Big Data in the Life Sciences: Bibliography, Social and Behavioral Sciences (Part 4)

Author(s)

Valerie Racine

Description

A bibliography looking at the ethics of big data use in social and behavioral sciences.

Body

Policies and Guidance

Dove, Edward S, David Townend, Eric M. Meslin, Martin Bobrow, Katherine Littler, Dianne Nicol, Jantina de Vries, Anne Junker, Chiara Garattini, Jasper Bovenberg, Mahsa Shabani, Emmanuelle Levesqure, Bartha M. Knoppers. 2016. Ethics Review for International Data-Intensive Research. *Science* 351: 6280, March 25. 1399-1400.
The authors reviewed numbers of approaches to ethics review for large data sets relevant to human subjects and their protection (excluding clinical trials research) and identified three models that could inform a framework allowing mutual recognition of international ethics review. The models are reciprocity, delegation, and federation, and a chart listing advantages and disadvantages and examples of

projects for each model is provided.

Metcalf Jacob and Kate Crawford 2016. Where are human subjects in big data research? The emerging ethics divide. Big Data & Society 3(1): 1–14.
Authors illustrate how proposed changes to the regulations governing human-subjects research protections do not address certain harms caused by big data research that uses public datasets, and discuss what ethical protections "data subjects" might require.

The National Academies of Sciences, Engineering and Medicine. 2013. *Proposed Revisions to the Common Rule: Perspectives of Social and Behavioral Scientists: Workshop Summary*. Washington, DC: The National Academies Press.
The summary focuses on: 1. Evidence on the functioning of the Common Rule and of institutional review boards (IRBs). 2. Types and levels of risk and harms in social and behavioral sciences, and issues of severity and probability of harm. 3. Consent and special populations. 4. Protection of research participants. 5. Multidisciplinary and multisite studies. 6. The purview and roles of IRBs.

Books

Collman, Jeff, Sorin Adam Matei (eds.) 2013. *Ethical Reasoning in Big Data: an exploratory analysis.* Cham: Springer.
Looking at the field of computational social science, this books looks at the privacy and ethical implications of research in human affairs using big data.

Elmer, Greg, Ganaele Langlois and Joanna Redden (eds). 2015. *Compromised Data: From Social Media to Big Data.* New York: Bloomsbury.
Discusses how researchers perform critical research within a compromised social data framework influenced by biases, economic interests, etc., and how this can lead to a fundamental shift between research and the public good as well as new forms of control and surveillance.

Kitchin, Rob. 2014. *The Data Revolution: Big data, open data, data infrastructure and their consequences.* London: SAGE Publications.

This book discusses the technical shortcomings and the social, political, and ethical consequences of this 'data revolution', as well as providing an analysis of the potential implications to academic, business, and government practices.

Journal Articles

Ambrose, Meg Leta. 2014. "[Lessons from the Avalanche of Numbers: Big Data in Historical Context](#)." *I/S: A Journal of Law and Policy for the Information Society* 11(2): 201-277.
The big data revolution, like many changes associated with technological advancement, is often compared to the industrial revolution to create a frame of reference for its transformative power, or portrayed as altogether new. This article argues that between the industrial revolution and the digital revolution is a more valuable, yet overlooked period: the probabilistic revolution that began with the avalanche of printed numbers between 1820 and 1840. By comparing the many similarities between big data today and the avalanche of numbers in the 1800s, the article situates big data in the early stages of a prolonged transition to a potentially transformative epistemic revolution, like the probabilistic revolution. The widespread changes in and characteristics of a society flooded by data results in a transitional state that creates unique challenges for policy efforts by disrupting foundational principles relied upon for data protection. The potential of a widespread, lengthy transition also places the law in a pivotal position to shape and guide big data-based inquiry through to whatever epistemic shift may lie ahead.

Ball, Kirstie, MariaLaura Di Domenico, and Daniel Nunan. 2016. "Big Data Surveillance and the Body-subject." *Body & Society* 22 (2):58-81. doi: 10.1177/1357034X15624973.
This paper considers the implications of big data practices for theories about the surveilled subject who, analysed from afar, is still gazed upon, although not directly watched as with previous surveillance systems. The authors propose that this surveilled subject be viewed through a lens of proximity rather than interactivity, to highlight the normative issues arising within digitally mediated relationships. They interpret the ontological proximity between subjects, data flows and big data surveillance through Merleau-Ponty's ideas combined with Levinas' approach to ethical proximity and Coeckelberg's work on proximity in the digital age. This leads us to highlight how competing normativities, and normative dilemmas in these

proximal spaces, manipulate the surveilled subject's embodied practices to lead the embodied individual towards experiencing them in a local sense.

Bonilla, Diego Navarro. 2013. "Information Management professionals working for intelligence organizations: ethics and deontology implications." *Security & Human Rights* **24 (3/4):264-279. doi: 10.1163/18750230-02404005.**

Archive and information management experts trained in library science programs are ideal candidates for jobs in intelligence organizations. Their skills, abilities and knowledge are frequently required in at least two well-defined areas: open source information gathering and records management/archival organisation. Under the general overview of the debate between "big data vs. big narrative" this article focuses on the ethical challenges that affect this community of information professionals. As a key component of the so-called "intelligence culture", it will be also underlined the need for intensifying from our university classrooms the ethical dimension of information exploitation for security and defence purposes. The role played by these information profiles involved in multiple phases of the intelligence production process must be based not only on efficiency and efficacy criteria but also on deontology principles whose benefits are the fortification of democratic practice by intelligence services working in strong legal frameworks designed to guarantee fundamental rights.

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

The era of Big Data has begun. Computer scientists, physicists, economists, mathematicians, political scientists, bio-informaticists, sociologists, and other scholars are clamoring for access to the massive quantities of information produced by and about people, things, and their interactions. Diverse groups argue about the potential benefits and costs of analyzing genetic sequences, social media interactions, health records, phone logs, government records, and other digital traces left by people. Significant questions emerge. Will large-scale search data help us create better tools, services, and public goods? Or will it usher in a new wave of privacy incursions and invasive marketing? Will data analytics help us understand online communities and political movements? Or will it be used to track protesters and suppress speech? Will it transform how we study human communication and culture, or narrow the palette of research options and alter what 'research' means?

Given the rise of Big Data as a socio-technical phenomenon, we argue that it is necessary to critically interrogate its assumptions and biases. In this article, we offer six provocations to spark conversations about the issues of Big Data: a cultural, techno- logical, and scholarly phenomenon that rests on the interplay of technology, analysis, and mythology that provokes extensive utopian and dystopian rhetoric.

Chan, Anita. 2015. "Big data interfaces and the problem of inclusion." ***Media, Culture & Society*** **37 (7):1078-1083. doi: 10.1177/0163443715594106.**

A commentary on 'Critical Questions for Big Data' and the projection in the article of how 'limited access to big data creates new digital divides'. Pressing questions are indeed proliferating around not only what the actual relationship is between data and real world user behavior but also around defining what the very practices, knowledge sets, legal and technological infrastructures, and social norms are that guide the work of big data as a field itself. But how much of a difference does it make for academics and academic institutions to gain access to big data when the logics of commerce and commercial enclosure around data management, collection, and use are what increasingly get privileged?

Cooper, Anwen, and Chris Green. 2016. "Embracing the Complexities of 'Big Data' in Archaeology: the Case of the English Landscape and Identities Project." ***Journal of Archaeological Method & Theory*** **23 (1):271-304. doi: 10.1007/s10816-015-9240-4.**

This paper considers recent attempts within archaeology to create, integrate and interpret digital data on an unprecedented scale-a movement that resonates with the much wider so-called big data phenomenon. Using the example of the authors' work with a particularly large and complex dataset collated for the purpose of the English Landscape and Identities project (EngLaID), Oxford, UK, and drawing on insights from social scientists' studies of information infrastructures much more broadly, they make the following key points. Firstly, alongside scrutinising and homogenising digital records for research purposes, it is vital that we continue to appreciate the broader interpretative value of 'characterful' archaeological data (those that have histories and flaws of various kinds). Secondly, given the intricate and pliable nature of archaeological data and the substantial challenges faced by researchers seeking to create a cyber-infrastructure for archaeology, it is essential that we develop interim measures that allow us to explore the parameters and potentials of working with archaeological evidence on an unprecedented scale.

Crawford. 2014a. "[The Test We Can—and Should—Run on Facebook](#)." The Atlantic. July 2.
Discusses the Facebook emotional contagion experiment and what it means ethically for social researcher interested in doing large-scale user experimentation using social media and big data.

Crawford, Kate. 2014. "[When Big Data Marketing Becomes Stalking](#)." *Scientific American* 310 (4):14-14.
Can data brokers and marketers be trusted to regulate themselves?

Crawford, Kate, Mary L. Gray, and Kate Miltner. "Big Data Critiquing Big Data: Politics, Ethics, Epistemology| Special Section Introduction." International Journal of Communication 8 (2014): 1-11.

Why now? This is the first question we might ask of the big data phenomenon. Why has it gained such remarkable purchase in a range of industries and across academia, at this point in the 21st century? Big data as a term has spread like kudzu in a few short years, ranging across a vast terrain that spans health care, astronomy, policing, city planning, and advertising. From the RNA bacteriophages in our bodies to the Kepler Space Telescope, searching for terrorists or predicting cereal preferences, big data is deployed as the term of art to encompass all the techniques used to analyze data at scale. But why has the concept gained such traction now?

Crawford, Kate and Jacob Metcalf. 2016. "[Where are Human Subjects in Big Data Research? The Emerging Ethics Divide](#)." *Big Data and Society.*
There are growing discontinuities between the research practices of data science and established tools of research ethics regulation. Some of the core commitments of existing research ethics regulations, such as the distinction between research and practice, cannot be cleanly exported from biomedical research to data science research. Such discontinuities have led some data science practitioners and researchers to move toward rejecting ethics regulations outright. These shifts occur at the same time as a proposal for major revisions to the Common Rule—the primary regulation governing human-subjects research in the USA—is under consideration for the first time in decades. The authors contextualize these revisions in long-running complaints about regulation of social science research and argue data science should be understood as continuous with social sciences in this regard. The proposed regulations are more flexible and scalable to the methods of non-

biomedical research, yet problematically largely exclude data science methods from human-subjects regulation, particularly uses of public datasets.

Eastin, Matthew S., Nancy H. Brinson, Alexandra Doorey, and Gary Wilcox. 2016. "Living in a big data world: Predicting mobile commerce activity through privacy concerns." *Computers in Human Behavior* **58:214-220. doi: 10.1016/j.chb.2015.12.050.**
As advertisers increasingly rely on mobile-based data, consumer perceptions regarding the collection and use of such data becomes of great interest to scholars and practitioners. Recent industry data suggests advertisers seeking to leverage personal data offered via mobile devices would be wise to acknowledge and address the privacy concerns held by mobile users. Utilizing the theoretical foundation of communication privacy management (CPM), the current study investigates commonly understood privacy concerns such as collection, control, awareness, unauthorized secondary use, improper access and a newly adapted dimension of location tracking, trust in mobile advertisers, and attitudes toward mobile commerce, to predict mobile commerce engagement. Data from this study indicate that control, unauthorized access, trust in mobile advertisers, and attitude toward mobile commerce significantly predicted 43% of the variance in mobile commerce activity.

Easton-Calabria, Evan, and William L. Allen. 2015. "Developing ethical approaches to data and civil society: from availability to accessibility." *Innovation: The European Journal of Social Sciences* **28 (1):52-62. doi: 10.1080/13511610.2014.985193.**
This research note reflects on the gaps and limitations confronting the development of ethical principles regarding the accessibility of large-scale data for civil society organizations (CSOs). Drawing upon a systematic scoping study on the use of data in the United Kingdom (UK) civil society, it finds that there are twin needs to conceptualize accessibility as more than mere availability of data, as well as examine the use of data among CSOs more generally. In order to deal with the apparent "digital divide" in UK civil society the authors present a working model in which ethical concerns accompanying data utilization by civil society may be better accounted. This suggests there is a need for further research into the nexus of civil society and data upon which interdisciplinary discussion about the ethical dimensions of engagement with data, particularly informed by insight from the social sciences, can be predicated.

Ekbia, Hamid, Michael Mattioli, Inna Kouper, G. Arave, Ali Ghazinejad, Timothy Bowman, Venkata Ratandeep Suri, Andrew Tsou, Scott Weingart, and Cassidy R. Sugimoto. 2015. "Big data, bigger dilemmas: A critical review." *Journal of the Association for Information Science & Technology* **66 (8):1523-1545. doi: 10.1002/asi.23294.**

The recent interest in Big Data has generated a broad range of new academic, corporate, and policy practices along with an evolving debate among its proponents, detractors, and skeptics. While the practices draw on a common set of tools, techniques, and technologies, most contributions to the debate come either from a particular disciplinary perspective or with a focus on a domain-specific issue. A close examination of these contributions reveals a set of common problematics that arise in various guises and in different places. It also demonstrates the need for a critical synthesis of the conceptual and practical dilemmas surrounding Big Data. The purpose of this article is to provide such a synthesis by drawing on relevant writings in the sciences, humanities, policy, and trade literature. In bringing these diverse literatures together, we aim to shed light on the common underlying issues that concern and affect all of these areas. By contextualizing the phenomenon of Big Data within larger socioeconomic developments, we also seek to provide a broader understanding of its drivers, barriers, and challenges. This approach allows us to identify attributes of Big Data that require more attention-autonomy, opacity, generativity, disparity, and futurity-leading to questions and ideas for moving beyond dilemmas.

Flick, Uwe. 2015. "Qualitative Inquiry—2.0 at 20? Developments, Trends, and Challenges for the Politics of Research." *Qualitative Inquiry* **21 (7):599-608. doi: 10.1177/1077800415583296.**

After 20 years of Qualitative Inquiry, some current trends and challenges are outlined, which might affect the current state and further development of qualitative research in the near future. A central focus is their impact on the politics of qualitative research. Politics of inquiry addressing problems of societal relevance are challenged by the globalization and internationalization of qualitative enquiry or trends to big data in funding. Other relevant trends are expectations about archiving and reanalysis of qualitative data, the new interest in qualitative inquiry in the context of evidence, limitations coming from ethical reviews, and the limitation to mixed methods research. These trends are discussed here by using examples from current research projects. Locating qualitative inquiry in the future is discussed between being pushed aside by citizen research and taking over some

(sub)disciplines.

Fuller, Michael. 2015. "Big Data: new science, new challenges, new digital opportunities." *Zygon: Journal of Religion & Science* 50 (3):569-582. doi: 10.1111/zygo.12187.
The advent of extremely large data sets, known as 'big data,' has been heralded as the instantiation of a new science, requiring a new kind of practitioner: the 'data scientist.' This article explores the concept of big data, drawing attention to a number of new issues-not least ethical concerns, and questions surrounding interpretation-which big data sets present. It is observed that the skills required for data scientists are in some respects closer to those traditionally associated with the arts and humanities than to those associated with the natural sciences; and it is urged that big data presents new opportunities for dialogue, especially concerning hermeneutical issues, for theologians and data scientists.

Herther, Nancy K. 2014. "Global Efforts to Redefine Privacy in the Age of Big Data." *Information Today* 31 (6):1-36.
The article reports on efforts in specifying privacy in the big data era worldwide. It mentions retail firm Target Corp. and its creation of algorithms for determining pregnant teenagers. An overview of the Electronic Frontier Foundation's (EFF) rating system for the user privacy protection capability of social media sites and Internet search engines is also presented.

Holtzhausen, Derina. 2016. "Datafication: threat or opportunity for communication in the public sphere?" *Journal of Communication Management* 20 (1):21-36. doi: 10.1108/JCOM-12-2014-0082.
The paper also exposes the potential for harm in the use of Big Data, as well as its potential for improving society and bringing about social justice. Originality/value – The value of this paper is that it introduces the concept of datafication to communication studies and proposes theoretical foundations for the study of Big Data in the context of strategic communications. It provides a theoretical and social foundation for the inclusion of the public sphere in a definition of strategic communication and emphasizes strategic communicators' commitment to the public sphere as more important than ever before. It highlights how communication practice and society can impact ach other positively and negatively and that Big Data should not be the future of strategic communication but only a part of it.

Honda, Laurie. 2017. "Case Study: "It Was A Matter of Life and Death": A YouTube Engineer's Decision to Alter Data in the 'It Gets Better Project'." Council for Big Data, Ethics, and Society.

In this case study, a YouTube engineer contemplates whether to subvert engineering best practices to bypass storage capacity limits on videos created for the It Gets Better Project, which aims to prevent self-harm by LGBTQ youth.

Horvitz, Eric and Deirdre Mulligan. 2015. "Data, privacy, and the greater good." *Science, Policy Forum*, 17 July 2015. 349 (6245): 253-255.

Large-scale aggregate analyses of anonymized data can yield valuable results and insights that address public health challenges and provide new avenues for scientific discovery. These methods can extend our knowledge and provide new tools for enhancing health and wellbeing. However, they raise questions about how to best address potential threats to privacy while reaping benefits for individuals and to society as a whole. The use of machine learning to make leaps across informational and social contexts to infer health conditions and risks from nonmedical data provides representative scenarios for reflections on directions with balancing innovation and regulation.

Jaeger, Jaclyn. 2016. "Think the FTC isn't monitoring big data? Think again." *Compliance Week* 13 (146):26-27.

The Federal Trade Commission of the US released a report in January 2016 warning companies about the sort of ethical, legal and compliance risks they could face when using data analytics practices that are counter to consumer protection and equal opportunity law. The report also poses a series of questions to companies to consider when using big data to mitigate these risks.

Johnson, Jeffrey A. 2014. "From open data to information justice. *Ethics and Information Technology.* 16 (4):263-274.

This paper argues for subsuming the question of open data within a larger question of information justice, with the immediate aim being to establish the need for rather than the principles of such a theory. The author shows that there are several problems of justice that emerge as a consequence of opening data to full public accessibility, and are generally a consequence of the failure of the open data movement to understand the constructed nature of data. The author examines the problems of the embedding of social privilege in datasets as the data is constructed, the differential capabilities of data users (especially differences between citizens and

''enterprise'' users), and the norms that data systems impose through their function as disciplinary systems. In each cases he shows that open data has the quite real potential to exacerbate rather than alleviate injustices.

Lazer, David. The rise of the social algorithm. *Science* **348 (6239):1090-1091.doi: 10.1126/science.aab1422**
Humanity is in the early stages of the rise of social algorithms: programs that size us up, evaluate what we want, and provide a customized experience. This quiet but epic paradigm shift is fraught with social and policy implications. The evolution of Google exemplifies this shift. It began as a simple deterministic ranking system based on the linkage structure among Web sites—the model of algorithmic Fordism, where any color was fine as long as it was black (1). The current Google is a very different product, personalizing results (2) on the basis of information about past searches and other contextual information, like location. On page 1130 of this issue, Bakshy et al. (3) explore whether such personalized curation on Facebook prevents users from accessing posts presenting conflicting political views.

Kernaghan, Kenneth. 2014. "Digital dilemmas: Values, ethics and information technology." *Canadian Public Administration* **57 (2):295-317. doi: 10.1111/capa.12069.**
In writings on public administration, the subject areas of values and ethics and of information technology (IT) have received substantial, but largely separate, attention. The public administration community can benefit by drawing on scholarship in the field of information and computer ethics and developing its own body of research with a view to sensitizing public servants to the effects of changes in IT on values and ethics. This article focuses on developments in the use of IT (for example, self-service technologies, Big Data, the Internet of Things) as a basis for assessing their implications for public sector values and ethics.

King, Gary. 2011. "Ensuring the Data-Rich Future of the Social Sciences." *Science* **331 (6018):719-721. doi: 10.1126/science.1197872.**
Massive increases in the availability of informative social science data are making dramatic progress possible in analyzing, understanding, and addressing many major societal problems. Yet the same forces pose severe challenges to the scientific infrastructure supporting data sharing, data management, informatics, statistical methodology, and research ethics and policy, and these are collectively holding back progress. The author addresses these changes and challenges and suggest what can be done.

Kosinski, Michal, Sandra C. Matz, Samuel D. Gosling, Vesselin Popov, and David Stillwell. 2015. "Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines." *American Psychologist* 70 (6):543-556. doi: 10.1037/a0039210.

Facebook is rapidly gaining recognition as a powerful research tool for the social sciences. It constitutes a large and diverse pool of participants, who can be selectively recruited for both online and offline studies. Additionally, it facilitates data collection by storing detailed records of its users' demographic profiles, social interactions, and behaviors. With participants' consent, these data can be recorded retrospectively in a convenient, accurate, and inexpensive way. Based on our experience in designing, implementing, and maintaining multiple Facebook-based psychological studies that attracted over 10 million participants, we demonstrate how to recruit participants using Facebook, incentivize them effectively, and maximize their engagement. We also outline the most important opportunities and challenges associated with using Facebook for research, provide several practical guidelines on how to successfully implement studies on Facebook, and finally, discuss ethical considerations.

Macer, Tim. 2016. "Analytics, Ethics and Market Research." *Research World* 2016 (56):41-43. doi: 10.1002/rwm3.20326.

Big data can mean big risks if you don't have a sound legal and ethical basis for integrating the data into your research. This is the view of three individuals I spoke to, whose job it is to consider the legal and ethical use of big data, and how to reconcile the rights of private individuals over the use of their data with the opportunities for economic and social benefits it can bring.

Martin, Kirsten. 2016. "Data aggregators, consumer data, and responsibility online: Who is tracking consumers online and should they stop?" *Information Society* 32 (1):51-63. doi: 10.1080/01972243.2015.1107166.

The goal of this article is to examine the strategic choices of firms collecting consumer data online and to identify the roles and obligations of the actors within the current network of online tracking.

Martin, Kirsten E. 2015. "Ethical Issues in the Big Data Industry." *MIS Quarterly Executive* 14 (2):67-85.

Big Data combines information from diverse sources to create knowledge, make

better predictions and tailor services. This article analyzes Big Data as an industry, not a technology, and identifies the ethical issues it faces. These issues arise from reselling consumers' data to the secondary market for Big Data. Remedies for the issues are proposed, with the goal of fostering a sustainable Big Data Industry.

McNeely, Connie L., and Jong-on Hahm. 2014. "The Big (Data) Bang: Policy, Prospects, and Challenges." *Review of Policy Research* **31 (4):304-310. doi: 10.1111/ropr.12082.**

Big data is increasingly the cornerstone on which policy making is based. However, with potential benefits and applications come challenges and dilemmas. In this set of symposium articles, authors examine the promise and problems of big data, exploring associated prospects, risks, parameters, and payoffs from a variety of perspectives. The articles address myriad challenges in the handling of big data sets, such as collection, validation, integrity, and security; ontological issues attending data analytics and conceptual transformations; the foundations of big data collection for social science research; the gap between the acquisition of data and its use to advance discovery and innovation; the costs and benefits of using big data in decision making and analysis; and, finally, related problems of privacy, security, and ethics.

Contributor(s)

Valerie Racine
Karin Ellison
Joseph Herkert

Rights

Use of Materials on the OEC
License
CC BY-NC-SA

Resource Type

Bibliography

Parent Collection

Big Data in the Life Sciences Collection

Topics

Big Data

Controversies

Discipline(s)

Social and Behavioral Sciences