

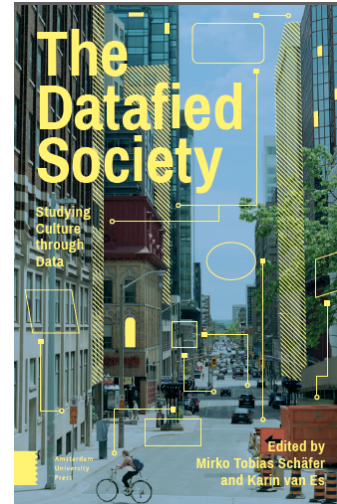
Mirko Tobias Schäfer and Karin van Es (Eds.), **The Datafied Society: Studying Culture through Data**, Amsterdam, The Netherlands: Amsterdam University Press, 2017, 266 pp., \$89.10 (hardcover), \$31.50 (paperback).

Reviewed by
Ester Appelgren
Södertörn University, Sweden

Will the current practice of data and digital analysis harm humanities scholarship? Or will increasing access to data and easy-to-use tools for digital analysis instead strengthen the ability to critically interpret culture and contemporary life?

These two perspectives permeate **The Datafied Society: Studying Culture through Data**, edited by Mirko Tobias Schäfer and Karin van Es. The contributors include distinguished as well as more junior scholars, predominately from Central Europe with an emphasis on the Netherlands and Germany, although North America and the UK are represented as well. Because the authors come from different subdisciplines of media and communication studies, the book contains a broad range of perspectives and examples as well as contrasting purposes for why researchers engage in studies of the increasingly datafied society. Ranging from cultural critique to descriptive studies of production, *The Datafied Society* is an important contribution to literature focused on data analysis for research purposes. The book serves as an introduction for humanistic research that is aided by data analysis, and I believe the intended audience includes humanities scholars and students who are just starting to use data in their research. The editors state that it is common for humanities scholars to take part in transdisciplinary project groups and that these scholars often have to “negotiate the terms of their cooperation and, in particular, find ways to reconcile their disparate epistemic positions” (p. 31). Not being a humanities scholar myself, I recognize such negotiations when working on transdisciplinary projects, and each chapter of this book offered me a piece to the puzzle and a greater understanding of the reasons certain scholars approach data and digital research methods with such caution. Therefore, I believe the book may also be useful to scholars outside of the humanities.

The backdrop to the book is not, as one may think, primarily the increasing abundance of data in society, often referred to as “big data.” Nor is it intended to be a sequel to best sellers such as Viktor Mayer-Schönberger and Kenneth Cukier’s *Big Data* (2013). In fact, Es and Schäfer prefer to acknowledge the development as a “computational turn,” and their motivation for writing the book originated in 2013 at the Utrecht Data School, where students and researchers do explorative data analysis in cooperation with municipalities, creative industries, and academic partners (Datafied Society, 2017). Against this backdrop, I believe the book is intended to inspire those who want to use digital methods in their research. The truly inspiring sections that really make you want to “get your hands dirty” are Manovich’s two chapters on the image analysis toolkit, Cultural Analytics, and the concluding interview with Mercedes Bunz. However, many of the essays, even though they are intended to inspire, focus on justifying the use of data in the humanities, addressing and thus highlighting negative critiques rather than using the space to primarily



encourage researchers with the great examples that are available. As part of the lengthy debate around the digital humanities, critiques needs to be addressed, of course, and Masson, for example, states the inevitable: "With the introduction of digital research tools, and tools for data research specifically, humanistic scholarship seems to get increasingly indebted to positivist traditions" (p. 25). Gosvig Olesen notes in an account of the uses of the Cinematrix tool in analysis of stylistic structure in films that such a tool "would make humanities scholars uncomfortable" (p. 47). Perhaps this reluctance is important and a natural part of the development process. In this context, Uricchio suggests that reluctance has to do with "a new order of things" (p. 126) when technology that amplifies the individual agency is introduced, giving as an example resistance to the introduction of the printing press and the debate around three-point perspective. The nineteen chapters are motivational but also something of a deterrent for those who want to do scholarly work within the humanities, as everything is described with caution, perhaps overemphasizing the vulnerabilities of using data in humanistic research.

The book is arranged into four sections. The first section, "Studying Culture through Data," presents the tensions as well as the opportunities data analysis provides in the humanities. The authors position humanistic data research against other research domains with an emphasis on explaining current challenges. The second section, "Data Practices in Digital Data Analysis," focuses on algorithms, black boxes, mathematics, and statistics. The central message of this section is that the learning curve for researchers wanting to approach data as a research object is quite steep, but there is a need to learn more. The idea is not to use programming and mathematical models to build something, but to use such skills to understand how programs and mathematical models work in order to be able to investigate creations that come out of them and thus be able to assess their impact on society. The third section of the book provides guidelines for data ethics by discussing ethical factors that researchers must take into consideration during all stages of the research process. Because data analytics tools are often created for commercial rather than research purposes, researchers must be particularly cautious when turning to these tools. There is, as of yet, no established one-size-fits-all approach for Internet research, and Markham and Buchanan (p. 203) encourage researchers to focus on a person's relationship to their information in each context when making ethical considerations. Furthermore, as stated by Schie, Westra, and Schäfer (p. 191), it is often the case that neither the tools nor the collection of data is illegal when it is openly collected from sources where individuals may have shared personal data, but that does not mean that the process is ethical. What is not mentioned in the book, however, is that legislation such as the General Data Protection Regulation (European Parliament, 2016) will soon force companies and authorities to structure their internal processes in such a way as to achieve a more sustainable individual privacy, and digital integrity management is rapidly changing corporate data ethics as well. The third section of this book is an important aid to all individuals who engage in digital analysis for research purposes.

The book concludes with a section based on interviews. Four scholars are interviewed about their critical standpoints on aspects of the datafied society. The section gives the reader an overview of other scholarly works related to the datafication of society and also implicitly positions the book in relation to previous literature. For example, in the interview with Nick Couldry, he describes the necessity to resist the "myth of big data" in order to produce knowledge about "the social." Carolin Gerlitz talks about data point critique, where she argues that researchers need to understand what they are counting when using

or aggregating metrics. Evgeny Morozov offers a nontechnological determinist view on algorithms and big data, and Mercedes Bunz wonders about what we are hoping to gain socially from digital technology. She believes that people need to be more curious about what algorithms can do and that the only way for anyone to become more curious is to become more digitally literate.

Throughout my reading of the book, I wondered, without really finding an answer, whether digital research methods based on various types of data, in particular behavioral data, are indeed currently introducing novel ways of conducting research or if new possibilities to explore behavior are normalized and adjusted to fit into existing research traditions, practices, and conventions. (See, for example, Singer, 2005; Lasorsa, Lewis, & Holton, 2012; and Lewis & Usher, 2016, who have observed similar developments in the field of journalism). Schäfer and Es argue that, for the humanities, adjusting digital research methods is somewhat preferable because “rather than importing questions and methods from the hard sciences, humanities scholars should develop their own approaches and sensitivities in working with data that will reflect the humanities’ traditions” (p. 15). This is more or less where we arrive at the end of the nineteen chapters of *The Datafied Society*, where Mercedes Bunz concludes by stating that algorithms are a cultural technique and those who have always been experts in human culture are the humanities scholars. As experts in human knowledge, ever since the invention of first writing and then books, they must now also develop their own dialogue with data.

References

- Datafied Society. (2017). *Utrecht Data School*. Retrieved from <http://datafiedsociety.wp.hum.uu.nl/research-themes/>
- European Parliament. (2016). *The EU General Data Protection Regulation*. Retrieved from <http://eur-lex.europa.eu/eli/reg/2016/679/oj>
- Lasorsa, D. L., Lewis, S.C., & Holton, A. E. (2012). Normalizing Twitter: Journalism practice in an emerging communication space. *Journalism Studies*, 13(1), 19–36.
- Lewis, S., & Usher, N. (2016). Trading zones, boundary objects, and the pursuit of news innovation: A case study of journalists and programmers. *Convergence*, 22(5), 543–560.
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. London, UK: John Murray.
- Singer J. B. (2005). The political j-blogger: “Normalizing” a new media form to fit old norms and practices. *Journalism*, 6(2), 173–198.