BIG DATA AND JOURNALISM
Epistemology, expertise, economics, and ethics

Seth C. Lewis and **Oscar Westlund**

Big data is a social, cultural, and technological phenomenon—a complex amalgamation of digital data abundance, emerging analytic techniques, mythology about data-driven insights, and growing critique about the overall consequences of big-data practices for democracy and society. While media and communication scholars have begun to examine and theorize about big data in the context of media and public life broadly, what are the particular implications for journalism? This article introduces and applies four conceptual lenses—epistemology, expertise, economics, and ethics—to explore both contemporary and potential applications of big data for the professional logic and industrial production of journalism. These distinct yet inter-related conceptual approaches reveal how journalists and news media organizations are seeking to make sense of, act upon, and derive value from big data during a time of exploration in algorithms, computation, and quantification. In all, the developments of big data potentially have great meaning for journalism's ways of knowing (epistemology) and doing (expertise), as well as its negotiation of value (economics) and values (ethics). Ultimately, this article outlines future directions for journalism studies research in the context of big data.

KEYWORDS algorithms; big data; computational journalism; epistemology; expertise; media economics; media innovation; journalism ethics; technology

Introduction

Big data, the buzzword *du jour*, carries with it all manner of hype and hope—and hesitation about its social consequences (Crawford, Miltner, and Gray 2014). By one account, "Big data is poised to reshape the way we live, work, and think" (Mayer-Schönberger and Cukier 2013, 190), and by another, "The data explosion [will change] how we do business. Every interaction, every communication, every touchpoint creates a digital breadcrumb—a piece of data that can be analyzed and manipulated" (cited in Dwoskin 2014). While it is impossible to know the prescience of such predictions, there is a growing body of evidence that *something* important is changing in the nature of data —in the volume and variety of its digital representation, in its collection and analysis on a massive scale, and in its ultimate potential for yielding social, cultural, and monetary value, even as these developments simultaneously raise troubling questions about privacy, accuracy, and ethics (boyd and Crawford 2012). Big data is made available by the growing ubiquity of mobile devices, sensors, "smart" machines, digital trace data, digital

repositories and archives, and other fragments of social and natural activity represented by clicks, tweets, likes, GPS coordinates, timestamps, and so on (for a related discussion of the "internet of things," see Howard forthcoming). The sheer volume of digital data and its boundless growth is staggering (Mayer-Schönberger and Cukier 2013), and just as significant is the increasing ease with which standard computer software can manage and manipulate data sets that once required supercomputers, thereby magnifying this episode of digital data exploration (Manovich 2012).

This "big data moment" is not merely a technological transition toward data deluge. Rather, it is a sociotechnical phenomenon with cultural, economic, and political origins and implications; it is, indeed, a mythology as much as a science or business (boyd and Crawford 2012; Crawford et al. 2014, 1664). As such, the very term "big data" deserves scrutiny of the kind applied to the likes of "Web 2.0" (Coleman 2013). While we take up such concerns later in this article, for now it suffices to suggest, as Boellstorff (2013) does, that even while "there is no unitary phenomenon 'big data' … the impact of big data is real and worthy of sustained attention." Moreover, questions of data privacy and surveillance raised by the Edward Snowden revelations, among other incidents, make big data a matter of public as well as professional concern.

In this article, we focus on four concepts that highlight the (potential) implications of big data for journalism: *epistemology*, *expertise*, *economics*, and *ethics*. These "Four E's," chosen from our review of the literature and observations of changes in different news, media, and information technology industries, appear to touch on salient dynamics of big data and its meaning for journalism, both for academic theory and professional practice. While the big-data phenomenon raises many relevant questions for news media, some of the most essential have to do with the legitimation of new claims about knowledge and truth (epistemology); the negotiation of occupational status, authority, and skill sets as new specializations are developed and deployed (expertise); the potential for and challenges of new efficiencies, resources, innovations, value creations, and revenue opportunities (economics); and the concerns raised by these developments for the norms and values that guide human decision-making and technological systems design (ethics). These four concepts serve as a dynamic set of lenses through which to view contemporary developments and envision future research opportunities.

Defining Big Data

Big data is a rather plastic concept, assuming different meanings in different contexts for different purposes—from policing to city planning to predicting preferences for breakfast cereal (Crawford et al. 2014). In strict computing terms, big data refers to data sets that are too large for standard computer memory and software to process. Or, put another way, "when the *volume*, *variety* and *velocity* of the data are increased, the current techniques and technologies may not be able to handle storage and processing of the data" (Suthaharan 2014, 70, emphasis added). Beyond technical specifications, however, big data can refer as much to *processes* surrounding data—and the resulting *products* of information about a great many people, places, and things—as to the scope of data itself. As Mayer-Schönberger and Cukier (2013, 6) describe it: "Big Data refers to our newfound ability to crunch a vast quantity of information, analyze it instantly, and draw sometimes astonishing conclusions from it."

From yet another perspective, and one perhaps more relevant for social researchers trying to understand the wider implications involved, big data can be defined as a social, cultural, and technological phenomenon that sits at the interplay of three dynamics:

(1) *Technology*: maximizing computation power and algorithmic accuracy to gather, analyze, link, and compare large data sets.
(2) *Analysis*: drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims.
(3) *Mythology*: the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy (boyd and Crawford 2012, 663).

The third of those—mythology—calls up a critical stance that is essential but often lost amid the clamor in business, policy, and even some scholarly corners for "big data solutions" (Crawford et al. 2014). Such a critique means worrying less about the "bigness" of data and attending more to how data comes to be seen as "big" in social relevance and normative valence. Moreover, this means recognizing the discursive work being performed by the term itself: "big data" may be as much a marketing term and a techno-utopian vision as it is a material phenomenon. Like terms such as "platform" (Gillespie 2010) and "digital divide" (Epstein et al. 2011), a term like "big data" does not emerge out of thin air but rather is "drawn from the available cultural vocabulary by stakeholders with specific aims, and carefully massaged so as to have particular reso- nance for particular audiences inside particular discourses" (Gillespie 2010, 359). Impor- tantly, to follow Gillespie's reasoning on the politics of "platforms" and terms like it, "These are efforts not only to sell, convince, persuade, protect, triumph or condemn, but to make claims about what these technologies are and are not, and *what should and should not be expected of them*" (359, emphasis added).

Big data invokes a wide range of normative claims and practical implications for journalism as a professional practice and an organizational production—from knowl- edge work and economic rationale to practical skills and philosophical ethics. Because of its contested nature, "big data" as a term is a messy business (Crawford et al. 2014), and yet it remains the most succinct way of referring to a larger and complicated set of factors at play in technology and society as well as in technology and journalism. In this sense, we draw upon boyd and Crawford's definition of big data as a social, cul- tural, and technological phenomenon—one representing various philosophies, prac- tices, and promises. For journalism, big data embodies emerging ideas about, activities for, and norms connected with data sets, algorithms, computational methods, and related processes and perspectives tied to quantification as a key paradigm of informa- tion work. We assume that big data is neither good nor bad for journalism but never- theless freighted with potential and pitfall, depending on how it is imagined and implemented—and, crucially, toward what purposes and in whose interests.

Big Data in the Journalism Profession and Media Industry

The literature points to the need for a conceptual starting point for the study of big data in the salient case of journalism. Like the scientific, corporate, and government

sectors broadly, the media industry must confront the question of what to *do* with all this data. This comes as many legacy news organizations struggle to find their way amid disruptions to the professional authority, business models, and traditional logics of news production and distribution (Anderson, Bell, and Shirky 2012; Lewis 2012; Ryfe 2012; Usher 2014). Journalism has long been familiar with data and databases as an object of news work and journalistic evidence, as evident in decades of computer-assisted reporting (CAR) and even older forms of information visualization (for history and discussion, see Cox 2000; Fink and Anderson 2014; Howard 2014; Powers 2012). But the database—which Manovich (1999) boldly suggests is to the digital era what narrative, in novels and cinema, was to the modern era—has assumed a particularly conspicuous role in contemporary journalism (Schudson 2010). Indeed, the larger turn toward digitization of information in recent times has been connected with a greater role in journalism for the *techniques* of computer and data sciences—from programming and algorithms to machine learning and probability models (Diakopoulos 2014; Stavelin 2014)—as well as the *ethos* of open-source software development and its emphasis on making data sets transparent and interactive (Lewis and Usher 2013; Parasie and Dagiral 2013). This take-up can be seen in the recent formation of data-specialist teams at leading news organizations, the development of data-focused journalism education courses and degrees, and the data-centric practices of "explainer" news startups like FiveThirtyEight and Vox (see Anderson 2013; Gynnild 2014; Howard 2014; Pitt 2014). Thus, *data*—whether "big" in the sense of being too complex for traditional database management software, or simply "big" in its potentially transformative import—has taken on particular relevance for news. "The open question in 2014," one report noted, "is not whether data, computers, and algorithms can be used by journalists in the public interest, but rather how, when, where, why, and by whom" (Howard 2014, 4).

Beyond the data-driven journalism on the editorial side of news media organizations, data manipulation has equal if not greater interest for business-side strategies and applications, particularly as market considerations have come to the fore amid declining revenues and fragmenting audiences (for a discussion of contemporary advertising and marketing dynamics, see Turow 2011; Couldry and Turow 2014). Moreover, as big data gains significance, technologists are needed to identify and appropriate suitable technological systems and solutions from external providers, or develop and reconfigure such systems and solutions themselves. These technologists, in turn, may form an important bridging function in negotiating technological systems and solutions across the editorial and business domains of the organization (Lewis and Westlund 2014; Westlund 2011)—potentially complicating sharp divisions that have long existed (or been assumed to exist) between news and business/marketing departments (Coddington, forthcoming; Klinenberg 2005). Big data as a social, cultural, and technological phenomenon (boyd and Crawford 2012) thus serves as a conceptual lens through which to understand how journalism—as both professional field and commercial enterprise—is seeking to make sense of, act upon, and derive value from the growing array of digital data in public life.

This article, like this special issue as a whole, does not celebrate or fetishize big data in the context of journalism. Nor do we assume that most news media organizations are "working" with big data in the same way that astronomers, biologists, and corporate data-miners are analyzing vast troves of data. Rather, we argue that the

implications of big data for journalism, while in many cases still hypothetical, require conceptual heuristics that can guide future research. Just as journalism is trying to make sense of big data, so too must journalism studies develop toolkits for understanding what it means to the practice and perception of news. The various agents in news media organizations may make sense of the big-data phenomenon in diverse ways, and consequently also approach it in different ways: by resisting, adapting to, intervening against, or shaping it. Moreover, they may act proactively or reactively, focusing on none, one, or several of the strategic objectives that reside at the intersection of editorial, business, and technological interests and practices (Lewis and Westlund 2014). What is needed, therefore, is a conceptual starting point.

Conceptual Lenses

The full spectrum of research at the interplay of data, computation, and journalism has been reviewed and covered well by Coddington's (2014) contribution to this special issue. Instead, this article will introduce concepts and future research questions for exploring this sociotechnical phenomenon in journalism. There are, of course, a great number of entry points for such an undertaking. In the journalism studies literature thus far, related conceptual development and research agendas have mostly focused on "computational journalism" (Anderson 2013; Flew et al. 2012; Stavelin 2014) or "computational exploration in journalism" (Gynnild 2014). But missing in these analyses of "computation" is a more focused treatment of "big data," as a concept and a phenomenon (with some exceptions, such as Fairfield and Shtein 2014). By contrast, in the media and communication literature more broadly, scholars have critically examined such things as big-data target marketing and media production (Couldry and Turow 2014), the role of smartphone users as appropriators of big data (Nafus and Sherman 2014), big data's metaphorical framing (Puschmann and Burgess 2014), and the implications of big data for communication research methods (González-Bailón 2013). Our review of these analyses reveals a recurring focus on at least four themes that have particular relevance for the case of journalism, as introduced above: *epistemology*, *expertise*, *economics*, and *ethics*. These concepts open up lines of inquiry for understanding journalism in the context of big data.

In the sub-sections that follow, we (1) introduce the concepts generally; (2) discuss their meaning for big data and journalism; (3) provide examples or possible applications, often in the context of news production and distribution; and (4) raise questions for future research. In doing this, we build on our previous work in the context of news media organizations by keeping an eye on inter-relationships among types of social *actors* (e.g., journalists, businesspeople, and technologists), technological *actants* (e.g., algorithms and content management systems), and *audiences* (e.g., whether passive or active)—all interpolated through the *activities* of media production and distribution (Lewis and Westlund 2014; Westlund and Lewis 2014).

Epistemology

Epistemology, as a theory of knowledge, differs from ontology. In the philosophy of science, ontology refers to fundamental inquiries into the nature of existence. That

is, ontology refers to what is said to "be" in the world, or the "science of what is" (Smith 2001, 79). While the world undoubtedly exists in the form of nature, people, or events, any attempt to represent the world will, in fact, turn into some sort of re-presentation, with inherent limitations. Epistemology thus points to the nature and boundaries of human knowledge about the world and the determination of truth in that process of re-presentation. The term derives from the Greek *episteme*, which means knowledge, and *epistanai*, which means to understand. A fundamental issue in episte-mology concerns the work of legitimizing certain types of information as knowledge relative to others. The academy, like other knowledge-producing fields of practice, long has developed epistemologies that shape what counts in this regard (Schon 1995).

Among the most influential knowledge-producing institutions of the modern era, journalism has a distinct epistemology. It outlines the "*rules, routines* and *institutional-ized procedures* that operate within a social setting and decide the form of the knowl-edge produced and the knowledge claims expressed (or implied)" (Ekström 2002, 260, original emphasis), as well as shapes the justifications that journalists make—to them-selves and others—in defense of their truth claims (Ettema and Glasser 1998). Put sim-ply, this epistemology is about how journalists know what they know—and why that matters for the knowledge practices and products in which they are engaged. Research into news production has shown how journalists develop methods for adjudicating knowledge claims in a routinized fashion (Tuchman 1978), for instance by adhering to ideals such as objectivity and practices such as multiple sourcing (Wiik 2010). However, the introduction of various technologies into news work has raised questions about the relative knowledge value associated with "technologically specific forms of work," from photojournalism of the past to programmer-journalism of the present (Powers 2012). The introduction of computers into news work, in the form of early computer-assisted reporting (CAR) nearly a half-century ago, pointed to hope for "precision journalism"— the potential for achieving greater accuracy through the use of databases, surveys, and an overall combination of computer and social science (Meyer 1973). While computers and their capacity for data analysis have improved significantly since that time, journal-ists have found it challenging to move beyond their established epistemology. This is true of the CAR tradition, one built on the belief that "data have no journalistic value on their own" and therefore journalists must work to find the story "hidden" in the data (Parasie and Dagiral 2013, 859); and it is true of the normative paradigm that positions journalists as essential knowledge-producers for society, bound up in their professional control of news information (Lewis 2012). Amid this tension are questions about the role that technology might play in developing the capacities for and practices of knowl-edge production in journalism—for example, in the form of augmented reality for digi-tal storytelling (Pavlik and Bridges 2013) or technological systems for customizing diverse types of news for diverse types of audiences (Westlund 2013; Gynnild 2014).

By extension, big data offers similar opportunities for rethinking the epistemolo-gies of journalism. For example, the extraction, combination, and analysis of big data may reveal new possibilities for investigative journalism (Parasie 2014), and with it a potentially stronger link between the production of social facts in journalism and mat-ters of "science" and "precision." Beyond the realm of investigative journalism, where data has long held a special role in the CAR tradition, such beliefs about the power of data are spreading to other types of news amid the "quantitative turn" in journalism broadly (Coddington 2014). While ideas about "data" and corresponding "computation"

enabling greater and more rigorous forms of knowledge are commonplace among some academics, practitioners, and pundits (see discussion in Anderson 2013), these ideas should be scrutinized for their underlying assumptions. Data should not to be taken as a proxy for the "science of what is," in the ontological sense, but rather as one form of epistemological knowledge in which numbers carry great significance. Big data, like any data, does not represent an objective truth. As Gitelman (2013) and others have stressed, "raw data is an oxymoron." The figures yielded by big data—even if enormous, robust, and highly correlated—still require interpretation.

Big data, as a set of technological processes as well as a key source of power, thus opens new paths for imagining how journalistic investigations develop epistemologically relevant revelations. More broadly, big data has implications for rethinking the epistemologies of news production and news distribution.

First, news production may be understood as a process of access/observation, selection/filtering, and processing/editing (Domingo et al. 2008)—each component connected with journalistic epistemology. Consider four brief examples: first, access/observation may involve computerized gatekeeping, or "watchdogging in code," through which journalists use actants to continuously and automatically monitor what politicians are doing, as Stavelin (2014) explored in the Norwegian context. A second example, involving all three steps of news production, regards programmer-journalists at *The Los Angeles Times*: they developed an algorithm to record earthquake notifications, process such alerts into epistemological facts, and facilitate easy editing and rapid publication upon human approval. A third example relates to how organizations may offer customers and readers forms of technology-led customization of news (Westlund 2013), as in Narrative Science's automated publishing of news stories based on financial reports released by public companies (Carlson 2014). A fourth example involves the Truth Teller prototype at *The Washington Post*, which combines speech-to-text algorithms with databases of "facts" to fact-check political speech in real-time. In each case, big data and related approaches present new facets for understanding the epistemology of transforming raw *information* into journalistic *truth*.

Turning to news distribution, big data is connected with emerging representations of digital journalism such as infographics, interactive data visualizations, and customizable probability models, among others (Howard 2014; Smit et al. 2014). These news products, in turn, carry certain epistemological assumptions about how audiences might acquire knowledge, as users are encouraged to "play" with the data to comprehend a particular and personalized version of the news narrative. For instance, some news organizations have sought to make data sets more accessible, transparent, and exploratory for users, in line with the ethos of open-source software and open-government advocacy (Lewis and Usher 2013; Parasie and Dagiral 2013). Others have invited audiences to participate on a massive collaborative project, as in the case of the public radio station WNYC asking listeners to build small DIY computer sensors to contribute local temperature readings to help predict the arrival of cicadas (Pitt 2014). Processes of news delivery and audience engagement in this big-data context thus present new questions not only about participation in journalism but also, and perhaps especially, about the legitimation of knowledge in and through such data-driven participation.

This brings us to a set of emerging research questions. In his analysis of television news, Ekström (2002) conceptualized journalistic epistemology in three parts: form of knowledge (i.e., medium-specific concerns, in his case those associated with television

as a media form), production of knowledge (i.e., professional norms and routines), and public acceptance of knowledge claims (i.e., the conditions for social legitimacy). Each of these perspectives, brought into conversation with our agent-oriented emphasis above, leads to various questions for future research: How might the particular form of big-data journalism be associated with particular types of knowledge claims? What are the institutionalized routines and procedures that social actors adopt to guide the production of data-backed knowledge claims, and how are such routines conditioned to ensure that claims are legitimate and justified? In what ways do audiences learn from certain knowledge claims made via big-data news products, and what types of conditions must be satisfied to ensure that publics accept such claims? Ultimately, to care about how and why publics come to accept certain knowledge claims is to care about the social actors behind such processes: the experts and their expertise.

Expertise

In the broad sense, the term "expertise" comes from the Latin root *experiri*, meaning "to try," and generally refers to "the know-how, the capacity to get a task accomplished better and faster because one is more experienced"—hence *expertus*, or "tried" (Eyal 2013, 869). Because *expert* (the social actor) is connected with *expertise* (the specialized know-how), the study of expertise has long been associated with the sociology of professions, or the study of how some actors become seen as experts relative to others. This line of research has examined how occupations (and the experts who constitute them) work to forge and maintain "jurisdictional control" over the boundaries around a body of abstract knowledge and the application of that knowledge through work practices—thus allowing professionals to claim autonomy, authority, and other social and material benefits associated with being granted special recognition in society (Abbott 1988, 60). Whether in an occupational context or beyond, expertise thus functions as a key boundary marker setting apart those possessing specialized knowledge and experience. If epistemology directs attention to the *process* through which information becomes recognized as legitimate knowledge, expertise points to the *people* behind the processes. Even more broadly, as Eyal (2013, 863) has shown, expertise may be less of a professional attribution or a "real" thing possessed by individuals and more of "a network linking together agents, devices, concepts, and institutional and spatial arrangements."

Are journalists truly experts then? Just as journalism holds contested status as a profession, lacking many of the protective trappings enjoyed by professions such as law and medicine, journalistic expertise also has something of an in-between nature: while journalists are known for certain skills, such as storytelling, that distinction tells us little about whether journalists are actually experts *in something* (Schudson and Anderson 2008). Moreover, journalism's claim to social expertise through its "professional logic" (Lewis 2012), a bargain to control the production and distribution of news on society's behalf, is beset by challenges to authority (Carlson, forthcoming). Against that backdrop, Reich (2012) has offered a new picture of journalistic expertise. Drawing on Collins and Evans' (2007) influential typology of expertise—one that presumes expertise is a "real" feature of socialization and experience, and therefore may be classified across social domains—Reich argues that journalists can be understood as "interactional" experts:

their expertise lies in their ability to work with and among other types of experts, ultimately synthesizing and translating others' specialized knowledge for non-experts. Furthermore, Reich suggests that journalists develop a *bipolar* form of interactional expertise because they also must manage interactions with lay audiences, thus negotiating a dual process of engagement that constitutes its own kind of expertise. Finally, within the sociology of knowledge, there is a greater recognition for a competence-based approach: "Expertise is now seen more and more as something practical—something based on what you can do rather than what you can calculate or learn" (Collins and Evans 2007, 23; Reich 2012). Skills, technical or otherwise, are thus recognized as key benchmarks of expert distinction, in journalism as in other social domains.

These frameworks, whether constructivist or normative in nature, offer useful entry points for conceptualizing what may become of journalistic expertise amid big data. These might be described as (1) social interactions, (2) networked interactions, and (3) skill sets.

Social Interactions

As Reich (2012) notes in elaborating on journalism as bipolar interactional expertise, some newsworkers may have primarily source-interactional expertise (e.g., long-time beat reporters), others may have primarily audience-interactional expertise (e.g., editors who hear from readers), and others may have some combination (e.g., columnists and commentators). How might the social, cultural, and technological nature of big data affect the character of these interactions with sources and audiences? For example, consider what journalists, long oblivious to their readers and viewers (Lowrey 2009), now know about their audiences via digital metrics, and how that affects decisions about what and how to present as news (Tandoc 2014). Moreover, what of the interactions around expertise that may be happening internally? Large news companies increasingly are hiring data scientists and other technical experts to make sense of data —both data as source material for journalistic storytelling as well as data on audiences for business purposes. How might expertise be developed and made manifest in and through these emerging types of interactions with data? For instance, at what point does a journalist become conversant in the "language of data" such that she can claim expertise in "interviewing" data as an expert source?

Networked Interactions

Next, consider the more socio-technical interactions of expertise that might occur between journalists (as actors) and machines (as actants). From Eyal's (2013) approach, the growing deployment of algorithms and automation in journalism might entail new arrangements of "networked expertise," altering how we imagine what it is that journalists know and how they represent that knowledge to the world. In this vein, Anderson (2013) has shown how the dividing lines of expertise between "original" reporting and "parasitic" news aggregation are hardly clear-cut. In fact, networks of social actors and technological actants, when viewed holistically, yield complicated renderings of journalistic expertise under different conditions of digitization. Thus, big-data forms of journalism,

such as data journalism (Howard 2014) or computational journalism (Stavelin 2014), raise questions about how human expertise is embedded in and through technical capacities, some of which may be programmed to perform relatively autonomously.

Skill Sets

Finally, as Collins and Evans (2007) remind us, expertise is manifest in actual, practical skills. Big data as a phenomenon and approach prioritizes certain skills, such as data analysis, computer programming, and visualization, drawn from disciplinary origins such as computer science, mathematics, and statistics (Mayer-Schönberger and Cukier 2013). What might these mean for journalism? Journalists have long worked with "data" of various kinds, building many Pulitzer Prize-winning investigations around data-driven analyses of public institutions. What is different now is that news organizations increasingly need computer programming, sophisticated back-end databases, and data science techniques—in essence, "code"—to comprehend increasingly ubiquitous data, and to publish it in ways that allow users to explore the data for themselves (Parasie and Dagiral 2013). *Data* and *code* thus constitute skills-based forms of expertise that news organizations are working to cultivate. Yet bridging the skills gap between journalists and technologists, or helping journalists develop such data-and-code skills, is neither easy nor broadly institutionalized as yet (Howard 2014; Lewis and Usher 2014). Nevertheless, consider how journalistic expertise might change if more journalists learned to write basic software, as Stavelin (2014) proposes, or if external "algorithmists"—expert reviewers of big-data analysis and predictions (Mayer-Schönberger and Cukier 2013)—were invited to critique computational journalism, much as ombudsmen function for news organizations today. To incorporate such skills into the news production and distribution process not only might lead to technological innovations but also alter the notion of what truly *counts* as expertise in journalism, perhaps complicating notions of status and authority both within and beyond the newsroom and the field at large (Ananny 2013a). Ultimately, perhaps these shifts have less to do with adapting interactional expertise through revised relationships between journalists and non-journalist experts, and more to do with the development or public display of a kind of "original" contributory expertise (Collins and Evans 2007), in which computationally minded journalists *are* the experts as they follow the methods of computer science.

Economics

Economics, with its etymology in the "management of household," is a discipline that studies the behaviors of agents in households and organizations, focusing on how resources are managed to achieve certain ends. Typically these agents are assumed to act rationally, making choices about how to use limited resources toward desired outcomes and strategic goals. Applied to the context of communication, *media economics* is defined as "the study of how media industries use scarce resources to produce content that is distributed among consumers in a society to satisfy various wants and needs" (Albarran 2002, 5). In all, such a focus takes up questions of media management, media business studies, and media innovations—covering the range of managerial strategies and tactics associated with media organizations and industries.

Many legacy news media companies around the world—especially local newspapers—have benefited financially from their standing as oligopolies or monopolies in a distinct geographical market, leading to impressive profit margins relative to other industries in recent decades (Picard 2010). Contemporary research, of course, has shown how such firms face a shrinking advertising base, fragmenting audiences, and rising competition from mobile, social, and digital media (Anderson et al. 2012). Amid a general call for media companies to innovate (Storsul and Krumsvik 2013; Westlund and Lewis, 2014), big data represents an opportunity for value creation through revised business processes as well as new products and services. Big data has obvious relevance for business-side revenue opportunities, allowing media companies to better understand and serve particular audiences and advertisers. Nevertheless, these developments come with corresponding concerns about the ultimate social and political outcomes of a world dominated by personalized digital media, where targeted advertising based on data mining leads to pressures on media companies (including news organizations) to personalize content in response (Couldry and Turow 2014). The march toward big-data personalization, in this view, threatens the very ecology of common knowledge upon which representative democracy depends (Sunstein 2009).

But seen another way, how might big data afford new value creation without undermining the church–state divide between business and editorial concerns that, for many journalists, is central to professional autonomy? Big data, in a basic sense, promises economic efficiency by enabling "more observation at less cost" (Crawford et al. 2014, 1666), as in the case of labor-saving "robot journalism" (Carlson, 2014) and technologically automated forms of journalism more generally (Westlund 2013). In another sense, big data may be associated with augmenting, rather than displacing, human labor by catalyzing new types of technologically enabled forms of news work (Powers 2012) or by allowing journalists to function more like "knowledge managers" who better gather, organize, and analyze disparate information flows in a community (Lewis and Usher 2013). Altogether, big data raises questions about the relative status and precarity of journalistic labor (Deuze 2008) amid the pursuit of new value propositions.

We consider, in brief, two ways of envisioning the value-creation opportunities for a journalism leveraging big data.

The first involves social actors, especially journalists but also technologists, manually drawing upon large data sets to report and present news in ways that differentiate their work from the traditional storytelling paradigm, thereby creating value for audiences interested in new types of news as well as creating distinction relative to commodity news in the marketplace. This shift has been called "method journalism" (Madrigal 2014), moving from an *area of coverage* (a topic, beat, or location of interest) to focus instead on the *method of coverage*. Several news startups are emblematic of this change, built around method-oriented objectives. FiveThirtyEight is "a data journalism organization," per founder Nate Silver's manifesto; Circa focuses on structuring news "data" (i.e., quotes, facts, and other atomic units of information) for mobile devices; and *The New York Times*' The Upshot emphasizes "plain-spoken, analytical journalism" based on data-driven analysis (quoted in Madrigal 2014). While those are US examples, there is growing evidence of similar data journalism efforts emerging in Europe and elsewhere (Lewis and Usher 2014). In these and other ways, data-focused initiatives may emerge as key forms of differentiation, traditionally held as essential for

companies to succeed (Porter 1980; Dimmick 2003). News media organizations, for the most part, have yet to realize the potential strategic value of data as a business model, both for attracting audiences and for providing database resources that can be monetized via subscription and consulting services (Howard 2014; see also Aitamurto and Lewis 2013; Ananny 2013a).

The second approach involves using technological actants such as algorithms and applications to automatically gather, link, compare, and act upon big data of interest to audiences (Lewis and Westlund 2014). These algorithms can be tailored to fit with the personalized preferences and behaviors of individuals, promoting specific types of news to specific individuals (Thurman 2011). Importantly, journalists and technologists need to actively assess what actants they are to use for such purposes (if acquired from external providers), or how they are to be configured (if developed internally). A diverse set of social actors within the organization—from journalists to data analysts to marketers—may need to collaborate in inscribing the technological actants with logics and values for their operation, raising questions about how journalistic values are built into algorithmic news products. Settling such matters across the editorial–business divide, and in conjunction with technology design teams, is no small feat, as Westlund (2012) found in his study of a Swedish newspaper's development of a mobile news app. Nevertheless, as Anderson et al. (2012) have argued, automation offers an important yet underexplored avenue for news media to cut expenses (e.g., by no longer wasting resources on stories that a robot could write just as well) and simultaneously create value (e.g., by redeploying humans toward projects where they uniquely can contribute).

Ultimately, economics directs us to questions about how news media organizations might strategically manage their human and technological resources in relation to big data. The economics of big data is a challenge for media innovation that potentially involves the full range of social actors: journalists to derive editorial value from growing bodies of data, technologists to configure actants for automation and audiences in new ways, and businesspeople to negotiate marketers' demands for increasingly precise audience data and personalized media experiences. All of these developments raise overlapping questions relevant to research and theory. For journalism studies, to what extent do big-data processes contribute to a weakening wall between editorial and business concerns, with what kind of outcomes for news values and judgment as well as revenues and profits? How does an emphasis on value creation influence the distinct practices and products of news work, and how do such things compare across countries or cultures? For media economics and management, how do the development and deployment of big-data approaches correspond with certain theories of strategic management, innovation, and organizational behavior? What is uniquely relevant about the intersecting contexts of economic logics, news media, and big-data tools, technologies, and techniques?

Ethics

The term "ethics" derives from the Greek *ethos*, meaning "character" or "personal disposition" on the part of the individual, and relates to the Latin *mores* (or "morals") and its emphasis on the customs of a group. In this way, Ward (2010) notes, ethics is

both internally concerned with the personal decision-making of the individual and externally situated in relation to the rules of society. While the study of ethics encompasses a vast terrain of moral philosophy and theory (e.g., Williams 1973), it suffices for our purposes to conclude that ethics is concerned with appropriate practice within a framework of moral principles.

Ethical behavior and moral reasoning have long been a concern in communication, media, and journalism (Christians et al. 1983). As the foundation for norms and values, ethical standards serve an essential function in orienting journalists, especially, to work in ways that promote honesty, accuracy, transparency, and public service, among other ideals (Singer 2007; Ward 2010). For journalists, ethical codes and conduct serve not only to guide their choices but also to define who they are as professionals. In this latter sense, ethics are associated with professionalism and its power as a form of boundary work, delineating insiders from outsiders and encouraging journalists to distance themselves from people and practices considered inferior or unfamiliar (Carlson and Lewis, forthcoming). The challenge, however, is that journalists are facing all manner of unfamiliarity in the media landscape, not least with regard to technological processes and practices. To reflexively resist or reject such change, framing the unfamiliar as a threat to normative principles, is "to cross the line from using ethics as a legitimate cause for concern to using ethics as a crutch" (Singer 2014, 67), forgoing opportunities to think productively, from the beginning, about how to adapt to and along with innovations as they arise.

Big data, of course, augurs potentially big changes for journalism: from shifting articulations of knowledge (epistemology) to highly technical skill sets and analytical understandings (expertise) to vexing questions of value creation, monetization, and media innovation (economics)—all of which are related to larger transitions in technology and society of which journalism is but one part (Couldry 2012). Such change is neither positively inevitable nor inevitably positive, and yet the specter of it raises questions about how the agents in news media organizations might respond to it all: pragmatically in transforming activities of media production, and philosophically in transforming norms and values that guide behavior. Moreover, the big data phenomenon is freighted with its own set of ethical quandaries—about user privacy, information security, and data manipulation, among others (Crawford et al. 2014)—that deserve scrutiny and reflection as journalists determine how to negotiate innovations associated with it.

Below, we briefly discuss three examples that illustrate the intersecting norms of journalism and big data.

First, consider the process of publishing data or making large data sets publicly available online. An ethos of openness is shared among many data journalists (Howard 2014), though journalists broadly have struggled to embrace such openness as a professional norm (Lewis 2012). In the spirit of open source, data journalists often seek to make complete data sets and programming code open to public examination —creating, in some cases, an invitation for users to contribute and collaborate. Such a movement may help journalism reinvent itself for the digital era, Lewis and Usher (2013) suggest, by integrating norms such as iteration, tinkering, transparency, and participation that are connected with the social, cultural, and technological framework of digital technologies. Yet often there are underlying problems with such public data, whether provided to news organizations by governments or political institutions

(Schudson 2010), or perhaps assembled by news organizations through crowdsourc-ing, data-scraping, or other means. Such problems may go unnoticed either because of the size of the data involved or because of the attractiveness of making it freely available. For example, big data—such as millions of tweets collected around a partic-ular event—are often assumed to represent the social world, yet they have deeply structural "signal problems," with little or no representation coming from less-con-nected communities (Crawford 2013). Meanwhile, journalists must weigh the benefits of open data against the risks of personal harm that may come with publication, par-ticularly at a time when private information can be so easily shared and searched (Howard 2014). Altogether, even well-intentioned efforts to revise journalistic norms through big data run up against ethical questions embedded in the organization, analysis, and dissemination of such data.

Second, consider social science research ethics. Drawing on social science meth-ods is not new for journalists, but the process has accelerated and expanded in this data-rich environment (Howard 2014). Meanwhile, at the same time that journalists are embracing such techniques, "social scientists are undergoing a fundamental shift in the ethical structure that has defined the moral use of these techniques," rethinking what it means to protect individuals from harm and allow for informed consent in a world of big-data research methods involving millions of human subjects (Fairfield and Shtein 2014, 38). Journalists, of course, are not subject to institutional review boards, yet they should be cautious: just because certain content is publicly accessible does not mean that it was intended to be made public to everyone (boyd and Crawford 2012).

Third, consider the ethics of technological systems design, or how computer sys-tems come to embody certain values (Nissenbaum 2001). To the extent that big data implies a large-scale, technology-led turn in journalism (Westlund 2013), what happens as humans embed technological actants like algorithms with some assumptions, norms, and values, and not others? In effect, as machines take on a greater role in news judg-ment, encouraging certain kinds of news selection and consumption over others, how are they "taught" to act ethically? Is there an ethics of algorithms (Kraemer et al. 2011)? Such an ethics will need to unpack various factors of selection, interpretation, and anticipation, revealing "how algorithms structure how we can *see* a concern, why we think it *probably matters*, and *when* we might act on it" (Ananny 2013b, 6, original emphasis), all issues of deep relevance for journalistic decision-making. Altogether, this attention to code and structure behind technology encourages a study of the "black boxes" of big data, working to uncover "the power structures, biases, and influences that computational artifacts exercise in society" (Diakopoulos 2014).

Ultimately, many issues remain to be addressed in future scholarship on ethics at the intersection of big data and journalism. At its core, ethics is concerned with appro-priate practice; researchers thus may examine how social actors in diverse news media firms make sense of, domesticate, and innovate in relation to big-data practices. Importantly, scholars should seek to understand the moral reasoning and ethical decision-making among the full range of media-work agents (Lewis and Westlund 2014), each representing relatively distinct positions with regards to epistemology, expertise, and economics. How are data-centric ethics constructed and rationalized (and also written into technological actants) comparatively across national, sociocultural, and technological contexts? Assuming ethical constructs change over time as the big-data

phenomenon unfolds, what characterizes this change, and what implications does it have for journalism at large?

Conclusion

This article has defined and discussed big data as a social, cultural, and techno-logical phenomenon, in the particular context of journalism. Introducing and applying four conceptual lenses—*epistemology*, *expertise*, *economics*, and *ethics*—the article has systematically explored both contemporary and potential applications of big data for the professional logic and industrial production of journalism. The discussion of these distinct yet inter-related conceptual lenses has shown how journalists and news organizations are seeking to make sense of, act upon, and derive value from big data during a time of exploration in algorithms, computation, and quantification. In all, the developments of big data potentially have great meaning for journalism's ways of *knowing* (epistemology) and *doing* (expertise), as well as its negotiation of *value* (economics) and *values* (ethics). Observers such as Howard (2014) argue that data and associated activities will only matter more in the years ahead, with results that may be utopian, dystopian, or otherwise. To assess such significance and outcomes going forward, scholars may benefit from considering the distinct contributions of epistemology, expertise, economics, and ethics as conceptual footholds, while acknowledging and applying the holistic interplay among them. These approaches, we have argued, are but starting points for undertaking future research on big data and the opportunities and challenges that it poses for journalism, media, and society.

REFERENCES

Abbott, Andrew D. 1988. *The System of Professions: An Essay on the Division of Expert Labor*. Chicago, IL: University of Chicago Press.

Aitamurto, Tanja, and Seth C. Lewis. 2013. "Open Innovation in Digital Journalism: Examining the Impact of Open APIs at Four News Organizations." *New Media & Society* 15 (2): 314–331. doi:10.1177/1461444812450682.

Albarran, Alan B. 2002. *Media Economics: Understanding Markets, Industries and Concepts*. Ames, IA: Blackwell.

Ananny, Mike. 2013a. "Press-Public Collaboration as Infrastructure: Tracing News Organizations and Programming Publics in Application Programming Interfaces." *American Behavioral Scientist* 57 (5): 623–642. doi:10.1177/000276421246936.

Ananny, Mike. 2013b. "Toward an Ethics of Algorithms: Observation, Probability, and Time." Paper presented at the Governing Algorithms Conference, New York, May 16–17.

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

Boellstorff, Tom. 2013. "Making Big Data, in Theory." *First Monday* 18 (10). 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.

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

Carlson, Matt, and Seth C. Lewis, eds. forthcoming. *Boundaries of Journalism: Professionalism, Practices, and Participation*. New York: Routledge.

Carlson Matt. forthcoming. *Journalistic Authority*. New York: Columbia University Press.

Christians, Clifford G., Kim B. Rotzoll, and Mark Fackler, eds. 1983. *Media Ethics: Cases and Moral Reasoning*. New York: Routledge.

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

Coddington, Mark. forthcoming. "The Wall Becomes a Curtain: Revisiting Journalism's News–business Boundary." In *Boundaries of Journalism: Professionalism, Practices, and Participation*, edited by Matt Carlson and Seth C. Lewis. New York: Routledge.

Coleman, E. Gabriella. 2013. *Coding Freedom: The Ethics and Aesthetics of Hacking*. Princeton, NJ: Princeton University Press.

Collins, Henry, and Robert Evans. 2007. *Rethinking Expertise*. Chicago, IL: University of Chicago Press.

Couldry, Nick. 2012. *Media, Society, World: Social Theory and Digital Media Practice*. Cambridge: Polity.

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, 1710–1726. http://ijoc.org/index.php/ijoc/article/view/2166.

Cox, Melissa. 2000. "The Development of Computer-assisted Reporting." Paper presented at AEJMC, Phoenix, August 9-12.

Crawford, Kate. 2013. "The Hidden Biases in Big Data." *Harvard Business Review*, April 1. http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/.

Crawford, Kate, Kate Miltner and Mary L. Gray. 2014. "Critiquing Big Data: Politics, Ethics, Epistemology." *International Journal of Communication* 8, 1663–1672. http://ijoc.org/ index.php/ijoc/article/view/2167/1164.

Deuze, Mark. 2008. "Understanding Journalism as Newswork: How it Changes, and How it Remains the Same." *Westminster Papers in Communication and Culture* 5 (2): 4–23.

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

Dimmick, John W. 2003. *Media Competition and Coexistence: The Theory of the Niche*. Mahwah, NJ: Lawrence Erlbaum.

Domingo, David, Thorsten Quandt, Ari Heinonen, Steve Paulussen, Jane B. Singer, and Marina Vujnovic. 2008. "Participatory Journalism Practices in the Media and Beyond: An International Comparative Study of Initiatives in Online Newspapers." *Journalism Practice* 2 (3): 326–342. doi:10.1080/17512780802281065.

Dwoskin, Elizabeth. 2014. "Beware 'Big Data' Hype, Report Warns." *The Wall Street Journal*, May 27. http://blogs.wsj.com/digits/2014/05/27/beware-big-data-hype-reports-warn/.

Ekström, Mats. 2002. "Epistemologies of TV Journalism: A Theoretical Framework." *Journalism* 3 (3): 259–282. doi:10.1177/146488490200300301.

Epstein, Dmitry, Erik C. Nisbet, and Tarleton Gillespie. 2011. "Who's Responsible for the Digital Divide? Public Perceptions and Policy Implications." *The Information Society* 27 (2): 92–104. doi:10.1080/01972243.2011.548695.

Ettema, James S., and Theodore L. Glasser. 1998. *Custodians of Conscience: Investigative Journalism and Public Virtue*. New York: Columbia University Press.

Eyal, Gil. 2013. "For a Sociology of Expertise: The Social Origins of the Autism Epidemic." *American Journal of Sociology* 118 (4): 863–907. doi:10.1086/66844.

Fairfield, Joshua, and Hannah Shtein. 2014. "Big Data, Big Problems: Emerging Issues in the Ethics of Data Science and Journalism." *Journal of Mass Media Ethics* 29 (1): 38–51. doi:10.1080/08900523.2014.863126.

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:10.1080/17512786.2011.616655.

Gillespie, Tarleton. 2010. "The Politics of 'Platforms'." *New Media & Society* 12 (3): 347–364. doi:10.1177/1461444809342738.

Gitelman, Lisa, ed. 2013. *"Raw Data" is an Oxymoron*. Cambridge, MA: MIT Press.

González-Bailón, Sandra. 2013. "Social Science in the Era of Big Data." *Policy & Internet* 5 (2): 147–160. doi:10.1002/1944-2866.POI32.

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.

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

Howard, Philip N. forthcoming. *Pax Technica: How the Internet of Things May Set us Free or Lock us Up*. New Haven, CT: Yale University Press.

Klinenberg, Eric. 2005. "Convergence: News Production in a Digital Age." *The ANNALS of the American Academy of Political and Social Science* 597 (1): 48–64. doi:10.1177/0002716204270346.

Kraemer, Felicitas, Kees van Overveld, and Martin Peterson. 2011. "Is There an Ethics of Algorithms?" *Ethics and Information Technology* 13 (3): 251–260. doi:10.1007/s10676-010-9233-7.

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., and Nikki Usher. 2013. "Open Source and Journalism: Toward New Frameworks for Imagining News Innovation." *Media, Culture and 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. 2014. "Actors, Actants, Audiences, and Activities in Cross-Media News Work: A Matrix and a Research Agenda." *Digital Journalism*. doi:10.1080/21670811.2014.927986.

Lowrey, Wilson. 2009. "Institutional Roadblocks: Assessing Journalism's Response to Changing Audiences." In *Journalism and Citizenship: New Agendas*, edited by Zizi Papacharissi, 44–67. Mahwah, NJ: Lawrence Erlbaum Associates.

Madrigal, Alexis C. 2014. "Method Journalism." *The Atlantic*, June 10. http://www.theatlantic.com/technology/archive/2014/06/method-journalism/372526/.

Manovich, Lev. 1999. "Database as Symbolic Form." *Convergence: The International Journal of Research Into New Media Technologies* 5 (2): 80–99. doi:10.1177/135485659900500206.

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.

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.

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

Nafus, Dawn, and Jamie Sherman. 2014. "This One Does Not Go Up To 11: The Quantified Self Movement as an Alternative Big Data Practice." *International Journal of Communication* 8, 11. http://ijoc.org/index.php/ijoc/article/view/2170.

Nissenbaum, Helen. 2001. "How Computer Systems Embody Values." *Computer* 34 (3): 120–119.

Parasie, Sylvain. 2014. "Data-Driven Revelation: Epistemological Tensions in Investigative Journalism in the Age of 'Big Data'." *Digital Journalism*. 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.

Pavlik, John V., and Frank Bridges. 2013. "The Emergence of Augmented Reality (AR) as a Storytelling Medium in Journalism." *Journalism & Communication Monographs* 15 (1): 4–59. doi:10.1177/1522637912470819.

Picard, Robert G. 2010. "A Business Perspective on Challenges Facing Journalism." In *The Changing Business of Journalism and its Implications for Democracy*, edited by A. L. David Levy and Rasmus Kleis Nielsen, 17–24. Oxford: Reuters Institute for the Study of Journalism, University of Oxford.

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

Porter, Michael E. 1980. *Competitive Strategy*. New York: Free Press.

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.

Puschmann, Cornelius, and Jean Burgess. 2014. "Metaphors of Big Data." *International Journal of Communication* 8, 20. http://ijoc.org/index.php/ijoc/article/view/2169.

Reich, Zvi. 2012. "Journalism as Bipolar Interactional Expertise." *Communication Theory* 22 (4): 339–358. doi:10.1111/j.1468-2885.2012.01411.

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

Schon, Donald A. 1995. "The New Scholarship Requires a New Epistemology." *Change* 27 (6): 27–34.

Schudson, Michael. 2010. "Political Observatories, Databases & News in the Emerging Ecology of Public Information." *Daedalus* 139 (2): 100–109. doi:10.1162/daed.2010.139.2.100.

Schudson, Michael, and Chris Anderson. 2008. "Objectivity, Professionalism, and Truth Seeking in Journalism." In *Handbook of Journalism Studies*, edited by Karin Wahl-Jorgensen and Thomas Hanitzsch, 88–101. New York: Routledge.

Singer, Jane B. 2007. "Contested Autonomy: Professional and Popular Claims on Journalistic Norms." *Journalism Studies* 8 (1): 79–95. doi:10.1080/14616700601056866.

Singer, Jane B. 2014. "Getting Past the Future: Journalism Ethics, Innovation, and a Call for 'Flexible First'." *Comunicação e Sociedade* 25: 67–82.

Smit, Gerard, Yael de Haan, and Laura Buijs. 2014. "Visualizing News: Make it Work." *Digital Journalism* 2 (3): 344–354. doi:10.1080/21670811.2014.897847.

Smith, Barry. 2001. "Objects and their Environments: From Aristotle to Ecological Ontology." In *The Life and Motion of Socio-Economic Units*, edited by Andrew Frank, Jonathan Raper and Jean-Paul Cheylan, 79–97. London: Taylor and Francis.

Stavelin, Erik. 2014. *Computational Journalism: When Journalism Meets Programming*. Unpublished dissertation. Bergen: University of Bergen.

Storsul, Tanja, and Arne H. Krumsvik. eds. 2013. *Media Innovations*. Gothenburg: Nordicom.

Sunstein, Cass R. 2009. *Republic.com 2.0*. Princeton: Princeton University Press.

Suthaharan, Shan. 2014. "Big Data Classification: Problems and Challenges in Network Intrusion Prediction with Machine Learning." *SIGMETRICS Perform. Eval. Rev.* 41 (4): 70–73. doi:10.1145/2627534.2627557.

Tandoc, Edson C. 2014. "Journalism is Twerking? How Web Analytics is Changing the Process of Gatekeeping." *New Media & Society*. doi:10.1177/1461444814530541.

Thurman, Neil. 2011. "Making 'The Daily Me': Technology, Economics and Habit in the Mainstream Assimilation of Personalized News." *Journalism* 12 (4): 395–415. doi:10.1177/1464884910388228.

Tuchman, Gaye. 1978. *Making News: A Study in the Construction of Reality*. New York: Free Press.

Turow, Joseph. 2011. *The Daily You: How the New Advertising Industry is Defining your Identity and Your Worth*. New Haven, CT: Yale University Press.

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

Ward, Stephen J. A. 2010. *Global Journalism Ethics*. Montreal: McGill-Queen's University Press.

Westlund, Oscar. 2011. *Cross-media News Work: Sensemaking of the Mobile Media (r)evolution*. Gothenburg: University of Gothenburg. https://gupea.ub.gu.se/bitstream/2077/28118/1/gupea_2077_28118_1.pdf.

Westlund, Oscar. 2012. "Producer-centric Versus Participation-centric: on the Shaping of Mobile Media." *Northern Lights: Film and Media Studies Yearbook* 10 (1): 107–121. doi:10.1386/nl.10.1.107_1.

Westlund, Oscar. 2013. "Mobile News: A Review and Model of Journalism in an Age of Mobile Media." *Digital Journalism* 1 (1): 6–26. doi:10.1080/21670811.2012.740273.

Westlund, Oscar, and Seth C. Lewis. 2014. "Agents of Media Innovations: Actors, Actants, and Audiences." *The Journal of Media Innovations* 1 (2): 10–35. doi:10.5617/jmi.v1i2.856.

Wiik, Jenny. 2010. *Journalism in Transition: The Professional Identity of Swedish Journalists*. JMG Book Series no. 64. Gothenburg: University of Gothenburg.

Williams, Bernard. 1973. *Ethics and the Limits of Philosophy*. Hoboken: Taylor & Francis.

Seth C. Lewis (author to whom correspondence should be addressed), School of Journalism and Mass Communication, University of Minnesota–Twin Cities, USA. E-mail: sclewis@umn.edu. Web: www.sethlewis.org

Oscar Westlund, Department of Journalism, Media and Communication, University of Gothenburg, Sweden. E-mail: oscar.westlund@jmg.gu.se. Web: www.oscar westlund.com