# 26

# DATA JOURNALISM AND THE PROMISE OF TRANSPARENCY

*Sylvain Parasie*

This chapter focuses on emerging journalistic practices that have been labeled 'data journalism'. Since the late 2000s, an increasing number of journalists have seen databases and algorithms as appropriate means for making news. Taking advantage of the growing availability of digital data, they promote a 'computational approach' to news production. This means that news production is viewed as a process based on a heterogeneous set of digital technologies and computing skills, where data are collected, analyzed, and presented to provide readers with various news products, from conventional articles to news applications or visualizations. In many countries, journalists have seen 'data journalism' or 'computational journalism' (for definitions, see Coddington 2015) as a way to address the multifaceted crisis facing their profession: the contraction of financial resources for news organizations, the decline of investigative reporting, and a relentless suspicion of partisanship. In other words, the 'data journalism' claim is that as they rely on data-processing technologies, news organizations are less dependent on their sources, demonstrate greater objectivity, and serve their customers more efficiently.

The main question of this chapter is how these emerging practices affect the ways in which media organizations produce scandals or collective indignation by revealing government dysfunction, showing injustice, or pointing out individual misconduct. Data journalism has generated many promises in this respect, from both inside and outside journalism – especially in technological worlds. Like previous technological innovations (photography, television), databases and algorithms are expected to enhance transparency in our democracies by allowing journalists to investigate deeper and in a more robust way, to be less dependent on official sources, and to reach out more easily to the average citizen. But are these promises fulfilled? What are the pitfalls and how do news organizations adopt this computational perspective on news making?

This chapter builds not only on recent scholarship on data journalism, but also on more established literatures related to journalism, and on science and technology studies. Drawing on empirical investigations mostly conducted in the United States and Europe, I identify five lines of inquiry that have been partially investigated by scholars in sociology, communication science, and science and technology studies:

(1)   What is the history of data journalism and computational perspectives on news making?
(2)   How does the use of data-processing techniques affect the way that journalists and news makers work together to produce scandals or revelations based on digital data?

(3)   How have journalistic 'epistemologies' evolved, i.e. the conditions on which the knowl-
        edge claims made by journalists appear to be valid for the journalists themselves?
(4)   What are the implications of data journalism for the journalistic profession? What does
        it concern?
(5)   How are citizens involved in the making of indignation, and to what extent do data jour-
        nalists successfully involve their readers in this respect?

Before addressing these research questions, we shall begin with the promises related to data
journalism as regards the contribution of media to scandals.

## A promise of transparency

Both inside and outside journalism, many see the adoption of data-processing technologies in
news organizations as a way to enhance the democratic contribution of journalism (Parasie and
Dagiral 2013). One general argument is that, because more and more actions – from govern-
ments, administration, firms, NGOs, or ordinary citizens – are associated with the production
of digital data, journalists have to adopt digital technologies in order to make the world more
transparent. As Tim Berners-Lee, the inventor of the Web, put it, 'now it's also going to be
about poring over data and equipping yourself with the tools to analyze it and picking out
what's interesting. And keeping it in perspective, helping people out by really seeing where it
all fits together, and what's going on in the country' (Arthur 2010). For news organizations, the
promise of data-processing techniques is that they will afford the opportunity to make govern-
ments, powerful firms, and other organizations more accountable.

Several reasons are usually put forward to support the idea that data journalism enhances the
accountability function of journalism. The first one is that in Western democracies, govern-
ments are more and more inclined, and even compelled, to release their documents in the form
of digital data. This phenomenon is explained both by longstanding dynamics within public
administration – notably the 'new public management' initiated in the 1980s – and scientific
communities, and by more recent mobilization such as the 'open data' advocacy which has
promoted the idea that public authorities should release their documents as digital data that are
usable by citizens (Yu and Robinson 2012). As a consequence, news organizations can more
easily collect data about a wider range of social activities: administrative records, government
decisions, environmental pollution, health or educational issues, and so on.

Another reason is the fact that in the last decade new techniques and infrastructures have
emerged, which make it easier for news organizations to transform a wide range of documents
(printed papers, books, pictures, etc.) into digital data. Spawned by computer engineering,
computer science, and statistics, these techniques have increased what is often called 'the data-
fication' of the world: in other words, our collective ability to turn documents into data. This
process is closely associated with the growth of digital platforms that produce huge amounts of
data about online activities. Thus, news organizations can have access to a greater number of
documents in the form of digital data, but also to data-processing techniques from the so-called
'data science' community, at the interface of computer science and statistics.

Journalists and scholars have also pointed out the financial cost of investigative reporting to
claim that data-processing technologies make it cheaper for news organizations to do investiga-
tive or quality reporting (Cohen, Hamilton and Turner 2011; Hamilton 2016). It is common
knowledge that, in most Western countries, many news organizations have radically reduced
the resources devoted to investigative reporting (Schudson 2010a; see Siles and Boczkowski
2012). Data journalism and computational journalism are consequently viewed as promising

ways of reducing the costs of investigative reporting (Hamilton 2016). One reason is that these technologies are said to serve investigation by allowing for the analysis of extensive sets of digital data, which is supposed to cost less than traditional human reporting. Another reason is that by automating the production of standardized articles based on digital data, news organizations are believed to be able to invest more resources in investigation (see Young and Hermida 2015).

Yet these promises are not to be taken for granted. The adoption of data-processing techniques by news organizations has unquestionably led to the revelation of significant scandals or affairs. WikiLeaks's revelations in 2010 of the abuses committed by the US army against civilians in Iraq, or more recent investigations on tax evasion carried out by the International Consortium of Investigative Journalists, are good examples of greater transparency regarding governments and private actors. However, these techniques cannot be assumed to increase automatically the general level of transparency in our societies. 'Technological determinism', that is, the tendency to suppose that technology has an effect by itself, has to be challenged by scholars if they want truly to understand how these emerging practices affect the production of scandals or affairs by news organizations (see Anderson 2013).

## Historical background of data journalism

One way of adopting a critical approach to data journalism is to study the historical background of the use of data-processing techniques to make government more accountable. Several scholars in sociology and communication have shown that these journalistic practices are neither entirely driven by technology, nor imposed from outside of journalism (Schudson 2010b; Parasie and Dagiral 2013; Anderson 2016). Instead, they build on previous democratic norms and well-established journalistic practices. This is the point made by Michael Schudson, who shows that journalistic interest in databases dates back to the early twentieth century (Schudson 2010b). As he notes, in the 1920s philosopher and writer Walter Lippmann wrote that the world had already become too complex to be adequately grasped by the conventional tools of journalism. Lippmann called for the creation of political observatories and even what he called a 'machinery of record' that would provide newspapers with objective, exhaustive, and trusted information. In an era when computers did not exist, he imagined that journalists might be able to grasp every social, economic, or political issue through quantified information, since that was already possible at the time in sport and finance. Schudson concludes that the contemporary impetus for data or computational tools has a long history in US journalism. He also suggests that data journalism in the United States makes sense only if one considers a set of political changes that took place in recent US history. For instance, many NGOs such as political observatories that emerged in the 1970s have increased government transparency by producing huge amounts of quantified information about governments.

Other studies also suggest that data journalism builds on previous historical developments. Chris Anderson, for instance, analyzes how the 'social survey movement' in the 1910s generated new journalistic practices combining social science and investigative journalism (Anderson 2016; Anderson 2018). In the United States, some movements – such as 'Men and Religion Forward' – promoted quantitative approaches and social science in order to reform society, with some interest from newspapers which published charts or diagrams. But at the time, journalists did not consider that their professional role was to produce interpretations based on techniques which were moreover very difficult to integrate into the news making process. The situation changed in the late 1960s, as US investigative reporters began to use computer and quantitative social science in order to reveal government wrongdoings or inequalities within US society. Philip Meyer and David Burnham were among the first journalists to make major revelations

based on computer processing of public data, showing for instance that black people are the main victims of urban violence in New York. As Sylvain Parasie and Éric Dagiral (2013) have shown, these investigative reporters identified with the term 'computer-assisted reporting' that was gaining currency in US journalism in the late 1980s. They believed that quantitative social science and computers provided a more efficient way of revealing wrongdoings or public issues through public documents: courts records, public decisions, and so on. In their view, adopting a computational approach to investigation would allow journalism to serve its accountability function better (Meyer 1973). Rather than singling out individual wrongdoings, this approach would shed light on more systemic explanations of government dysfunction.

So far, few scholars have studied the historical background of data journalism, particularly outside of the USA. As these practices develop in European countries, for instance, it remains unclear how journalists build on shared professional norms when they develop an interest in quantification and data processing as a way to increase transparency and accountability.

## A new division of labor

In basing their work on digital data and algorithms, journalists collaborate with professionals and organizations that they were previously not used to working with (Gillespie, Boczkowski and Foot 2014). The revelation of scandals or public issues based on digital data and algorithms thus entails the contribution of a greater number of people and organizations. This has been noted by the researchers who have studied corpora of data-journalism projects praised by the profession (Loosen, Reimer and de Silva-Schmidt 2017). Whereas traditional investigative reporters usually work in small teams, these investigations require journalists to work with new actors – programmers, computer scientists, NGOs, etc. – and infrastructures such as databases, software, and digital platforms. As we shall see, these new collaborations raise new questions that scholars have only partially investigated.

The most significant aspect of data journalism is the renewed relationships of news organizations with people with a background in computer science, statistics, or Web development. Some of them have integrated newsrooms and have been recognized as being part of the journalistic profession, as 'data journalists', 'programmer-journalists', or 'news applications developers' (Royal 2010; Parasie and Dagiral 2013). In short, the division between journalists in charge of producing editorial content, on the one hand, and programmers in charge of publishing or displaying that content on the other, is being challenged, as it is more widely accepted that computational skills may contribute to the production of the editorial content itself. Not only have data journalists called for an end to this division, but the institutions of journalism themselves have acknowledged the journalistic contribution of programmers and developers. This is evidenced in the emergence of the international 'hacks/hackers' network, where coders and journalists share their experience (Lewis and Usher 2014), or in the fact that major US journalism foundations increasingly value news projects initiated by computer scientists or propose a data-oriented solution, to the detriment of more traditional news projects (Lewis 2011).

These new forms of collaboration between journalists and programmers raise several questions for research. One major question is: to what extent do the people with a background in computer science share some of the core journalistic values about what constitutes a scandal? Some authors have noted that programmers, developers, or designers often share a concern for the notion of free flow of information in society, which may differ significantly from the journalistic commitment to freedom of information (Annany and Crawford 2015). Does this entail a departure from what journalists have historically understood to constitute a scandal? Another

question is how journalists can develop data-processing skills, at least to interact with program-mers or designers. Recent research suggests difficulties in many newsrooms in the USA and Europe (Gynnild 2014; Karlsen and Stavelin 2014; Hannaford 2015). Another critical question is the extent to which the relationship between journalists and programmers can be described as balanced or imbalanced. As programmers design algorithms in order to collect, process, or visu-alize information, do they increase their control over the news-making process at the expense of traditional journalists? These are critical questions for future research.

A second aspect of this new division of labor is the role played by 'data providers'. In most cases, they are governments or public institutions more generally, which provide news organi-zations with digital data and statistics. Official agencies in charge of producing statistical infor-mation (US Census; Eurostat) remain the most common data sources for news organizations (Parasie and Dagiral 2013; Loosen, Reimer and de Silva-Schmidt 2017), but other data provid-ers have emerged. They may be individuals or small organizations leaking data from a private firm or government, for civic or commercial reasons. Such data leaks have aroused keen interest among journalists, especially around WikiLeaks and the emerging networks of investigative journalists. They may also be private companies selling digital data they collect from online platforms or public agencies ('data brokers'), or NGOs or political observatories releasing data for civic-oriented reasons.

These data providers raise important questions for research. Since the grounds on which they base their release of data are civic or commerce-oriented, on what conditions can news organi-zations trust them and reveal scandals based on their data? Decades ago, sociologists Harvey Molotch and Marilyn Lester (1974) analyzed the role of information promoters in the journal-istic making of public events. The same line of questions emerges here, as journalists have to work with data providers that have their own agenda which potentially conflicts with that of the news organization. Yet, it would be wrong to assume that news organizations cannot use data from unofficial sources. The main question is how journalists develop their own standard as they assess the quality of the data (Parasie 2015).

Even with 'open government data', journalists can find it very difficult to reveal scandals or affairs. As scholars have shown, the process of openness draws on a series of transformations that require significant work from administrations in order to ensure that the data remain intelligible outside of their organizations (Denis and Goëta 2017). As a consequence, journalists often face major challenges with open government data, as they struggle to articulate them to reveal scan-dals involving governments or private interests.

There is a need for scholars to analyze the effects of these new collaborations on the mak-ing of scandals. How does the growing availability of digital data affect the collective produc-tion of scandals? To what extent can media organizations rely on digital data coming from new sources without being manipulated by political or commercial interests? And does the growing involvement of people with a background in computer science imply new ways of considering scandals? These are major questions for research.

## A shift of epistemologies?

Another line of research focuses on the epistemological foundations informing data journalists' work as they reveal scandals or affairs based on data-processing techniques. Anchored in the sociology of knowledge, the study of journalistic 'epistemologies' implies the analysis of how journalists make knowledge claims that they collectively find acceptable – and not an evaluation of whether those claims are valid or not (Ettema and Glasser 1998). The question here is how the collection and processing of digital data affect the epistemologies on which journalists rely to

produce scandals. In the last decades, many activists have considered that the internet offered an opportunity to limit the media control over the production and broadcasting of scandals. Since the late 2000s, WikiLeaks has raised significant concerns by releasing a considerable volume of data about many governments' actions (Benkler 2011). More recently, computer scientists have argued that so-called 'big data' techniques are prompting major change in news organizations' production of scandals (Lewis 2015). They claim that the statistical processing of huge datasets by news organizations facilitates revelations about the nature of society without relying on any theoretical or normative assumptions. This, they argue, means a change towards greater objectivity, since most journalists usually make assumptions before investigating a topic (Ettema and Glasser 1998).

Scholars have analyzed the epistemological claims made by data journalists and people with a background in computer science. Studying data journalism in Chicago in 2010, Parasie and Dagiral pointed out that the rise of 'programmer-journalists' within news organizations challenged the epistemologies conveyed by established journalists, and especially computer-assisted reporters (Parasie and Dagiral 2013). Drawing on the study of news projects and interviews with data journalists, these authors highlight three propositions that partially challenge long-standing journalistic epistemological principles: (1) news in itself should be viewed as computer-processable data, and not only as a story hidden in the data; (2) readers should be given the opportunity not only to check the data but also to combine them and to use them for other goals; (3) as databases are regularly and automatically updated, online applications allow journalists and ordinary citizens to check decisions made by city officials and to monitor possible wrongdoings.

Yet most researchers stress that existing epistemologies tend to prevail in established news organizations. In the USA and in Europe, reporters working on computational news projects usually claim that they put these techniques at the service of their own journalistic standards. As Bucher (2017) found in several Danish, Norwegian, and Swedish news organizations, journalists emphasize the fact that they established their own priorities, without letting technology take the lead in any way. Interviewing reporters and data journalists from the *Los Angeles Times*, Young and Hermida (2015) analyzed how they collected homicide data from official sources and designed algorithms in order to generate automatically articles about every homicide, along with an interactive map. The authors' claim is that data-processing techniques are tailored to fit established journalistic norms: each homicide is viewed through various criteria that are relevant for journalists; various sources of data are collected and combined in order to identify inconsistencies; and so on. Other researchers confirm this tendency among journalists to stick to their established epistemological perspective on news making.

Although scholars have produced many ethnographic accounts of online news work in the last decade (Paterson and Domingo 2008; Boczkowski 2010), few have undertaken ethnographic research on how news organizations conduct particular investigations based on data and algorithms in order to examine 'journalistic epistemologies in action'. One researcher has focused on a particular news organization, analyzing how it revealed that the State of California had failed to enforce earthquake safety standards in public schools (Parasie 2015). On the basis of in-depth interviews with the team of reporters and data journalists in charge of this investigation, the author shows that the organization devoted huge resources so that the database and the algorithms match with the established epistemology of journalistic investigation. The study shows that if the investigative team succeeded in producing a revelation despite the poor quality of the data, it is mostly because the reporters remained skeptical about the objectivity of the data. Thanks to a great deal of organizational work, they managed to rely on the data to produce a revelation that had a political impact.

## Journalism as a profession

Data journalism can be viewed as an emerging 'professional segment' within journalism (Bucher and Strauss 1961). Data journalists are a coalition of people for whom some uses of technology provide the opportunity to accomplish better the traditional mission of journalism: mainly, informing citizens and holding governments accountable. Like any segments, they often challenge the norms of the profession as a whole. Several researchers have significantly contributed to the analysis of this process, but quite a few questions remain unanswered as regards media scandals.

Some scholars have started to map this emerging segment in the United States and Europe, describing the educational backgrounds, skills, tools, and values of the individuals who call themselves data journalists (Fink and Anderson 2015; De Maeyer et al. 2015). These researchers show the strong diversity of trajectories. Some of them had professional experience in computer science before joining news organizations; others had been first trained as journalists, and chose to develop their skills in web development, programming, or statistics. More research is needed here to grasp fully the diversity of this emerging professional segment. Specifically, it is important to study their connection to investigative journalism, and more generally the production of scandals.

Some scholars have studied the 'boundary work' done by journalists as they refused to acknowledge as 'journalistic' certain ways of producing scandals based on data processing. Coddington (2014) analyzed this collective work done by British and American journalists in the case of WikiLeaks. This organization has triggered critical reactions by most journalists, who refuse to see this as journalism. In order to dissociate themselves from WikiLeaks, these journalists claim that journalism involves many tasks such as selecting, editing, and interpreting documents, which WikiLeaks refuses to take on. Parasie and Dagiral (2013) have also analyzed criticism against data-oriented news projects, voiced by journalists who stress that the data is often not verified, whereas traditional journalists must verify every document.

Data journalists usually claim that they only use data-processing techniques to perform what traditional journalists do: collect information, analyze it, and display it to the audience. Several researchers have pointed out that journalists from established news organizations often consider data journalism as 'journalism as usual', underlining the necessity for any journalist to share common professional standards (Appelgren and Nygren 2014; Bucher 2017). In a provoking essay, Daniel Kreiss criticizes the definition of journalism that is often assumed by data journalists (Kreiss 2016). He argues that journalism cannot be defined as information production, for two reasons: first, because the good health of a democracy is not proportional to the volume of information that is provided to the citizens; and second, because this definition jeopardizes journalism as a profession, since computer scientists are better equipped to produce and process information on an industrial scale. Kreiss calls for another definition of journalism that emphasizes the production of a 'civic skepticism' in society, notably against public or private authorities. As we can see, the rise of data journalism gives professionals and scholars the opportunity to reconsider what the core missions of journalism are. But more research is needed on that topic.

## Involving citizens in the making of scandals

A last area of research questions focuses on the involvement of citizens in the making of scandals or affairs. Promoters of data journalism often claim that their techniques help citizens play a greater role in the process of making wrongdoings or scandals visible. As they can check the data themselves – in raw formats, or via online maps and web applications – citizens are said to be able to identify wrongdoings, inequalities, or public issues. Consequently, ordinary citizens are afforded the opportunity to contribute more substantially to the making of scandals

(Parasie 2013). This idea has been supported by many 'open data' advocates, often from the computer world, and implemented by many news organizations.

News organizations have reached out to their readers through web applications displaying data about a wide range of political, social, economic, or environmental issues. Users can check measures of air quality at street level, parliamentarians' expenses accounts, crime rates in cities or neighborhoods, and so on. In some cases, newspapers have crowd-sourced some tasks that needed too many resources inside the newsroom, such as data collection, cleaning, verifying, or analysis (Diakopoulos 2015). For instance, *The Guardian* obtained high levels of participation from its readers when it asked them to identify which parliamentarians' expenses could be worth investigating (Daniel and Flew 2010). This led to a major scandal in the UK. But most established organizations have been reluctant to get their readers involved in the making of scandals, both because it challenged their professional role as journalists, and because it proved to be very difficult to get ordinary citizens involved in such time-consuming tasks.

As a few scholars have pointed out, most organizations have used data-processing techniques to provide their readers with more personalized information (Parasie and Dagiral 2013), which has important implications for media scandals. Whereas traditional scandals usually focus on problems that are removed from readers, these techniques allow readers to situate themselves in relation to a particular problem. For instance, an investigation on air pollution would identify government responsibilities at national level, offering readers the possibility to check measures of air quality in their neighborhood. This phenomenon is in line with previous research on digital journalism (Boczkowski 2004), which shows the reluctance of established news organizations to use fully the interactive potentials of digital technologies.

And yet we know very little about the people who actually check the data provided by news organizations. What are their backgrounds and motives? Are they only interested in making decisions as consumers, or do they seek to satisfy more civic-oriented interests? There is a major issue here, as many critics and scholars express a strong concern that citizens may be provided with information that revolves only around themselves as individuals (Sunstein 2018).

## Conclusion

As news organizations access more information in the form of digital data, the way they can contribute to the making of scandals appears to be significantly affected. Therefore, it becomes urgent to analyze how these emerging media practices may increase government accountability, the transparency of powerful interests, and more generally our collective ability to make public issues visible. Social science has an important part to play here, not only by offering tools to describe and explain social phenomena emerging from the data, but also by providing an alternative to technological determinism – that is the tendency to assume that data-processing techniques automatically increase the general level of transparency in our societies. At the intersection of sociology, communication science, and science and technology studies, several scholars have made significant contributions to account for these emerging media practices. They have identified several research questions that are worth exploring further if we wish to understand fully the evolution of our democracies.

## References

Anderson, C. W. (2013) 'Towards a sociology of computational and algorithmic journalism', *New media & society*, 15, 7, pp. 1005–1021.
Anderson, C. W. (2016) 'Social survey reportage: Context, narrative, and information visualization in early 20th century American journalism', *Journalism*, 18, 1, pp. 81–100.

Anderson, C. W. (2018) *Apostles of Certainty: Data Journalism and the Politics of Doubt*. New York: Oxford University Press.

Annany, M. and K. Crawford (2015) 'A liminal press: Situating news app designers within a field of networked news production', *Digital journalism*, 3, 2, pp. 192–208.

Appelgren, E. and G. Nygren (2014) 'Data Journalism in Sweden: Introducing new methods and genres of journalism into "old" organizations', *Digital Journalism*, 2, 3, pp. 394–405.

Arthur, C. (2010) 'Analysing data is the future for journalists, says Tim Berners-Lee', *The Guardian*, 22 November. Available online at: http://www.guardian.co.uk/media/2010/nov/22/dataanalysis-tim-berners-lee, accessed on 10 April 2018.

Benkler, Y. (2011) 'A free irresponsible press: Wikileaks and the battle over the soul of the networked fourth estate', *Harvard Civil Rights-Civil Liberties Law Review*, 46, pp. 311–397.

Boczkowski, P. J. (2004) *Digitizing the News: Innovation in Online Newspapers*. Cambridge and London: MIT Press.

Boczkowski, P. J. (2010) *News at Work: Imitation in an Age of Information Abundance*. Chicago: University of Chicago Press.

Bucher, R. and A. Strauss (1961) 'Professions in process', *American Journal of Sociology*, 66, 4, pp. 325–334.

Bucher, T. (2017) 'Machines don't have instincts': Articulating the computational in journalism', *New media & society*, 19, 6, pp. 918–933.

Coddington, M. (2014) 'Defending judgment and context in "original reporting": Journalists' construction of newswork in a networked age', *Journalism*, 15, 6, pp. 678–695.

Coddington, M. (2015) 'Clarifying journalism's quantitative turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting', *Digital Journalism*, 3, 3, pp. 331–348.

Cohen, S., J. T. Hamilton and F. Turner (2011) 'Computational journalism', *Communications of the ACM*, 54, 10, pp. 66–71.

Daniel, A. and T. Flew (2010) 'The Guardian reportage of the UK MP expenses scandal: A case study of computational journalism', in *Record of the Communications Policy and Research Forum 2010*, Network Insight Pty. Ltd., Sydney, pp. 186–194.

De Maeyer, J., M. Libert, D. Domingo, F. Heinderyckx and F. Le Cam (2015) 'Waiting for data journalism: A qualitative assessment of the anecdotal take-up of data journalism in French-speaking Belgium', *Digital Journalism*, 3, 3, pp. 432–446.

Denis, J. and S. Goëta (2017) 'Rawification and the careful generation of open government data', *Social Studies of Science*, 47, 5, pp. 604–629.

Diakopoulos, N. (2015) 'Algorithmic accountability: Journalistic investigation of computational power structures', *Digital Journalism*, 3, 3, pp. 398–415.

Ettema, J. S. and T. L. Glasser (1998) *Custodians of Conscience: Investigative Journalism and Public Virtue*. New York: Columbia University Press.

Fink, K. and C. W. Anderson (2015) 'Data Journalism in the United States: Beyond the "usual suspects"', *Journalism Studies*, 16, 4, pp. 467–481.

Gillespie, T., P. J. Boczkowski and K. A. Foot (eds.) (2014) *Media Technologies: Essays on Communication, Materiality and Society*. Cambridge and London: MIT Press.

Gynnild, A. (2014) 'Journalism innovation leads to innovation journalism: The impact of computational exploration on changing mindsets', *Journalism*, 15, 6, pp. 713–730.

Hamilton, J. (2016) *Democracy's Detectives: The Economics in Investigative Journalism*. Cambridge and London: Harvard University Press.

Hannaford, L. (2015) 'Computational journalism in the UK newsroom', *Journalism education*, 4, 1, pp. 6–21.

Karlsen, J. and E. Stavelin (2014) 'Computational journalism in Norwegian newsrooms', *Journalism practice*, 8, 1, pp. 34–48.

Kreiss, D. (2016) 'Beyond administrative journalism: Civic skepticism and the crisis in journalism', in Jeffrey C. Alexander, Elizabeth Butler Breese and Maria Luengo (eds.), *The Crisis of Journalism Reconsidered: Democratic Culture, Professional Codes, Digital Future*. Cambridge: Cambridge University Press, pp. 59–76.

Lewis, S. (2011) 'Journalism innovation and participation: An analysis of the Knight News Challenge', *International Journal of Communication*, 5, pp. 1623–1648.

Lewis, S. C. (2015) 'Journalism in an era of big data: Cases, concepts, and critiques', *Digital Journalism*, 3, 3.

Lewis, S. C. and N. Usher (2014) 'Code, collaboration, and the future of journalism: A case study of the Hacks/Hackers global network,' *Digital Journalism*, 2, 3, pp. 383–393.

Loosen, W., J. Reimer and F. De Silva-Schmidt (2017) 'Data-driven reporting: An on-going (r) evolution? An analysis of projects nominated for the Data Journalism Awards 2013–2016', *Journalism*, first published online, DOI: 10.1177/1464884917735691

Meyer, P. (1973) *Precision Journalism: A Reporter's Introduction to Social Science Methods*. Bloomington: Indiana University Press.

Molotch, H. and M. Lester (1974) 'On the strategic use of routine events, accidents and scandals', *American Sociological Review*, 39, 1, pp. 101–112.

Parasie, S. (2013) 'Scandal machines: Towards a moral sociology of databases', *Réseaux*, 178–179, pp. 127–161.

Parasie, S. (2015) 'Data-driven revelation? Epistemological tensions in investigative journalism in the age of "big data"', *Digital Journalism*, 3, 3, pp. 364–380.

Parasie, S. and E. Dagiral (2013) 'Data-driven journalism and the public good: "Computer-assisted-reporters" and "programmer-journalists" in Chicago', *New media & society*, 15, 6, pp. 853–871.

Paterson, C. and D. Domingo (eds.) (2008) *Making Online News: The Ethnography of New Media Production*. London: Peter Lang.

Royal, C. (2010) 'The journalist as programmer: A case study of the *New York Times* interactive news technology department', International Symposium on Online Journalism, The University of Texas at Austin.

Schudson, M. (2010a) 'News in crisis in the United States: Panic – and beyond', in D. A. L. Levy and R. K. Nielsen (eds.) *The Changing Business of Journalism and its Implications for Democracy*. Oxford: Reuters Institute for the Study of Journalism, pp. 95–106.

Schudson, M. (2010b) 'Political observatories, databases & news in the emerging ecology of public information,' *Daedalus*, 139, 2, pp. 100–109.

Siles, I. and P. J. Boczkowski (2012) 'Making sense of the newspaper crisis: A critical assessment of existing research and an agenda for future work', *New Media & Society*, 14, 8, pp. 1375–1394.

Sunstein, C. R. (2018) *# Republic: Divided Democracy in the Age of Social Media*. Princeton: Princeton University Press.

Young, M. L. and A. Hermida (2015) 'From Mr. and Mrs. outlier to central tendencies: Computational journalism and crime reporting at the Los Angeles Times', *Digital Journalism*, 3, 3, pp. 381–397.

Yu, H. and D. G. Robinson (2012) 'The new ambiguity of "open government"', *UCLA Law Review Discourse*, 59, pp. 178–208.