Digital Journalism

Journalism In An Era Of Big Data

Seth C. Lewis[a]

[a] School of Journalism and Mass Communication, University of Minnesota–Twin Cities, USA
Published online: 27 Nov 2014.

CrossMark

Click for updates

PLEASE SCROLL DOWN FOR ARTICLE

JOURNALISM IN AN ERA OF BIG DATA
Cases, concepts, and critiques

Seth C. Lewis

This special issue examines the changing nature of journalism amid data abundance, computational exploration, and algorithmic emphasis—developments with wide meaning in technology and society at large, and with growing significance for the media industry and for journalism as practice and profession. These data-centric phenomena, by some accounts, are poised to greatly influence, if not transform over time, some of the most fundamental aspects of news and its production and distribution by humans and machines. While such expectations may be overblown, the trend lines are nevertheless clear: large-scale datasets and their collection, analysis, and interpretation are becoming increasingly salient for making sense of and deriving value from digital information, writ large. What such changes actually mean for news, democracy, and public life, however, is far from certain. As such, this calls for scholarly scrutiny, as well as a dose of critique to temper much celebration about the promise of reinventing news through the potential of "big data." This special issue thus explores a range of phenomena at the junction between journalism and the social, computer, and information sciences. These phenomena are organized around the contexts of digital information technologies being used in contemporary newswork—such as algorithms and analytics, applications and automation—that rely on harnessing data and managing it effectively. What are the implications of such developments for journalism's professional norms, routines, and ethics? For its organizations, institutions, and economics? For its authority and expertise? And for the epistemology that undergirds journalism's role as knowledge-producer and sense-maker in society?

Before getting to those questions, however, let us begin more prosaically: *What is the big deal about big data?* That may be a curious way to open a special issue on the subject, but the question is an important starting point for at least three reasons. First, it is *the* question being asked, whether directly or indirectly, in many policy, scholarly, and professional circles, on many a panel at academic and trade conferences, and across the pages of journals and forums in seemingly every discipline. This is especially true in the social sciences and humanities generally and in communication, media, and journalism specifically. While exploring the methods of computational social science (Lazer et al. 2009; see also Busch 2014; Lewis, Zamith, and Hermida 2013; Mahrt and Scharkow 2013; Oboler, Welsh, and Cruz 2012), scholars are also wrestling with the conceptual implications of digital datasets and dynamics that, in sheer size and scope, may challenge how we think about the nature of mediated communication (Boellstorff 2013; Bruns 2013; Couldry and Turow 2014; Driscoll and Walker 2014; Karpf 2012). Second, this opening query calls up the skepticism that is quite needed, for there is good reason to question not only whether big data is a "thing," but also in whose interests, toward what purposes, and with what consequences the very term is being

promulgated as a "solution" to unlocking various social problems (Crawford, Miltner, and Gray 2014; cf. Morozov 2013). Finally, to open with such an audacious question is to acknowledge at the outset that the processes and philosophies associated with big data, in the broadest sense, are very much in flux: an indeterminate set of leading-edge activities and approaches that may prove to be innovative, inconsequential, or something else entirely. What, then, *is* the big deal?

It is for this reason that I emphasize the deliberate naming of this special issue: "Journalism *in an era of* big data." While historical hindsight can make any naming of an "era" a fool's game, there also seems to be broad agreement that, in the developed world of digital information technologies, we are situated in a moment of data deluge. This moment, however loosely bounded, is noted for at least two major developments that have accelerated in recent years. The first is the overwhelming volume and variety of digital information produced by and about human (and natural) activity, made possible by the growing ubiquity of mobile devices, tracking tools, always-on sensors, and cheap computing storage, among other things. As one report described it: "In a digitized world, consumers going about their day—communicating, browsing, buying, sharing, searching—create their own enormous trails of data" (Manyika et al. 2011, 1). "This data layer," noted another observer, "is a shadow. It's part of how we live. It is always there but seldom observed" (quoted in Bell 2012, 48). The second major development involves rapid advances in and diffusion of computing processing, machine learning, algorithms, and data science (Manovich 2012; Mayer-Schönberger and Cukier 2013; O'Neil and Schutt 2013; Provost and Fawcett 2013). Put together, these developments have enabled corporations, governments, and researchers to more readily navigate and analyze this shadow layer of public life, for better or worse, and much to the chagrin of critics concerned about consumer privacy and data ethics (boyd and Crawford 2012; Oboler, Welsh, and Cruz 2012). Thus, whether dubbed "big" or otherwise, this moment is one in which *data*—its collection, analysis, and representation, as well as associated data-driven techniques of computation and quantification—bears particular resonance for understanding the intersection of media, technology, and society (González-Bailón 2013).

Computation and Quantification in Journalism

What is the big deal, then, for journalism? By now, there is no shortage of accounts about the implications of technology change for the most fundamental aspects of gathering, filtering, and disseminating news; similarly, much has been written about such changes and their implications for journalistic institutions, business models, distribution channels, and audiences (for an overview of recent scholarly work in this broad terrain, see Franklin 2014; see also Anderson, Bell, and Shirky 2012; Lewis 2012; Ryfe 2012; Usher 2014). Yet, in comparison to the large body of literature, for instance, on the role of Twitter in journalism (Hermida 2013), the particular role of *data* in journalism—as well as interrelated notions of algorithms, computer code, and programming in the context of news—is only beginning to receive major attention in the scholarly and professional discourse. Among scholars, there is a rapidly growing body of work focused on unpacking the nature of computation and quantification in news. The scholarly approaches include case studies of journalists within and across news

organizations (e.g., Appelgren and Nygren 2014; Fink and Anderson 2014; Karlsen and Stavelin 2014; Parasie and Dagiral 2013), theoretical undertakings that often articulate concepts of computer science and programming in the framework of journalism (e.g., Anderson 2013; Hamilton and Turner 2009; Flew et al. 2012; Gynnild 2014; Lewis and Usher 2013), and analyses that take a more historical perspective in comparing present developments with computer-assisted reporting (e.g., Parasie and Dagiral 2013; Powers 2012). More oriented to journalism professionals, there are a growing number of handbooks on data journalism (Gray, Bounegru, and Chambers 2012), industry-facing reports on the likes of data (Howard 2014), algorithms (Diakopoulos 2014), and sensors (Pitt 2014), and conferences on "quantifying journalism" via data, metrics, and computation.

Data journalism, as Fink and Anderson (2014, 1) note bluntly, is seemingly "everywhere," based on the industry buzz and accelerating scholarly interest. "[W]hether and how data journalism actually exists as a thing in the world, on the other hand, is a different and less understood question." This special issue is a systematic effort to address that issue. It aims to outline the state of research in this emerging domain, bringing together some of the most current and critical scholarship on what is becoming of journalism—from its reporting practices to its organizational arrangements to its discursive interpretation as a professional community—in a moment of experimentation with digital data, computational techniques, and algorithmic forms of representing and interpreting the world.

"Journalism in an era of big data" is thus a way of seeing journalism as interpolated through the conceptual and methodological approaches of *computation* and *quantification*. It is about both the ideation and implementation of computational and mathematical mindsets and skill sets in newswork—as well as the necessary deconstruction and critique of such approaches. Taking such a wide-angle view of this phenomenon, including both practice and philosophy within this conversation, means attending to the social/cultural dynamics of computation and quantification—such as the grassroots groups that are seeking to bring pro-social "hacking" into journalism (Lewis and Usher 2013, 2014)—as well as the material/technological characteristics of these developments. It means recognizing that algorithms and related computational tools and techniques "are neither entirely material, nor are they entirely human—they are hybrid, composed of both human intentionality and material obduracy" (Anderson 2013, 1016). As such, we need a set of perspectives that highlight the distinct and interrelated roles of social actors and technological actants at this emerging intersection of journalism (Lewis and Westlund 2014a).

To trace the broad outline of journalism in an era of big data, we need (1) empirical cases that describe and explain such developments, whether at the micro (local) or macro (institutional) levels of analysis; (2) conceptual frameworks for organizing, interpreting, and ultimately theorizing about such developments; and (3) critical perspectives that call into question taken-for-granted norms and assumptions. This special issue takes up this three-part emphasis on *cases*, *concepts*, and *critiques*. Such categories are not mutually exclusive nor exhaustively reflective of what is covered in this issue; indeed, various elements of case study, conceptual development, and critical inquiry are evident in all of the articles here. In that way, these studies provide a blended set of theory, practice, and criticism upon which scholars may develop future research in this important and growing area of journalism, media, and communication.

Cases, Concepts, and Critiques

For a set of phenomena as uncertain as journalism in an era of big data, conceptual clarity is the first order of business. What used to be a coherent notion of computer-assisted reporting (CAR) in the 1990s "has splintered into a set of ambiguously related practices" that are variously described in terms such as computational journalism, data journalism, programmer-journalism, and so on (Coddington 2014). Reviewing the state of the field thus far, Mark Coddington finds "a cacophony of overlapping and indistinct definitions that forms a shaky foundation for deeper research into these practices." As data-driven forms of journalism become more central to the profession, "it is imperative that scholars do not treat them as simple synonyms but think carefully about the significant differences between the forms they take and their implications for changing journalistic practice as a whole." Against that backdrop, Coddington opens this special issue by clarifying this "quantitative turn" in journalism, offering a typology of three dominant approaches: computer-assisted reporting, data journalism, and computational journalism. While there are overlaps in practice among these forms of quantitative journalism, there are also key distinctions:

> CAR is rooted in social science methods and the deliberate style and public-affairs orientation of investigative journalism, data journalism is characterized by its participatory openness and cross-field hybridity, and computational journalism is focused on the application of the processes of abstraction and automation to information.

Having classified them as such, Coddington differentiates them further according to their orientation on four dimensions: (1) professional expertise or networked participation, (2) transparency or opacity, (3) big data or targeted sampling, and (4) a vision of an active or passive public. His typology points to "a significant gap between the professional and epistemological orientations of CAR, on the one hand, and both data journalism and computational journalism, on the other." Open-source culture, he suggests, is a continuum through which to see distinctions among these forms: CAR reflecting a professional, less "open" approach to journalism, on one end, with data journalism being situated as a professional–open hybrid in the middle, and computational journalism hewing most closely to the networked, participatory values of open source (cf. Lewis and Usher 2013).

Building on Coddington's conceptualization of quantitative journalism, C. W. Anderson (2014) offers a historically based critique that reveals, at least in the US context, how "the underlying ideas of data journalism are not new, but rather can be traced back in history and align with larger questions about the role of quantification in journalistic practice." He takes what he calls an "objects of journalism-oriented" approach to studying data and news, one that pays attention (in this case historically) to how data is embodied in material "objects" such as databases, survey reports, and paper documents as well as how journalists situate their fact-building enterprise in relation to those objects of evidence. This object orientation is connected with actor-network theory (ANT) and its way of seeing news and knowledge work as an "assemblage" of material, cultural, and practice-based elements. It allows Anderson to take "a longer historical trajectory that grapples with the very meaning of 'the quantitative' for the production of knowledge," with a particular emphasis on "the epistemological dimensions of these quantitative practices" (emphasis original). By examining several historical

tensions underlying journalists' use of data—such as the document-oriented shift from thinking about news products as "records" to thinking about them as "reports" that occurred in the early nineteenth century—Anderson offers an important critique. He challenges prevailing wisdom about the orderly progression of data and visualization, showing instead that "the story of quantitative journalism in the United States is less one of sanguine continuity than it is one of rupture, a tale of transformed techniques and objects of evidence existing under old familiar names." The ultimate payoff in this approach, he argues, is both a backward-looking reappraisal of history and a forward-looking lens for examining the quantitative journalism of the future: not merely in how it embraces big data, but "rather the ways in which it reminds us of *other* forms of information that are not data, *other* types of evidence that are not quantitative, and *other* conceptions of what counts as legitimate public knowledge" (original emphasis).

With its emphasis on epistemology and materiality, Anderson's historical account sets up the contemporary case study by Sylvain Parasie (2014). He examines the San Francisco-based Center for Investigative Reporting (CIR) to explore the question: To what extent does big-data processing influence how journalists produce knowledge in investigative reporting? Parasie extends (and critiques) previous research on journalistic epistemologies in two ways, firstly by more fully taking into account "how journalists rely on the material environment of their organization to decide whether their knowledge claims are justified or not." These material factors include databases and algorithms, which "are not black boxes providing unquestionable results, and [thus] we need to examine the material basis on which they collectively hold a specific output as being justified." Secondly, Parasie sheds light on "the often tortuous history of how justified beliefs are collectively produced in relation to artifacts," following the lead of Latour and Woolgar (1979) in their study of how science is produced in the laboratory. In studying a 19-month investigation by CIR, Parasie shows how a heterogeneous team of investigative reporters, computer-assisted reporters, and programmer-journalists works through epistemological tensions to develop a shared epistemic culture, one connected with the material artifacts of data-oriented technologies. In all, Parasie makes key distinctions between "hypothesis-driven" and "data-driven" paths to journalistic revelations, in line with Coddington's conceptual mapping; he also highlights the interplay of materiality, culture, and practice, much as Anderson prescribes.

These articles are followed by three that take up algorithms and automation, pointing to matters of "autonomous decision-making" (Diakopoulos 2014) and the journalistic consequences of such developments for organizational and professional norms and routines. In the first article, Mary Lynn Young and Alfred Hermida (2014) examine the emergence of computationally based crime news at *The Los Angeles Times*. Following Boczkowski's (2004) theorizing about technological adaptation in news media organizations, they find that "computational thinking and techniques emerged in a (dis)continuous evolution of organizational norms, practices, content, identities, and technologies that interdependently led to new products." Among these products was a series of automatically generated crime stories, or "robo-posts," to a blog tracking local homicides. This concept of "algorithm as journalist," they argue, raises questions about "how decisions of inclusion and exclusion are made, what styles of reasoning are employed, whose values are embedded into the technology, and how they affect public understanding of complex issues."

This interest in interrogating the algorithm is further developed in Nicholas Diakopoulos' (2014) provocative notion of "algorithmic accountability reporting," which he defines as "a mechanism for elucidating and articulating the power structures, biases, and influences that computational artifacts exercise in society." In effect, he argues for flipping the computational journalism paradigm on its head, at least in this instance: instead of building another computational tool to enable news storytelling, technologists and journalists instead can use reverse engineering to investigate the algorithms that govern our digital world and unpack the crux of their power: autonomous decision-making. Understanding algorithmic power, in this sense, means analyzing "the atomic decisions that algorithms make, including *prioritization*, *classification*, *association*, and *filtering*" (original emphasis). Furthermore, Diakopoulos uses five case studies to consider the opportunities and challenges associated with doing algorithm-focused accountability journalism. He thus contributes to the literature both a theoretical lens through which to scrutinize the relative transparency of public-facing algorithms as well as an empirical starting point for understanding the potential for and limitations of such an approach, including questions of human resources, law, and ethics.

Lastly among these three, Matt Carlson (2014) explains what begins to happen as "the role of big data in journalism shifts from reporting tool to the generation of news content" in the form of what he calls "automated journalism." The term refers to "algorithmic processes that convert data into narrative news texts with limited to no human intervention beyond the initial programming." Among the data-oriented practices emerging in journalism, he says, "none appear to be as potentially disruptive as automated journalism," insofar as it calls up concerns about the future of journalistic labor, news compositional forms, and the very foundation of journalistic authority. By analyzing Narrative Science and journalists' reactions to its automated news services, Carlson shows how this "technological drama" (cf. Pfaffenberger 1992) reveals fundamental tensions not only about the work practices of human journalists but also what a future of automated journalism may portend for "larger understandings of what journalism is and how it ought to operate." Among other issues going forward, he says, "questions need to be asked regarding whether an increase in algorithmic judgment will lead to a decline in the authority of human judgment."

Before rushing headlong into robot journalism, however, quantitative journalism in its most basic form is still searching for institutional footing in many parts of the world. In exploring the difficulties for data journalism in French-speaking Belgium, Juliette de Maeyer et al. (2014) offer a much-needed reminder that the take-up of such journalism is neither consistent nor complete. Moreover, they argue that journalism (and hence data journalism) must be understood "as a socio-discursive practice: it is not only the production of (data-driven) journalistic artefacts that shapes the notion of (data) journalism, but also the discursive efforts of all the actors involved, in and out of the newsrooms." By mapping the discourse within this small media system, they uncover "a cartography of who and what counts as data journalism," within which they find divisions around the duality of "data" and "journalism" and between "ordinary" versus "thorough" forms of data journalism. These discourses disclose the various obstacles, many of them structural and organizational, that hinder the development of data journalism in that region. Among their respondents who have engaged in the actual practice of data journalism,

there seems to be an overall feeling of resignation. There might have been a brief euphoric phase after the first encounter with the concept of data journalism, but journalists who return from trainings full of ideas and ambitious projects are quickly caught again in the constraints of routinized news production.

Like Anderson and Parasie in this issue, the authors draw upon Bruno Latour (2005), in this case to suggest that data journalism is clearly a "matter of concern" in French-speaking Belgium even while there is a relative absence of data journalism artifacts, or "matters of fact" that can be displayed as evidence. Overall, de Maeyer and colleagues demonstrate how data journalism may "exist as a discourse (re)appropriated by a range of actors, originating from different—and sometimes overlapping—social worlds," allowing us to understand the uneven and sometimes incoherent path through which experimentation may lead to implementation (or not).

Finally, and befitting the opening discussion about the big deal of big data, the concluding article takes up this question: If big data is a wide-scale social, cultural, and technological phenomenon, what are its particular implications for journalism? Seth Lewis and Oscar Westlund (2014b) suggest four conceptual lenses—epistemology, expertise, economics, and ethics—through which to understand the present and potential applications of big data for journalism's professional logic and its industrial production. These conceptual approaches, distinct yet interrelated, show "how journalists and news media organizations are seeking to make sense of, act upon, and derive value from big data." Ultimately, the developments of big data, Lewis and Westlund posit, may have transformative meaning for "journalism's ways of *knowing* (epistemology) and *doing* (expertise), as well as its negotiation of *value* (economics) and *values* (ethics)." As quantitative journalism becomes more central to journalism's professional core, and as computational and algorithmic techniques likewise become intertwined with the business models on which journalism is supported, critical questions will continually emerge about the socio-material relationship of big data, journalism, and media work broadly. To what extent are journalism's cultural authority and technological practices changing in the context of (though not necessarily because of) big data? And how might such changes be connected with news audiences, story forms, organizational arrangements, distribution channels, and news values and ethics, among many other things? The articles in this issue—their cases, concepts, and critiques—offer a starting point for exploring such questions in the future.

REFERENCES

Anderson, C. W. 2013. "Towards a Sociology of Computational and Algorithmic Journalism." *New Media & Society* 15 (7): 1005–1021. doi:10.1177/1461444812465137.

Anderson, C. W. 2014. "Between the Unique and the Pattern: Historical Tensions in Our Understanding of Quantitative Journalism." *Digital Journalism*. [This issue] doi:10.1080/21670811.2014.976407.

Anderson, C. W., Emily Bell, and Clay Shirky. 2012. *Post-industrial Journalism: Adapting to the Present*. New York: Tow Center for Digital Journalism, Columbia University.

Appelgren, Ester, and Gunnar Nygren. 2014. "Data Journalism in Sweden: Introducing New Methods and Genres of Journalism into 'Old' Organizations." *Digital Journalism* 2 (3): 394–405. doi:10.1080/21670811.2014.884344.

Bell, Emily. 2012. "Journalism by Numbers." *Columbia Journalism Review* 51 (3): 48–49.

Boczkowski, Pablo J. 2004. *Digitizing the News: Innovation in Online Newspapers*. Cambridge, MA: MIT Press.

Boellstorff, Tom. 2013. "Making Big Data, in Theory." *First Monday* 18 (10). doi:10.5210/fm.v18i10.4869. http://firstmonday.org/ojs/index.php/fm/article/view/4869/3750.

boyd, danah, and Kate Crawford. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." *Information, Communication & Society* 15 (5): 662–679. doi:10.1080/1369118X.2012.678878.

Bruns, Axel. 2013. "Faster than the Speed of Print: Reconciling 'Big Data' Social Media Analysis and Academic Scholarship." *First Monday* 18 (10). doi:10.5210/fm.v18i10.4879. http://firstmonday.org/ojs/index.php/fm/article/view/4879.

Busch, Lawrence. 2014. "A Dozen Ways to Get Lost in Translation: Inherent Challenges in Large Scale Data Sets." *International Journal of Communication* 8. http://ijoc.org/index.php/ijoc/article/view/2160.

Carlson, Matt. 2014. "The Robotic Reporter: Automated Journalism and the Redefinition of Labor, Compositional Forms, and Journalistic Authority." *Digital Journalism*. [This issue] doi:10.1080/21670811.2014.976412.

Coddington, Mark. 2014. "Clarifying Journalism's Quantitative Turn: A Typology for Evaluating Data Journalism, Computational Journalism, and Computer-Assisted Reporting." *Digital Journalism*. [This Issue] doi:10.1080/21670811.2014.976400.

Couldry, Nick, and Joseph Turow. 2014. "Advertising, Big Data and the Clearance of the Public Realm: Marketers' New Approaches to the Content Subsidy." *International Journal of Communication* 8. http://ijoc.org/index.php/ijoc/article/view/2166.

Crawford, Kate, Kate Miltner, and Mary L. Gray. 2014. "Critiquing Big Data: Politics, Ethics, Epistemology." *International Journal of Communication* 8. http://ijoc.org/index.php/ijoc/article/view/2167/1164.

Diakopoulos, Nicholas. 2014. *Algorithmic Accountability Reporting: On the Investigation of Black Boxes*. New York: Tow Center for Digital Journalism, Columbia University. doi:10.1080/21670811.2014.976411.

Diakopoulos, Nicholas. 2014. "Algorithmic Accountability: Journalistic Investigation of Computational Power Structures." *Digital Journalism*. [This Issue]. doi:10.1080/21670811.2014.976411.

Driscoll, Kevin, and Shawn Walker. 2014. "Working within a Black Box: Transparency in the Collection and Production of Big Twitter Data." *International Journal of Communication* 8. http://ijoc.org/index.php/ijoc/article/view/2171

Fink, Katherine, and C. W. Anderson. 2014. "Data Journalism in the United States: Beyond the 'Usual Suspects'." *Journalism Studies*. doi:10.1080/1461670X.2014.939852.

Flew, Terry, Christina Spurgeon, Anna Daniel, and Adam Swift. 2012. "The Promise of Computational Journalism." *Journalism Practice* 6 (2): 157–171. doi:10.1080/17512786.2011.61665.

Franklin, Bob. 2014. "The Future of Journalism: In an Age of Digital Media and Economic Uncertainty." *Digital Journalism* 2 (3): 254–272.

González-Bailón, Sandra. 2013. "Social Science in the Era of Big Data." *Policy & Internet* 5 (2): 147–160. doi:10.1002/1944-2866.POI328.

Gray, Jonathan, Liliana Bounegru, and Lucy Chambers. 2012. *The Data Journalism Handbook*. Sebastopol, CA: O'Reilly Media.

Gynnild, Astrid. 2014. "Journalism Innovation Leads to Innovation Journalism: The Impact of Computational Exploration on Changing Mindsets." *Journalism* 15 (6): 713–730. doi:10.1177/1464884913486393.

Hamilton, James T., and Fred, Turner. 2009. "Accountability through Algorithm: Developing the Field of Computational Journalism." Report from the Center for Advanced Study in the Behavioral Sciences, Summer Workshop 27-41, Duke University, Durham, NC.

Hermida, Alfred. 2013. "#Journalism: Reconfiguring Journalism Research about Twitter, One Tweet at a Time." *Digital Journalism* 1 (3): 295–313. doi:10.1080/21670811.2013.808456.

Howard, Alexander Benjamin. 2014. *The Art and Science of Data-Driven Journalism*. New York: Tow Center for Digital Journalism, Columbia University.

Karlsen, Joakim, and Eirik Stavelin. 2014. "Computational Journalism in Norwegian Newsrooms." *Journalism Practice* 8 (1): 34–48. doi:10.1080/17512786.2013.813190.

Karpf, David. 2012. "Social Science Research Methods in Internet Time." *Information, Communication & Society* 15 (5): 639–661. doi:10.1080/1369118X.2012.665468.

Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor-Network-theory*. New York: Oxford University Press.

Latour, Bruno, and Steve Woolgar. 1979. *Laboratory Life. the Construction of Scientific Facts*. Princeton, NJ: Princeton University Press.

Lazer, David, Alex (Sandy) Pentland, Lada Adamic, and Sinan Aral et al. 2009. "Life in the Network: The Coming Age of Computational Social Science." *Science* 323 (5915): 721–723.

Lewis, Seth C. 2012. "The Tension between Professional Control and Open Participation: Journalism and Its Boundaries." *Information, Communication & Society* 15 (6): 836–866. doi:10.1080/1369118X.2012.674150.

Lewis, Seth C., Rodrigo Zamith, and Alfred Hermida. 2013. "Content Analysis in an Era of Big Data: A Hybrid Approach to Computational and Manual Methods." *Journal of Broadcasting & Electronic Media* 57 (1): 34–52. doi:10.1080/08838151.2012.761702.

Lewis, Seth C., and Nikki Usher. 2013. "Open Source and Journalism: Toward New Frameworks for Imagining News Innovation." *Media, Culture & Society* 35 (5): 602–619. doi:10.1177/0163443713485494.

Lewis, Seth C., and Nikki Usher. 2014. "Code, Collaboration, and the Future of Journalism: A Case Study of the Hacks/Hackers Global Network." *Digital Journalism* 2 (3): 383–393. doi:10.1080/21670811.2014.895504.

Lewis, Seth C., and Oscar Westlund. 2014a. "Actors, Actants, Audiences, and Activities in Cross-Media News Work: A Matrix and a Research Agenda." *Digital Journalism*. doi:10.1080/21670811.2014.927986.

Lewis, Seth C., and Oscar Westlund. 2014b. "Big Data and Journalism: Epistemology, Expertise, Economics, and Ethics." *Digital Journalism*. [This issue] doi:10.1080/21670811.2014.976418.

Mahrt, Merja, and Michael Scharkow. 2013. "The Value of Big Data in Digital Media Research." *Journal of Broadcasting & Electronic Media* 57 (1): 20–33. doi:10.1080/08838151.2012.761700.

de Maeyer, Juliette, Manon Libert, David Domingo, François Heinderyckx, and Florence Le Cam. 2014. "Waiting for Data Journalism: A Qualitative Assessment of the Anecdotal Take-up of Data Journalism in French-speaking Belgium." *Digital Journalism*. [This issue] doi:10.1080/21670811.2014.976415.

Manovich, Lev. 2012. "Trending: The Promises and the Challenges of Big Social Data." In *Debates in the Digital Humanities*, edited by M. K. Gold, 460–475. Minneapolis, MN: The University of Minnesota Press.

Manyika, James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela H. Byers. 2011. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute. http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation.

Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston, MA: Houghton Mifflin Harcourt.

Morozov, Evgeny. 2013. *To Save Everything, Click Here: The Folly of Technological Solutionism*. New York: PublicAffairs.

Oboler, Andre, Kristopher Welsh, and Lito Cruz. 2012. "The Danger of Big Data: Social Media as Computational Social Science." *First Monday* 17 (7–2). http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3993/3269.

O'Neil, Cathy, and Rachel Schutt. 2013. *Doing Data Science: Straight Talk from the Frontline*. Sebatspol, CA: O'Reilly Media.

Parasie, Sylvain. 2014. "Data-driven Revelation: Epistemological Tensions in Investigative Journalism in the Age of 'Big Data'." *Digital Journalism*. [This issue] doi:10.1080/21670811.2014.976408.

Parasie, Sylvain, and Eric Dagiral. 2013. "Data-driven Journalism and the Public Good: 'Computer-Assisted-Reporters' and 'Programmer-journalists' in Chicago." *New Media & Society* 15 (6): 853–871. doi:10.1177/1461444812463345.

Pfaffenberger, Bryan. 1992. "Technological Dramas." *Science, Technology & Human Values* 17 (3): 282–312. doi:10.1177/016224399201700302.

Pitt, Fergus. 2014. *Sensors and Journalism*. New York: Tow Center for Digital Journalism, Columbia University.

Powers, Matthew. 2012. "'In Forms That Are Familiar and Yet-to-Be Invented': American Journalism and the Discourse of Technologically Specific Work." *Journal of Communication Inquiry* 36 (1): 24–43. doi:10.1177/0196859911426009.

Provost, Foster, and Tom Fawcett. 2013. "Data Science and Its Relationship to Big Data and Data-Driven Decision Making." *Big Data* 1 (1): 51–59. doi:10.1089/big.2013.1508.

Ryfe, David M. 2012. *Can Journalism Survive? An Inside Look at American Newsrooms*. Cambridge; Malden, MA: Polity Press. Web.

Usher, Nikki. 2014. *Making News at The New York Times*. Ann Arbor: University of Michigan Press.

Young, Mary Lynn, and Alfred Hermida. 2014. "From Mr. and Mrs. Outlier to Central Tendencies: Computational Journalism and Crime Reporting at the Los Angeles Times." *Digital Journalism*. [This Issue] doi:10.1080/21670811.2014.976409.

Seth C. Lewis, School of Journalism and Mass Communication, University of Minnesota–Twin Cities, USA. E-mail: sclewis@umn.edu. Web: www.sethlewis.org