

Ganaele Langlois, Joanna Redden, and Greg Elmer (Eds.), **Compromised Data: From Social Media to Big Data**, New York, NY: Bloomsbury Academic, 2015, 296 pp., Bloomsbury: \$17.54 (PDF); Amazon: \$15.61 (paperback), \$19.36 (Kindle).

Reviewed by

Michael L. Kent

University of New South Wales (UNSW), Sydney Australia

As with many books, the title of this book previews the premise of the text. **Compromised Data: From Social Media to Big Data** is an examination of the integrity of online data and how it is used and abused by organizational and governmental agents. Edited by three Canadian scholars, with chapters written by scholars from Australia, Canada, Hong Kong, Singapore, UK, and the United States, *Compromised Data* is a truly international text examining issues that exist in almost every country and most parts of the world. The scope of the text is both macro (examining big data systems, social media, and political economy), and micro (utilizing case studies and examining cultural movements).



Compromised Data offers a self-referential set of essays that often make connections to the other chapters and ideas found in the volume and provide a surprisingly coherent set of ideas and assumptions for an edited volume. *Compromised Data* also has the feel of another cutting edge text of its day, *The Political Economy of Information* (1988), from two decades ago by Vincent Mosco (also a Canadian) and Janet Wasco (a U.S. scholar), who examined existing media, political economy, information as a social good, culture, and other concepts. The authors of the various chapters in *Compromised Data* explore issues that most technology professionals are probably already aware of and recognize, but have largely come to take for granted. The prevalence of the numerical mind-set and the power of big data to control our lives and the mind-set that is created by data, numbers, quantification, hierarchy, etc. is something that has come to dominate most people's lives (cf., Postman, 1993).

Indeed, even now as we read texts critical of the media industries and information technologies, we seem to be revisiting arguments about technological risks and exploitation that go back decades. Scholars such as Burnam (1984) warned about the dangers of databases of information being exploited by government agents. Negroponte (1995) saw the eventual shape that networked technologies would take more than a decade ahead of their emergence. Postman (1993) recognized the power that quantification and technology has to usurp control over our lives. Stoll (1995) understood the seductive power of the technologies that have emerged through social media. And Vallee (1982) recognized the power that databases of information would bring. *Compromised Data* is an updated version of many of the arguments from those texts.

The early warnings that were written prior to and at the birth of the Internet were prescient, much like the essays in *Compromised Data*. Indeed, as Vallee (1982) remarked, and as several of the authors of *Compromised Data* would probably agree, "Real power resides with those who set up the

structure for others to think about because they define what is available and what is not, what is recorded and what is forgotten" (p. 87). A theme taken up in chapter 10 by Renzi and Langlois, and by David Karpf in chapter 4, who writes about the manipulation of political petitions and using various new technologies for political purposes.

Compromised Data is organized into three sections—Data, Power, and Politics; Data Limit(ed); and Alt-Data. Langlois, Redden, and Elmer take the reader through 12 chapters, critiquing information technologies, social media, and big data. Although the text was published in 2015, it deals with many issues that have become notorious over the last 12 months, including compromised personal information; "big-data" being exploited for governmental, private, and corporate interests; and the many risks and harms that come from the exploitation of unchecked personal data. *Compromised Data* would make an excellent class text for an ethics, technology, public relations, or other course. A brief review of each chapter should shed some light on the content of the text.

The first chapter and introduction to *Compromised Data*, by the text's authors, outline the key issues to be taken up in the text, such as the inability for most people to access their personal information and social data, controlled by corporations and the government; the privileging, aggregation, marketization, management, and control of our personal information; as well as the concomitant "limiting opportunities for democratic engagement, participation, and contestation" (p. 1). The many compelling topics previewed by Langlois et al. are then taken up in the 11 chapters that follow.

Joanna Redden, in chapter 2, "Big Data as System of Knowledge: Investigating Canadian Governance," shares a very practical case study of governmental use of "big data" (and personal data), reporting on the results of primary data and personal interviews to explore the Canadian government's "volume, velocity, and variety" of data (p. 21) that is currently being gathered about citizens. As Redden notes, "There is a danger that an emphasis on data mining, data analysis, and data-informed decision-making will transform the nature of government" (p. 26). The data is increasingly controlled by the haves, who can understand and manipulate the data, and the data have-nots, or *us*, everyday citizens, who lack both access to the data, and the general ability to understand how it is being manipulated.

Chapter 3 offers another practical case study, by Ingrid Hoofd: "Data Mining Research and the Disintegration of Society: The 'Project X' Haren Riots." Hoofd examines "an unexpected and unprecedented spate of rioting" (p. 40) in the Dutch city of Haren, arguing for vigilance and apprehensiveness of the discourses around new technology and the "uncritical eagerness" with which technology has been embraced (p. 41).

Chapter 4, by David Karpf, offers a "Look at the Man Behind the Curtain: Computational Management in 'Spontaneous' Citizen Political Campaigning." Karpf examines three online activism sites—Change.org, MoveOn.org, and We the People (Petitions.Whitehouse.gov)—and identifies some interesting behind-the-scenes power dynamics regarding how petitions are enacted and citizen political campaigning works. Karpf provides a thorough analysis and explanation of the online petition process and traces the process from inception of an idea through success, as well as exploring the obstacles and issues involved in the process.

The second section begins with chapter 5, "Easy Data, Hard Data: The Politics and Pragmatics of Twitter Research After the Computational Turn," by Jean Burgess and Axel Bruns. They describe the "conceptual turn" that has taken place in the last few years toward an analysis of big data and the way that big data platforms like Twitter and Facebook shape and control access and how we study problems (cf. pp. 93–95). Burgess and Bruns discuss "platform politics and regimes of access" (p. 96), and the "changing scientific affordances" (p. 100) of social media data. In particular, their discussion of "easy" vs. "hard" data and how having access to the hard data is more difficult but provides infinitely more insight is very compelling. Burgess and Bruns argue that "there is a growing divide between the majority of researchers who are forced to work with easy data—pursuing the low-hanging fruit . . . and a minority of researchers . . . who . . . have the funds to pay" (p. 107) for better, harder, more useful data.

Greg Elmer, coeditor of the text, authored chapter 6, "Scraping the First Person." Elmer discusses the process of "data scraping" (using electronic tools to mine data sites for information), its influence over "business models and brands" (p. 114), and the ability of data scraping to evolve into an analytic method. For Elmer, a key concern with the emergence of data scraping technologies is that they fail to recognize the "first person" or individualized nature of the online, social media experience. Elmer focuses on several types of data scraping, including screen and interface scraping, crawler scraping, and API scraping, ultimately pointing to the problems and opportunities inherent in the machine logic for researchers interested in conducting social research.

Fenwick McKelvey, in chapter 7, "Openness Compromised? Questioning the Role of Openness in Digital Methods and Contemporary Critical Praxis," examines "the implications and consequences of openness-inspired digital methods" (p. 127). Using discussions of "web archiving and platform studies" (p. 128), McKelvey ultimately concludes that researchers may need to reorient themselves toward a "more experimental and less replicable mode of inquiry and engagement" (p. 128). Questions about the balance between safety and security of private information as well as allowing open information to remain open while at the same time preventing open data from being used exploitatively are examined.

Chapter 8, the last chapter in Part 2, is by Robert Gehl, "Critical Reverse Engineering: The Case of Twitter and TalkOpen." Gehl focuses on the concept of "'reverse engineering' to critique contemporary social media software as well as establish criteria for alternatives to that software" (p. 148). Gehl also uses Twitter's TalkOpen as a case study to help describe how critics might usefully approach the study of social media. Drawing on the literature of reverse engineering and his examination of TalkOpen, as well as insights from his book, *Reverse Engineering Social Media* (2014), Gehl provides several suggestions, including designing new social media for social justice purposes, designing social media platforms not based in surveillance, an expansion or focus on genuine social media, and a number of other interesting issues.

Part 3, "Alt-Data," begins with Sky Croeser and Tim Highfield's chapter, "Mapping Movements—Social Movement Research and Big Data: Critiques and Alternatives." Croeser and Highfield begin by reviewing some of the major issues that have been raised by others in the text about the way that social media have changed social research. Croeser and Highfield argue for a mixed-methods approach, combining participant observation, in-depth interviews, and big data methods. The authors use a case

study of the Occupy Oakland movement to help explain their “ethical” approach and some of the ethical and methodological issues at stake, such as the risks some people face from using big data analyses to highlight movement, members and activities, biases, etc.

Chapter 10 has Alessandra Renzi and Ganaele Langlois describing “Data Activism.” Renzi and Langlois begin with the “commonplace observation: whoever owns, controls, and has the right to access and analyze data holds tremendous power over individuals and populations” (p. 202). The power of data, however, is not just in possessing the data to produce knowledge, but also being able to “shape perceived realities” (p. 202). The people who control data wield political power. Renzi and Langlois review three vignettes of social media activism via metadata, big data, and what they refer to as “facial weaponization,” which explores how various technologies have been used to “unmask,” “make visible,” and “conceal” (p. 230), as well as focusing on various Occupy streams to explore their concepts and examples.

Chapter 11 is by Yuk Hui, “A Contribution to the Political Economy of Personal Archives.” Hui discusses a subtle but important issue that has become part of the digital age: “personal archiving.” Hui begins with a brief review of what archiving is—“preserving, selecting, and forgetting” (p. 227)—drawing attention to the nature of digital archiving, where so much information and data exists, but at the same time is lost and invisible to history. Hui argues that “the archive needs new rules and new archivist practices” (p. 227). In the body of the chapter, Hui addresses two questions: “What happened to the concept of the archive after digital industrialization?” and “What kind of archivist is necessary and desirable in the new configuration?” (p. 228). The chapter also discusses intriguing questions about archival care and practices, decentralization, archival culture, etc.

The final chapter explored in Part 3, chapter 12, by Lisa Blackman, “The Haunted Life of Data,” provides a final comment on the idea of alt-data. Blackman’s focus is a case study of “little data” and the serendipitous relationships and discoveries that can only come from examining messages rather than masses of information, what she calls the “micro-dynamics of data generation” (p. 248). Blackman describes a “hauntological” mode of inquiry that focuses on the “trail or traces” of events, via the “John Bargh priming controversy” (p. 249). She examines how various interpretative methodologies and messages lead to unique insights and conclusions that cannot even be imagined using data analytics and quantitative methodologies. Blackman’s essay represents what she refers to as a “political project” (p. 251), where she hopes to revive something of the archival and critical methodology that has been lost with the focus on big data and data analytics.

On balance, *Compromised Data* is an excellent text. The ideas and issues discussed are sure to be of interest to anyone interested in big data critique. More importantly, however, the ideas in the text represent extensions of decades-old issues that communication scholars and professionals need to be aware of, understand, and come to terms with.

References

- Burnham, D. (1984). *The rise of the computer state*. New York, NY: Vintage Books.
- Gehl, R. (2014). *Reverse engineering social media*. Philadelphia, PA: Temple University Press.
- Mosco, V., & Wasco, J. (Eds.). *The political economy of information*. Madison, WI: University of Wisconsin Press.
- Negroponte, N. (1995). *Being digital*. New York, NY: Alfred A. Knopf.
- Postman, N. (1993). *Technopoly: The surrender of culture to technology*. New York, NY: Vintage Books.
- Stoll, C. (1995). *Silicon snake oil: Second thoughts on the information highway*. New York, NY: Doubleday.
- Vallee, J. (1982). *The network revolution: Confessions of a computer scientist*. Berkeley CA: And/Or Press.