υπολογιστική δημοσιογραφία (computational journalism) 4.3, 4.4, 4.6, 4.7, 4.8, 6.1, 9.1
φεμινισμός (feminism) 4.1
φυλετικός διαχωρισμός (racial segregation) 2.2
χάρτες (maps)
ως δεδομένα (as data) 2.1, 2.2
ως οπτικοποιήσεις (as visualizations) 2.1, 2.2, 2.3, 2.4, 2.5, 3.5, 5.4, 8.2, 9.4, 9.6
ψηφιακή εγκληματολογία (digital forensics) 6.3
ψηφιακή έρευνα (digital research) 6.4
ψηφιακή οικουμενικότητα (digital universalism) 8.1
ai2html 5.4
API (Διεπαφή Προγραμματισμού Εφαρμογών –Application Programming Interface) 3.1, 6.2, 6.4
bot 6.2
cyborg 6.2
Facebook
αλγόριθμος «Άτομα που Ίσως Γνωρίζετε» («People You May Know» algorithm) 6.1
Αυτόχθονες χρήστες (Indigenous users) 9.4
δεδομένα χρήστη (user data) 5.1, 6.2
διαφημίσεις (ads) 5.1, 6.2, 6.5, 9.5
έλεγχος γεγονότων (fact-checking) 4.8
περιορισμοί πρόσβασης στα δεδομένα (data access restrictions) 6.2, 6.4
υποδομές δεδομένων (data infrastructure) 6.4
ψεύτικοι ακόλουθοι (fake followers) 6.4
ψεύτικοι λογαριασμοί (fake accounts) 6.4
Google
αλγόριθμος αυτόματης συμπλήρωσης (autocomplete algorithm) 6.1, 6.5
αποτελέσματα αναζήτησης (search results) 6.1, 6.5
hashtag για τη δημοσιογραφία δεδομένων (hashtags for data journalism) 7.1
metrics (μετρήσεις) 7.7, 8.2, 9.3, 9.5
newsgame 5.4
scraping 4.4, 6.2
scrollytelling 2.1, 5.2, 5.4
TF-IDF (Term Frequency-Inverse Document Frequency –Συχνότητα Όρων/Αντίστροφη Συχνότητα Εγγράφων) 4.3

Ευρετήριο Ονομάτων

Ανεξάρτητη Αρχή Κοινοβουλευτικών Προτύπων (Independent Parliamentary Standards Authority, IPSA) 9.5
Αραβική Άνοιξη (Arab Spring) 8.1
Αυστραλία (Australia) 3.2, 3.3
Αφγανιστάν (Afganistan) 8.3
Βραβεία Δημοσιογραφίας Δεδομένων (Data Journalism Awards) 2.6, 4.3, 5.1, 7.3, 9.2
Βραζιλία (Brazil) 4.6
Γιάου, Νέιθαν (Yau, Nathan) 5.4, 7.1
Γκρίλεϊ, Χόρας (Greeley, Horace) 8.2
Γουατεμάλα (Guatemala) 2.1
Δημόσια Ραδιοτηλεόραση της Αυστραλίας (Australian Broadcasting Corporation) 5.4
Διεθνής Σύμπραξη Ερευνητών Δημοσιογράφων (International Consortium of Investigative Journalists, ICIJ) 4.2, 5.6, 7.6, 9.2, 9.5
Ζι, Σάρα (Sze, Sarah) 1.2
Ινδία (India) 3.5
Ισραήλ (Israel) 2.4
Κένυα (Kenya) 8.3
Κίνα (China) 3.4
Κλίντον, Χίλαρι (Clinton, Hillary) 4.3, 9.6
Κολομβία (Colombia) 2.1, 2.3, 8.5
Κούβα (Cuba) 3.6
Λέσλι, Φρανκ (Leslie, Frank) 5.1
Μάγιερ, Φίλιπ (Meyer, Philip) 5.1, 7.3, 9.1
Μαορί, δεδομένα των (Māori data) 3.3
Μέι, Τερέζα (May, Theresa) 4.8
Μπογκοτά (Bogotá) 2.3
Νάιτινγκεϊλ, Φλόρενς (Nightingale, Florence) 8.2
Νέα Ζηλανδία (New Zeland) 3.3
Ντέντμαν, Μπιλ (Dedman, Bill) 9.5
Ντόιγκ, Στιβ (Doig, Steve) 7.2
Ονδούρα (Honduras) 2.1, 8.5
Οργανισμός Ηνωμένων Εθνών (United Nations) 2.1, 2.4
Παγκόσμιο Δίκτυο Ερευνητικής Δημοσιογραφίας (Global Investigative Journalism Network, GIJN) 7.1, 9.6
Πακιστάν (Pakistan) 8.3
Παλαιστίνη (Palaistine) 2.4
Παπούα-Νέα Γουϊνέα (Papua New Guinea) 4.5
Περού (Peru) 5.7, 8.1
Πρώτα Έθνη (First Nations) 3.3, 8.1
Ράμσφελντ, Ντόναλντ (Rumsfeld, Donald) 5.8

Σίλβερ, Νέιτ (Silver, Nate) 1.2, 5.1, 7.1, 7.3, 9.5, 9.6
Σνόου, Τζον (Snow, John) 8.2
Σνόουντεν, Έντουαρντ (Snowden, Edward) 1.2, 7.3
Τουρκία (Turkey) 2.6
Τραμπ, Ντόναλντ (Trump, Donald) 1.2, 4.3, 6.2, 9.1
Τσαλαμπί, Μόνα (Chalabi, Mona) 5.3
Φιλιππίνες (Philippines) 2.5
Χολμς, Νάιτζελ (Holmes, Nigel) 4.1
Χονγκ Κονγκ (Hong Kong) 6.2, 8.1
ACT NOW! 4.5
AdCollector 6.5
AlgorithmWatch 6.1, 6.5
Alibaba 3.4
Ampp3d 5.4
Anatomy of a Killing 5.1
Associated Press 4.8
Atlanta Journal-Constitution, The 4.7
BBC 4.4, 5.4, 9.5, 9.6
Belize 8.5
Bellingcat 1.2
Berliner Morgenpost 5.6, 7.2
Blockchain 8.2
Brexit 1.2, 6.4, 8.1, 9.1
Broken Homes 2.4
Bureau Local 9.5
Bureau of Investigative Journalism 5.6, 9.5
BuzzFeed News 1.2, 4.5, 4.6, 5.1, 6.2, 7.3, 7.4
California Civic Data Coalition 7.5
Cambridge Analytica 6.2, 6.4, 8.1
Center for Investigative Reporting 9.3
Chartbeat 7.7, 8.2
chicagocrime.org 9.4
Chicas Poderosas 8.5
China Judgments Online 3.4
Cicada Tracker 9.3
Correctiv 4.5, 5.6, 7.4
Counted, The 5.6
Dağ Medya 2.6
Data Journalism Handbook, The (2012) 1.2, 7.2, 9.6
Datablog 5.1, 7.3
Dear Mona 5.3
Dedupe.io 4.7
Depth of the Problem, The 5.4, 8.2
Diario, El 2.1
Dollars for Docs 5.6, 8.2, 9.2
Drone Warfare, βάση δεδομένων 5.6, 9.5

Economist, The 4.3, 4.4
EcoTender 3.2
Electionland 7.3
Enslaved Land 2.1
EnSym 3.2
Euros für Ärzte 5.6
Excel 2.6, 4.4, 4.5
Facebook 4.1, 5.1, 6.1, 6.2, 6.4, 6.5, 9.4
Faro, El 2.1
Financial Times, The 4.5, 5.5, 7.1
FiveThirtyEight 1.2, 4.3, 4.6, 5.1, 5.3, 5.4, 5.5, 7.1, 7.3, 8.2, 9.1, 9.6
Fixed Points Finding a Home 1.2
Gawker Media 7.7
Gênero e Número 4.6
GitHub 3.6, 4.5, 4.6, 8.4
Global I-Hub 4.2, 7.6
Global Witness 4.5
Google
- Alerts 2.6, 3.1, 9.3
- Analytics 6.3, 8.2
- Cloud Natural Language API (API του Google Cloud για επεξεργασία φυσικής γλώσσας) 4.3
- Earth 5.1
- Fusion Tables 4.8, 7.3
- News Lab 1.2, 3.4, 7.3
- Sheets 4.4, 5.6, 7.3

GreenSock Animation Platform 5.7
Guardian, The 3.1, 3.7, 4.3, 4.5, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 7.3, 8.2, 9.6
Guerra por el agua, La (Ο πόλεμος για το νερό) 5.7
Homicide Report 9.4
How to Win a Trade War 5.1
INDVIL 5.1
Internet Archive, The 7.2
JavaScript 4.4, 5.4, 5.6, 5.7, 7.2, 7.5
Jazeera, Al 2.4, 6.2
Jupyter Notebook 4.4, 4.5, 4.8, 8.4
Klaxon 7.5
Land Conflict Watch 3.5
License to Betray 4.7
Los Angeles Times 4.3, 4.6, 9.4
Machine Bias 6.1, 9.5
Maiam nayri Wingara Indigenous Data Sovereignty Collective 3.3
Markup, The 5.1, 6.5
Marshall Project, The 7.5

Migrants' Files, The 9.5
Monstrous Costs 4.1
Mozilla Festival 1.2, 7.3, 9.6
MPs' expenses scandal (Το σκάνδαλο για τα έξοδα των βουλευτών, κάλυψη Guardian) 7.3, 8.2, 9.5, 9.6
Murder Mysteries 7.2
Nación, La 3.7, 7.3
National Geographic 5.5
Nauru Files 5.6
Neue Zürcher Zeitung 5.5
New York Times, The 4.5, 4.6, 5.1, 5.2, 5.4, 5.5, 7.1, 7.2, 7.3, 7.5, 7.7, 8.4, 9.5, 9.6
NodeXL 7.1
NOS 5.8
OCAP 3.3
Offshore Leaks 4.2, 5.6, 9.4
OjoPúblico 5.7
Onion, The 7.7
Open Database of Deceased Workers in Turkey (Ανοιχτή Βάση Δεδομένων για τους Αποβιώσαντες Εργαζομένους στην Τουρκία) 2.6
Open Knowledge Foundation, OKFN 6.5, 7.3, 7.4
openDemocracy 3.8
OpenNews 3.7, 7.2
OpenRefine 2.5, 4.7, 7.3
Organized Crime and Corruption Reporting Project, OCCRP 5.6, 7.6
Our World in Data 5.5
Pajhwok Afghan News 8.3
Panama Papers 1.2, 2.1, 4.2, 5.6, 7.3, 7.6, 8.2, 9.2, 9.5
Paradise Papers 4.2, 5.6, 8.2, 9.8
PGP 4.2
PM2.5 3.4
Postdata.club 3.6
Pourover 7.5
ProPublica 3.7, 4.4, 4.6, 5.1, 5.4, 5.6, 6.1, 6.5, 7.3, 7.4, 8.2, 9.2, 9.4
Pudding, The 5.5, 7.1
Python 4.3, 4.4, 4.5, 4.7, 4.8, 8.4
Qing Yue 3.4
Quartz 5.4, 8.2
R 4.3, 4.5
RADAR 4.8
Rappler 2.5
React 4.4
Reading the Riots 7.3
Refine 4.4
ReproZip 7.2

Right of Way 2.5
Rising Lab 3.4
Ruby 4.4
Schufa 6.5
Seeing Data 5.2
Snow Fall 9.6
Spiegel, Der 5.6, 6.5
Spotlight 9.5
Sunday Times, The 4.4
Te Mana Raraunga 3.3
Telegraph, The 9.5
Thomson Reuters, Ίδρυμα (Thomson Reuters Foundation) 3.8
Times, The 4.4, 4.7
Twitter 4.4, 5.1, 5.2, 6.2, 6.4, 9.7
Uber 4.6, 6.1
Under the Dome 3.4
Universal, El 7.3, 9.5
Upshot, The 5.1, 7.1, 7.3, 9.6
Vox 4.3, 5.4
Vozdata 3.7
Wall Street Journal, The 4.8, 5.4
Washington Post, The 2.2, 4.3, 4.6, 9.3, 9.7
WeChat 3.4
WikiLeaks 7.3
Workbench 4.5, 4.7
Years You Have Left to Live 5.4
YouTube 5.1, 6.1, 6.2
Zones of Silence 7.3

Το Εγχειρίδιο της Δημοσιογραφίας Δεδομένων: Προς μια Κριτική Πρακτική Δεδομένων αποτελεί μια πλούσια και πανοραμική εισαγωγή στη δημοσιογραφία δεδομένων, που συνδυάζει τόσο τον κριτικό προβληματισμό όσο και την πρακτική διορατικότητα. Προσφέρει μια ετερόκλητη συλλογή από απόψεις για τον τρόπο με τον οποίο ασκείται η δημοσιογραφία δεδομένων σε ολόκληρο τον κόσμο και για τις ευρύτερες συνέπειες της «δεδομενοποίησης» των ειδήσεων, με αποτέλεσμα να είναι χρήσιμο τόσο ως εγχειρίδιο όσο και ως βιβλίο-πηγή για αυτόν τον αναδυόμενο τομέα. Με περισσότερα από 50 κεφάλαια γραμμένα από κορυφαίους ερευνητές και επαγγελματίες της δημοσιογραφίας δεδομένων, διερευνά το έργο που απαιτείται για να καταστούν οι τεχνολογίες και τα δεδομένα παραγωγικά για δημοσιογραφικούς σκοπούς. Παρουσιάζει, επίσης, μια «εσωτερική ματιά» στην κοινωνική ζωή των συνόλων δεδομένων, των υποδομών δεδομένων και των ιστοριών δεδομένων, στις αίθουσες σύνταξης, στους οργανισμούς μέσων ενημέρωσης, στις start-up επιχειρήσεις, στις οργανώσεις της κοινωνίας των πολιτών και όχι μόνο. Στο βιβλίο θα βρείτε τις ενότητες «Θέματα με Δεδομένα», «Συγκεντρώνοντας τα Δεδομένα», «Δουλεύοντας με Δεδομένα», «Η Εμπειρία της Εργασίας με Δεδομένα», «Διερευνώντας τα Δεδομένα, τις Πλατφόρμες και τους Αλγορίθμους», «Οργάνωση της Δημοσιογραφίας Δεδομένων», «Μαθαίνοντας τη Δημοσιογραφία Δεδομένων από Κοινού» και «Χωροθέτηση της Δημοσιογραφίας Δεδομένων».

Ο Τζόναθαν Γκρέι είναι Λέκτορας Σπουδών Κρίσιμων Υποδομών στο Τμήμα Ψηφιακών Ανθρωπιστικών Επιστημών στο King's College του Λονδίνου, συνιδρυτής του Public Data Lab και ερευνητικός συνεργάτης του Digital Methods Initiative στο Πανεπιστήμιο του Άμστερνταμ.

Η Λιλιάνα Μπούνεγκρου είναι Λέκτορας Ψηφιακών Μεθόδων στο Τμήμα Ψηφιακών Ανθρωπιστικών Επιστημών στο King's College του Λονδίνου, συνιδρύτρια του Public Data Lab και ερευνητική συνεργάτιδα του Digital Methods Initiative στο Πανεπιστήμιο του Άμστερνταμ.

Οι δυο τους συνέβαλαν εξίσου στην επιμέλεια του πρωτότυπου έργου.

ISBN 978-618-85968-1-8