Amsterdam University Press

Chapter Title: The Datafication of Journalism: Strategies for Data-Driven Storytelling and Industry—Academy Collaboration
Chapter Author(s): Damian Radcliffe and Seth C. Lewis

Book Title: The Data Journalism Handbook
Book Subtitle: Towards A Critical Data Practice
Book Editor(s): Liliana Bounegru, Jonathan Gray
Published by: Amsterdam University Press. (2021)
Stable URL: https://www.jstor.org/stable/j.ctv1qr6smr.47

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

This book is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0). To view a copy of this license, visit https://creativecommons.org/licenses/by-sa/4.0/.

Amsterdam University Press is collaborating with JSTOR to digitize, preserve and extend access to *The Data Journalism Handbook*

Electronic copy available at: https://ssrn.com/abstract=3907053
This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms

44. The Datafication of Journalism: Strategies for Data-Driven Storytelling and Industry–Academy Collaboration

Damian Radcliffe and Seth C. Lewis

Abstract

How are journalism and academia responding to the datafication of their professions, and how can they collaborate more effectively on data-driven work?

Keywords: journalism, academia, collaboration, datafication, data work, researcher–journalist collaborations

We live in a world driven and informed by data. Data increasingly influences how policy and political decisions are made (Höchtl et al., 2016; Kreiss, 2016), informs the design and functionality of the cities we live in (Glaeser et al., 2018), as well as shapes the types of news, products and information that we have access to—and consume—in the digital age (Diakopoulos, 2019; Lewis, 2017; Lewis & Westlund, 2015; Thurman et al., 2019; Usher, 2016; Zamith, 2018). The full power and potential of data, for good or ill, is only just beginning to be realized (Couldry & Mejias, 2019; O'Neil, 2016; Schroeder, 2018).

Governments, universities and news media have long made use of data and statistics to find patterns and explain the world. But with the growth in digital devices and the massive trace data they produce—about our clicks, likes, shares, locations, contacts and more—the sheer volume of data generated, as well as the increase in computing power to harness and analyze such data at scale, is staggering. Making sense of all that data, in many cases, is arguably the biggest challenge, and is deeply fraught with ethical determinations along the way (Crawford et al., 2014). It is a riddle that

Bounegru, L. and J. Gray (eds.), *The Data Journalism Handbook: Towards a Critical Data Practice*. Amsterdam: Amsterdam University Press, 2021
DOI 10.5117/9789462989511_CH44

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

policy makers, businesses, researchers, activists, journalists and others are contending with—and one that will not be so easily resolved by "big-data solutions" or, in vogue today, the glittering promise of artificial intelligence (Broussard, 2018; Broussard et al., 2019).

In this chapter, building on our respective observations of practice (Radcliffe) and research (Lewis) regarding data and journalism, we outline how the worlds of journalism and academia are responding to the datafication of their professions as well as the broader datafication of public life. Ultimately, our aim is to offer recommendations for how these two fields, which historically have shared a rather uneasy relationship (Carlson & Lewis, 2019; Reese, 1999), might more productively work together on data-centric challenges.

The poet John Donne wrote that "no man is an island." In a data-driven world, no profession should be either.

Journalism and Data-Driven Storytelling: Five Strategic Considerations

The use of data to tell stories, and make sense of the world around us, is not wholly new.[1]

In Victorian England, physician John Snow produced a map that plotted cholera cases in central London. It enabled him to identify a pump on Broad Street as the cause of a particularly fatal, and geographically focused, outbreak of the disease in 1855 (see Figures 44.1 and 44.2). Snow's influential analysis does not look too dissimilar from disease maps produced with modern tools of data analysis and visualization.

In another example, Florence Nightingale's visualizations "of the causes of mortality in the army in the East" ("Worth a Thousand Words," 2013) helped to demonstrate the role that sanitation (or lack thereof) played in causing the death of British soldiers fighting in the Crimean War (see Figure 44.3). Her designs still feel remarkably contemporary.

Alongside these efforts, around the same time, Horace Greeley's work for *The New York Tribune* in the mid-19th century exposed how a number of elected officials (including a young Abraham Lincoln) were claiming expenses greater than they were eligible for (Klein, 2015). Although the world has moved on (Greeley's work focused on distances typically travelled by horseback), this type of important investigative work continues to be a journalistic staple (Larcinese & Sircar, 2017; see also Barrett, 2009; "A

1 See Anderson's chapter in this book for a look at different genealogies of data journalism.

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

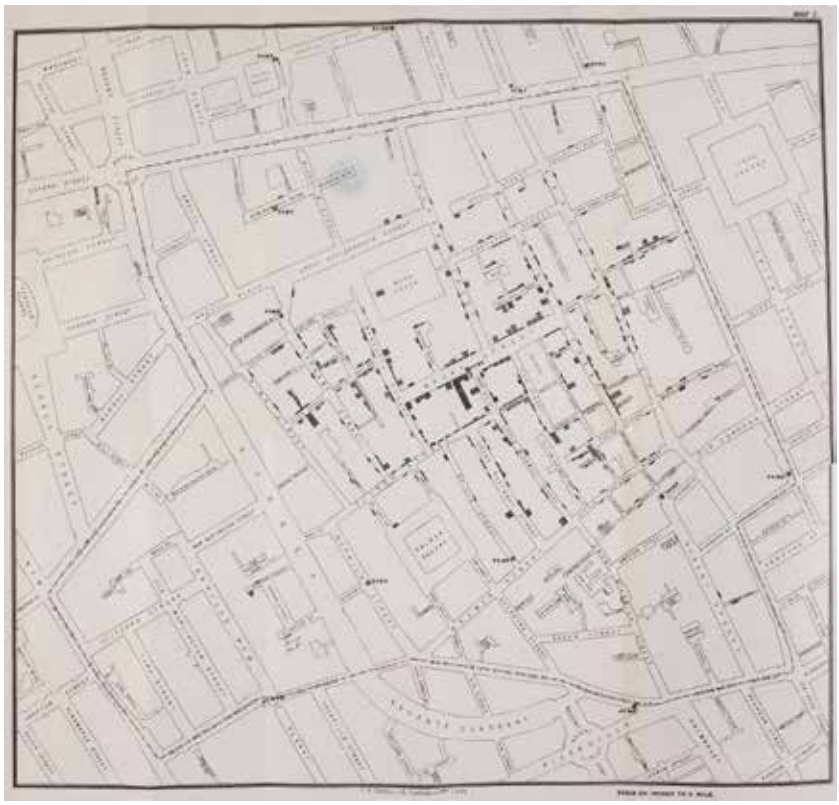


Figure 44.1. Map of London produced by physician John Snow, plotting cholera cases in central London in 1855. Source: British Library. https://www.bl.uk/collection-items/john-snows-map-showing-the-spread-of-cholera-in-soho-london-1855.

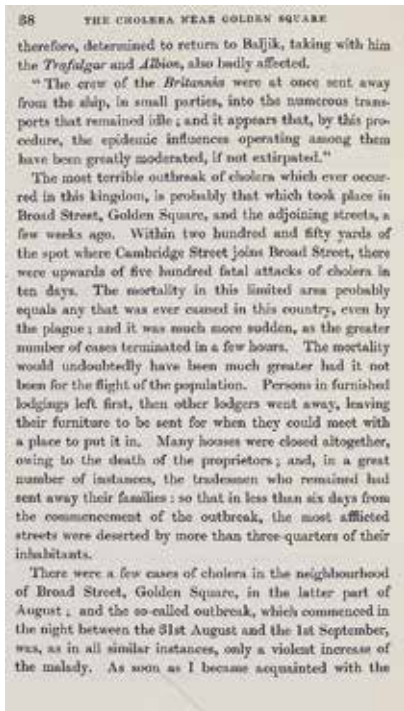
Chronology of the Senate Expenses Scandal," 2016; "Expenses Scandal an Embarrassing Start," 2017; "MPs' Expenses: Full List," 2009; "Q&A: MP Expenses," 2009; Rayner, 2009; "Senate Expenses Scandal," n.d.).

These historic examples, coupled with more contemporary case studies (such as those identified by the annual Data Journalism Awards), can act as powerful sources of inspiration for journalists.[2] They demonstrate how data-driven approaches may be used to hold authority to account (ICIJ, n.d.), highlight important social injustices (Lowenstein, 2015), as well as visualize and showcase the extraordinary ("2016 Year in Review," 2016).

While data has long been a part of journalism, as reflected in the emergence of "computer-assisted reporting" during the late 20th century, recent

2 https://www.datajournalismawards.org/. See also Loosen's discussion of the awards in her chapter in this book.

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053



38 THE CHOLERA NEAR GOLDEN SQUARE

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

Figure 44.2. Text of an 1855 newspaper story documenting cholera cases in central London. Source: British Library. https://www.bl.uk/learning/images/makeanimpact/publichealth/large12734.html

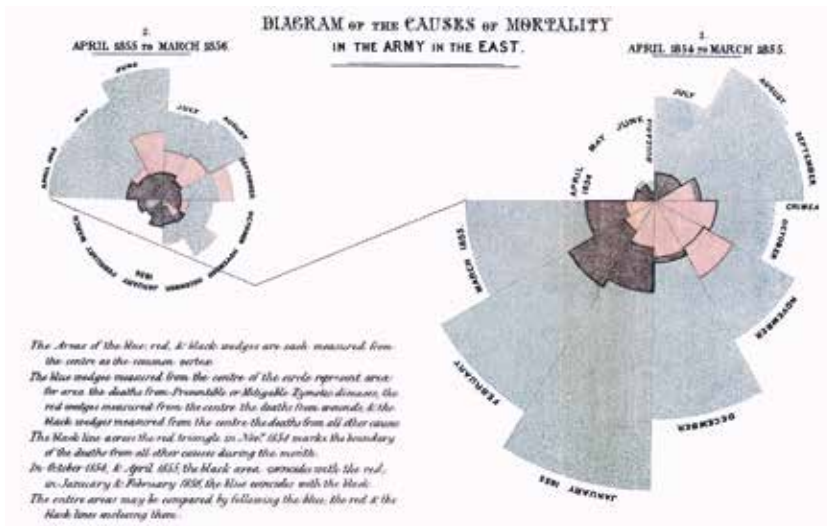


Figure 44.3. "Diagram of the causes of mortality in the army in the East," by Florence Nightingale. Source: Wikimedia. https://upload.wikimedia.org/wikipedia/commons/1/17/Nightingale-mortality.jpg

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

developments in the availability and accessibility of data-driven techniques have amplified opportunities for distinctly data-driven journalism (for a history, see Anderson, 2018; for an overview of data journalism, see Gray et al., 2012). It is against this backdrop that news organizations around the world—particularly the best-resourced ones, but increasingly smaller newsrooms as well—are using data to inform their journalistic work, both in telling stories with data (Hermida & Young, 2019) as well as in using data (in the form of digital audience metrics) to influence story selection as well as to measure and improve the impact of their work (Zamith, 2018).

Here are five key messages for newsrooms and journalists looking to do more with data:

Data alone does not tell stories. We still need journalists. For all of the data we have access to, we still need journalists to make sense of it, by providing context and interrogating the data in the same way as any other source.

As Steve Johnson (2013) of Montclair State University has noted: "Readers don't care about the raw data. They want the story within the data." Commenting on data about lower Manhattan provided by an early open-data portal, EveryBlock, he observed:

> There were reports on what graffiti the city said it had erased each month, by neighborhoods. But what was missing was context, and photos. If I'm a reporter doing a story on graffiti, I want to show before and after photos, AND, more importantly, I want to know whether the city is successfully fighting the graffiti artists, i.e., who is winning. The raw data didn't provide that. (Johnson, 2013)

More recent "data dumps" such as the Paradise Papers and Panama Papers also emphasize this point. In this instance, sources had to be cross-referenced and contextualized—a time-consuming process that took many journalists months to do. However, without this interrogation of the sources by journalists (as opposed to concerned citizens), the full impact of the data could not be realized. These principles are as applicable at the local level as they are in stories of national and international import (Radcliffe, 2013).

Data, in itself, is seldom the story. It needs to be unpacked and its implications explained, if the full meaning behind it is to be understood.

You don't have to go it alone. Collaboration is often key. Collaboration has been a watchword of the networked age, and a key element in the ongoing blending of journalism and technology sensibilities—including the integration of "hacks" (journalists) and "hackers" (coders) (Lewis & Usher, 2014, 2016) as well as the broader interplay of news organizations and their

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

communities around shared concerns. The essence of such "networked journalism" (Beckett, 2010; Van der Haak et al., 2012) or "relational journalism" (Boczkowski & Lewis, 2018; Lewis, 2019) is the underlying belief that more might be accomplished through cooperative activity.

This approach is applicable to many beats and stories, including those involving large volumes of data. As *The Guardian* showed in their 2009 analysis of British MPs' expenses, concerned citizens and members of the public can work in tandem with journalists to analyze data sets and provide tips ("MPs' Expenses: *The Guardian*," 2009; Rogers, 2009a, 2009b). More recently, research by Stanford's James T. Hamilton (2016) and others (Sambrook, 2018) has identified the importance of collaboration—both across organizations and in the deployment of different disciplines—for many newsrooms, when it comes to producing high-quality, high-impact investigative journalism.

The amount of data that many new organizations are contending with, coupled with ongoing challenging economic circumstances, means that partnerships, the use of specialists, volunteers and wide-ranging skill sets, are often a necessity for many newsrooms. And, a collaborative approach is increasingly essential from both a financial and journalistic standpoint.

How you present your data matters. Journalists have access to a wide range of tools, techniques and platforms through which to present data and tell stories accordingly.[3] As a result, it is important to determine which tools are most appropriate for the story you are trying to tell.

Data visualizations, graphs, infographics, charts, tables and interactives—all can help to convey and drive home a story. But which one (or ones) you use can make all the difference.[4]

As our colleague Nicole Dahmen has noted, one way to do this is through data visualization. "Visuals catch audience attention . . . [and] . . . are processed 60 times faster than text" (as cited in Frank et al., 2015). When used well, they can help to bring a story alive in a manner that text alone cannot.

The Washington Post's online feature "The Depth of the Problem," which shows how deep in the ocean the black box from the missing Malaysia Airlines flight 370 could be, is a good example of this ("The Depth of the Problem," n.d.; see Figure 26.5). The reader scrolls down the page to 1,250 feet, the height of the Empire State Building; past 2,600 feet, the depth of giant squids; and below 12,500 feet, where the *Titanic* sits; to 15,000 feet, where the black box was believed to be.

3 https://www.journaliststoolbox.org/2018/03/11/online_journalism/

4 See https://www.import.io/post/8-fantastic-examples-of-data-journalism/ for eight stories which use different techniques and consider swapping them.

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

"You're not just reading how deep that plane is," Dahmen has said. "You can see and engage and really experience and understand how deep they suspect that plane to be."

Determining your approach may be influenced by both the story you want to tell and the data literacy and preferences of your audience. Either way, your data-driven stories should be well designed so that audiences do not struggle to understand what is being shown or how to interact with the data (Radcliffe, 2017b, 2017c).

Place your work in a wider context. Alongside these considerations, journalists working with data also need to be cognizant of wider developments, in terms of the consumption of content and attitudes towards journalism.

Think mobile: In 2012, the Pew Research Center found that over half of those who owned a smartphone or a tablet accessed news content on those devices (Mitchell et al., 2012; "News Consumption on Digital Devices," 2017); just four years later, more than seven in ten Americans (72%) reported getting news via mobile (Mitchell et al., 2016). As mobile news consumption continues to grow, so too it is imperative that news organizations provide a positive mobile experience for all of their content, including data-rich material.

Make it personal: In an era of personalization and algorithmically generated media experiences, this can include creating opportunities for audiences to interrogate data and understand what it means for them. *ProPublica*'s Dollars for Docs investigation (Tigas et al., 2019), which enables patients to see the payments made by pharmaceutical and medical device companies to their doctors, is one example of this technique in action.

Protecting your sources: Journalists need to know how to protect data as well as how to analyze it. Protecting yourself, and your sources, may well require a new approach—including new skill sets—to handling sensitive data and whistleblowers (Keeble-Gagnere, 2015). Encryption coupled with anonymity (as witnessed in the Panama Papers) is one way to do this.

Harnessing new technologies: Blockchain is just one tool that may protect and support data and investigative work (IJNET, 2016). As Walid Al-Saqaf of Södertörn University (Sweden) (as cited in Bouchart, 2018) has explained: "Blockchain preserves data permanently and prevents any manipulation or fraud. That means that if governmental data is there it can't be removed or changed once it is published." Machine learning is another technology already being used in this space, and one which will only grow (Bradshaw, 2017).

Rebuilding public trust: With trust in journalism at near-record lows, it is incumbent on all journalists to work towards remedying this (Knight Foundation, 2018; Nicolau & Giles, 2017). For those working with data, this means being transparent about the data you are working with, providing

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

links to the original sources, and ensuring that original data files are available for download. Showing your work—what Jay Rosen (2017) calls "the new terms for trust in journalism"—allows readers to see the raw materials you worked with (and interpreted), and thereby opens a door to transparency-based trust in news.

The influence of data on your work is/will be wider than you might think. Finally, it is impossible to overlook the role that data also plays in shaping acts of journalism. We need to remember that the datafication of journalism is not just influencing data storytelling but also the wider journalistic profession (Anderson, 2018; Usher, 2016).

Analytics tools such as Google Analytics, Chartbeat and Omniture are omnipresent in newsrooms, giving journalists more information about the reading habits of their audiences than ever before. These quantitative insights, coupled with qualitative insights (see, e.g., programs like Metrics for News, developed by the American Press Institute), are informing the work of newsrooms large and small.

As highlighted in white papers published by Pars.ly[5] and in recent academic research (Belair-Gagnon & Holton, 2018; Cherubini & Nielsen, 2016; Ponsford, 2016; Radcliffe et al., 2017; Zamith, 2018), it's clear that data is playing a pivotal role both in the positioning of stories (including literally how they are placed on homepages and promoted on social media) and in the decision making around what stories get covered.

Levi Pulkkinen, a Seattle-based reporter and editor and former senior editor of the *Seattle Post-Intelligencer*, argues that much of this data suggests that newsrooms need to do some things differently. "I think there's a hesitancy in the newspaper industry among reporters to not recognize that what the metrics are telling us is that we need to change the content," Pulkkinen (as cited in Radcliffe, 2017a) says, indicating that public affairs reporting (among other beats) may be ripe for change. "They like when we can tell them a whole story, or tell them an important story . . . but they don't need us to just act as a kind of stenographer of government" (as cited in Radcliffe, 2017a).

Moving Forward: Five Ideas for Industry–Academy Collaboration

Data is shaping and informing acts of journalism across virtually all newsrooms and reporting beats. It can be a tool for telling specific stories—as exemplified among established players such as *The Guardian* and newer

5 https://www.parse.ly/resources/case-studies/

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

entities such as *FiveThirtyEight* and *Quartz* (Ellis, 2015; Seward, 2016)—as well as an important source for editorial and resource-driven decision making.

But beyond discrete stories and strategies, data portends a larger sea change in journalism. For better or worse, an embrace of quantification may well have major implications for what have been described as the Four Es of big data and journalism: Epistemology (what journalism knows), expertise (how journalism expresses that knowledge), economics (journalism's market value) and ethics (journalism's social values) (Lewis & Westlund, 2015). The data-related implications are therefore far-reaching—for how we teach, practice and research journalism.

We believe that, too often, the worlds of academia and news industry fail to recognize the generative potential that could come through greater collaboration between them (much like our point about collaborative journalism, above). As both parties grapple with the possibilities afforded by datafication, we contend that closer relationships between journalists and academics could be mutually beneficial. Below we outline five starting points to explore.

More partnerships between classrooms and newsrooms. The work undertaken by Paul Bradshaw offers a clear indication of how to do this. As part of the new MA in Data Journalism offered at Birmingham City University in the UK, Paul and his students have partnered with a number of news organizations, such as *The Daily Telegraph* (Bradshaw, 2018), the BBC, ITN, the *Manchester Evening News*, *The Guardian* and the Centre for Investigative Journalism.[6] To extend this teaching-based partnership to improve research, these news organizations could open up their data journalism processes to (participant) observation by ethnographers, with the expectation that such scholarship would lead not only to peer-reviewed academic publication but also to public-facing reports that are intended for industry—like the kind produced by the Tow Center for Digital Journalism and the Reuters Institute for the Study of Journalism.

Undertake classroom projects with potential news value. Jan Goodey, a journalism lecturer at Kingston University in west London, has also demonstrated the ability to turn class projects into tangible reporting, having identified some potential conflicts of interest in UK local government. Their research—which included submitting, tracking and analyzing 99 separate FOI requests—revealed that these bodies were investing pension funds in fracking companies, while at the same time also acting as arbiters for planning proposals submitted by this nascent industry (Goodey & Milmo,

6 http://bcu.ac.uk/media/courses/data-journalism-ma-2018-19

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

2014). In some cases, students and their professors may have a longer time horizon to explore data projects, thus allowing them to do forms of data journalism that are elusive for journalists overwhelmed by ceaseless daily deadlines.

Reverse-engineer these relationships. Given the resource challenges that most newsrooms face, journalists could more frequently approach students and academics with stories that could benefit from their help. Arizona State University's Steve Doig, who won a 1993 Pulitzer Prize for Public Service at *The Miami Herald*,[7] for a series which showed how weakened building codes and poor construction practices exacerbated damage caused by Hurricane Andrew a year before, actively consults on computer-assisted reporting problems.[8] He won the George Polk Award (2012) for Decoding Prime, an analysis of suspect hospital billing practices for the California Watch investigative organization.[9] His is an advising and consultancy model—with faculty and potential student involvement—that others could emulate.

Open the door to researchers and independent critique. Journalists are known to rely on academics as frequent sources for news stories, but they are often less comfortable opening themselves up to academic scrutiny. Compounding this problem are increasingly strident organizational directives against taking surveys or speaking to researchers without permission from upper management. But, just as journalists need good source material to do their work, for academics to do good research *about* journalism requires their having better access than they presently do. This is especially pertinent as researchers seek to understand what datafication means for journalism (Baack, 2015)—for how journalists use metrics (Belair-Gagnon & Holton, 2018; Christin, 2018; Ferrer-Conill & Tandoc, 2018), for how they tell stories in new ways (Hermida & Young, 2019; Toff, 2019; Usher, 2016) and so on. A little less defensiveness on the part of news organizations could go a long way towards developing a mutually beneficial relationship: Researchers get better access to understanding how data fits in journalism, and in turn news organizations can gain independent evaluations of their work and thus better appraise, from a critical distance, how they are doing.

Ensure your research is accessible. On the flip side, academics could do much more to ensure the openness and accessibility of their work. By now, dozens of academic studies have been produced regarding the "datafication

7 http://www.pulitzer.org/winners/miami-herald

8 https://www.flickr.com/photos/juggernautco/sets/72157607210036175/

9 https://cronkite.asu.edu/about/faculty-and-leadership/faculty/doigbio

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

of journalism," with a particular emphasis on the evolution of tools for data storytelling and its impact on journalistic ethics and approaches (for an overview, see Ausserhofer et al., 2020). These studies could have tremendous relevance for news organizations. But too often they are locked behind academic journal paywalls, obscured by the overuse of jargon and altogether situated in such a way that makes it hard for journalists to access, let alone understand, the transferable lessons in this research. Where possible, industry outreach and engagement could be an integral part of the publication process, so that the benefits of these insights resonate beyond the journals—such as through rewritten briefs or short explainers for trade-press venues, such as Nieman Journalism Lab, or websites designed to disseminate academic work to lay audiences, such as *The Conversation*.

Conclusion

Data journalism, in the words of famed data journalist Simon Rogers (2012), now data editor at Google, is "a great leveler." Because of its emergent character, virtually anyone can try it and become proficient in it. "Data journalism is the new punk," he says (Rogers, 2012). This means that "many media groups are starting with as much prior knowledge and expertise as someone hacking away from their bedroom" (Rogers, 2012).[10]

Data journalism, of course, has a long history, with antecedents in forms of science and storytelling that have been around for more than a century (Anderson, 2018).[11] But as a nascent "social world" (Lewis & Zamith, 2017) within journalism—a space for sharing tools, techniques and best practices across news organizations and around the globe—data journalism is at a particular inflection point, amid the broader datafication of society in the 21st century.

There is a corresponding opportunity, we argue, for critical self-reflection: For examining what we know about data journalism so far, for outlining what remains to be explored, and particularly for pursuing a path that recognizes the mutual dependence of journalism as practice and pedagogy, industry and academy. For journalism to make sense of a world awash in data requires better recognizing, self-critically, the limitations and generative possibilities of data-driven approaches—what they reveal, what they don't and how they can be improved.

10 See also Simon Rogers' chapter in this book.

11 See also Anderson's chapter in this book.

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

Works Cited

2016 year in review: Highlights and heartbreaks. (2016, December 29). *Los Angeles Times*. https://www.latimes.com/local/california/la-me-updates-best-year-review-2016-htmlstory.html#every-shot-kobe-bryant-ever-took-all-30-699-of-them

Anderson, C. W. (2018). *Apostles of certainty: Data journalism and the politics of doubt*. Oxford University Press.

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) 950–973. https://doi.org/10.1177/1464884917700667

Baack, S. (2015). Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data & Society*, *2*(2), 2053951715594634. https://doi.org/10.1177/2053951715594634

Barrett, D. (2009, May 17). MPs' expenses: How they milk the system. *The Telegraph*. https://www.telegraph.co.uk/news/newstopics/mps-expenses/5335101/MPs-expenses-how-they-milk-the-system.html

Beckett, C. (2010). *The value of networked journalism.* 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), 492–508. 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. In P. J. Boczkowski & Z. Papacharissi (Eds.), *Trump and the media* (pp. 177–185). MIT Press. https://doi.org/10.7551/mitpress/11464.003.0028

Bouchart, M. (2018, February 1). A data journalist's new year's resolutions. *Medium*. https://medium.com/data-journalism-awards/a-data-journalists-new-year-s-resolutions-474ef92f7e8f

Bradshaw, P. (2017, December 14). Data journalism's AI opportunity: The 3 different types of machine learning & how they have already been used. *Online Journalism Blog*. https://onlinejournalismblog.com/2017/12/14/data-journalisms-ai-opportunity-the-3-different-types-of-machine-learning-how-they-have-already-been-used/

Bradshaw, P. (2018, February 15). Wanted: MA Data Journalism applicants to partner with *The Telegraph*. Medium. https://medium.com/@paulbradshaw/wanted-ma-data-journalism-applicants-to-partner-with-the-telegraph-8abd15426of3

Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world*. MIT Press.

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

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), 673–695. 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), 642–650. 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* (SSRN Scholarly Paper ID 2739328). Social Science Research Network. 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), 1382–1415. https://doi.org/10.1086/696137

A chronology of the Senate expenses scandal. (2016, July 13). *CBC News*. https://www.cbc.ca/news/politics/senate-expense-scandal-timeline-1.3677457

Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media*, *20*(4), 336–349. 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*, 1663–1672. https://ijoc.org/index.php/ijoc/article/view/2167

The depth of the problem. (n.d.). *The Washington Post*. https://www.washingtonpost.com/apps/g/page/world/the-depth-of-the-problem/931/

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

Ellis, J. (2015, June 23). Quartz maps a future for its interactive charts with Atlas. Nieman Lab. https://www.niemanlab.org/2015/06/quartz-maps-a-future-for-its-interactive-charts-with-atlas/

Expenses scandal an embarrassing start to 2017 for Australia's embattled Prime Minister. (2017, January 13). *The Indian Express*. https://indianexpress.com/article/world/expenses-scandal-an-embarrassing-start-to-2017-for-australias-embattled-prime-minister/

Ferrer-Conill, R., & Tandoc, E. C. (2018). The audience-oriented editor. *Digital Journalism*, *6*(4), 436–453. https://doi.org/10.1080/21670811.2018.1440972

Frank, A., Yang, Y., & Radcliffe, D. (2015, December 18). The mainstreaming of data reporting and what it means for journalism schools. *Journalism.co.uk*. https://www.journalism.co.uk/news/the-mainstreaming-of-data-journalism-and-what-it-means-for-journalism-schools/s2/a593608/

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), 114–137. https://www.hbs.edu/faculty/Pages/item.aspx?num=51012

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

Goodey, J., & Milmo, C. (2014, April 27). Exclusive: Local authorities have "conflict of interest" on fracking investments. *The Independent*. http://www.independent.co.uk/news/uk/politics/exclusive-local-authorities-have-conflict-of-interest-on-fracking-investments-9294590.html

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

Hamilton, J. T. (2016). *Democracy's detectives: The economics of investigative journalism*. Harvard University Press.

Hermida, A., & Young, M. L. (2019). *Data journalism and the regeneration of news*. 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), 147–169. https://doi.org/10.1080/10919392.2015.1125187

ICIJ. (n.d.). The Panama Papers: Exposing the rogue offshore finance industry. https://www.icij.org/investigations/panama-papers/

IJNET. (2016, August 5). How blockchain technology can boost freedom of the press. International Journalists' Network. https://ijnet.org/en/story/how-blockchain-technology-can-boost-freedom-press

Johnson, S. (2013, February 8). Sorry EveryBlock, you never learned how to write a headline. *The Hudson Eclectic*. https://hudsoneclectic.com/2013/02/08/sorry-everyblock-you-never-learned-how-to-write-a-headline/

Keeble-Gagnere, G. (2015, November 5). Encryption for the working journalist: Communicating securely. *Journalism.co.uk*. https://www.journalism.co.uk/news/encryption-for-the-working-journalist/s2/a580914/

Klein, S. (2015, March 17). Antebellum data journalism: Or, how big data busted Abe Lincoln. *ProPublica*. https://www.propublica.org/nerds/antebellum-data-journalism-busted-abe-lincoln

Knight Foundation. (2018, June 26). 10 reasons why Americans don't trust the media. *Medium*. https://medium.com/trust-media-and-democracy/10-reasons-why-americans-dont-trust-the-media-d0630c125b9e

Kreiss, D. (2016). *Prototype politics: Technology-intensive campaigning and the data of democracy*. 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*, 75–99. https://doi.org/10.1016/j.ejpoleco.2016.12.006

Lewis, S. C. (2017). Digital journalism and big data. In B. Franklin & S. Eldridge (Eds.), *The Routledge companion to digital journalism studies* (pp. 126–135). 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), 44–47. https://doi.org/10.1177/1464884918808134

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

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), 383–393. 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), 543–560. 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), 447–466. https://doi.org/10.1080/21670811.2014.976418

Lewis, S. C., & Zamith, R. (2017). On the worlds of journalism. In P. J. Boczkowski & C. W. Anderson (Eds.), *Remaking the news: Essays on the future of journalism scholarship in the digital age* (pp. 111–128). MIT Press.

Lowenstein, J. K. (2015, February 13). How I used data-driven journalism to reveal racial disparities in U.S. nursing homes. *Storybench*. https://www.storybench.org/how-i-used-data-driven-journalism-to-reveal-racial-disparities-in-u-s-nursing-homes/

Mitchell, A., Gottfried, J., Barthel, M., & Shearer, E. (2016, July 7). *The modern news consumer: News attitudes and practices in the digital era*. Pew Research Center. https://www.journalism.org/2016/07/07/pathways-to-news/

Mitchell, A., Rosenstiel, T., & Christian, L. (2012). Mobile devices and news consumption: Some good signs for journalism. In Pew Research Center (Ed.), *The state of the news media: An annual report on American journalism*. https://www.pewresearch.org/wp-content/uploads/sites/8/2017/05/State-of-the-News-Media-Report-2012-FINAL.pdf

MPs' expenses: Full list of MPs investigated by *The Telegraph*. (2009, May 8). *The Telegraph*. https://www.telegraph.co.uk/politics/0/mps-expenses-full-list-mps-investigated-telegraph/

MPs' expenses: *The Guardian* launches major crowdsourcing experiment. (2009, June 23). *The Guardian*. https://www.theguardian.com/gnm-press-office/crowdsourcing-mps-expenses

News consumption on digital devices. (2017, August 21). Pew Research Center. https://www.journalism.org/2012/03/18/mobile-devices-and-news-consumption-some-good-signs-for-journalism/5-news-consumption-on-digital-devices/

Nicolau, A., & Giles, C. (2017, January 16). Public trust in media at all time low, research shows. *Financial Times*. https://www.ft.com/content/fa332f58-d9bf-11e6-944b-e7eb37a6aa8e

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Ponsford, D. (2016, July 31). "Heartbroken" reporter Gareth Davies says *Croydon Advertiser* print edition now "thrown together collection of clickbait." *Press Gazette*. https://www.pressgazette.co.uk/heartbroken-

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

reporter-gareth-davies-says-croydon-advertser-print-edition-now-thrown-together-collection-of-clickbait/

Q&A: MP expenses row explained. (2009, June 18). *BBC News*. http://news.bbc.co.uk/2/hi/uk_news/politics/7840678.stm

Radcliffe, D. (2013). Hyperlocal media and data journalism. In J. Mair & R. L. Keeble (Eds.), *Data journalism: Mapping the future* (pp. 120–132). Abramis Academic Publishing.

Radcliffe, D. (2017a). Local journalism in the Pacific Northwest: Why it matters, how it's evolving, and who pays for it (SSRN Scholarly Paper ID 3045516). Social Science Research Network. https://papers.ssrn.com/abstract=3045516

Radcliffe, D. (2017b). Data journalism in the US: Three case studies and ten general principles for journalists. In J. Mair, R. L. Keeble, & M. Lucero (Eds.), *Data journalism: Past, present and future* (pp. 197–210). Abramis Academic Publishing.

Radcliffe, D. (2017c, November 28). 10 key principles for data-driven storytelling. *Journalism.co.uk*. https://www.journalism.co.uk/news/10-key-principles-for-data-driven-storytelling/s2/a713879/

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. https://doi.org/10.7916/D8XP7BGC

Rayner, G. (2009, May 8). MPs' expenses: Ten ways MPs play the system to cash in on expenses and allowances. *The Telegraph*. 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

Reese, S. D. (1999). The progressive potential of journalism education: Recasting the academic versus professional debate. *Harvard International Journal of Press/Politics*, *4*(4), 70–94. https://doi.org/10.1177/1081180X9900400405

Rogers, S. (2009a, June 18). How to crowdsource MPs' expenses. *The Guardian*. https://www.theguardian.com/news/datablog/2009/jun/18/mps-expenses-houseofcommons

Rogers, S. (2009b, June 19). MPs' expenses: What you've told us. So far. *The Guardian*. http://www.theguardian.com/news/datablog/2009/sep/18/mps-expenses-westminster-data-house-of-commons

Rogers, S. (2012, May 24). Anyone can do it. Data journalism is the new punk. *The Guardian*. https://www.theguardian.com/news/datablog/2012/may/24/data-journalism-punk

Rosen, J. (2017, December 31). Show your work: The new terms for trust in journalism. *PressThink*. https://pressthink.org/2017/12/show-work-new-terms-trust-journalism/

Sambrook, R. (Ed.). (2018). *Global teamwork: The rise of collaboration in investigative journalism*. Reuters Institute for the Study of Journalism.

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053

Schroeder, R. (2018). *Social theory after the Internet: Media, technology, and globalization*. UCL Press.

Senate expenses scandal. (n.d.). *HuffPost Canada*. https://www.huffingtonpost.ca/news/senate-expenses-scandal/

Seward, Z. M. (2016, May 10). Atlas is now an open platform for everyone's charts and data. *Quartz*. https://qz.com/679853/atlas-is-now-an-open-platform-for-everyones-charts-and-data/

Thurman, N., Lewis, S. C., & Kunert, J. (2019). Algorithms, automation, and news. *Digital Journalism*, *7*(8), 980–992. https://doi.org/10.1080/21670811.2019.1685395

Tigas, M., Jones, R. G., Ornstein, C., & Groeger, L. (2019, October 17). Dollars for docs. *ProPublica*. https://projects.propublica.org/docdollars/

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), 873–889. https://doi.org/10.1177/1464884917731655

Usher, N. (2016). *Interactive journalism: Hackers, data, and code*. University of Illinois Press.

Van der Haak, B., Parks, M., & Castells, M. (2012). The future of journalism: Networked journalism. *International journal of communication*, *6*, 2923–2938.

Worth a thousand words. (2013, October 7). *The Economist*. https://www.economist.com/christmas-specials/2013/10/07/worth-a-thousand-words

Zamith, R. (2018). Quantified audiences in news production: A synthesis and research agenda. *Digital Journalism*, *6*(4), 418–435. https://doi.org/10.1080/21670811.2018.1444999

About the Authors

Damian Radcliffe is the Carolyn S. Chambers Professor in Journalism, a Professor of Practice, and an affiliate of the Department for Middle East and North Africa Studies (MENA), at the University of Oregon.

Seth C. Lewis (PhD, University of Texas at Austin) is the founding holder of the Shirley Papé Chair in Emerging Media in the School of Journalism and Communication at the University of Oregon and is a visiting fellow with the Reuters Institute for the Study of Journalism and with Green Templeton College at the University of Oxford.

This content downloaded from 73.67.149.124 on Tue, 17 Aug 2021 20:41:15 UTC
All use subject to https://about.jstor.org/terms
Electronic copy available at: https://ssrn.com/abstract=3907053