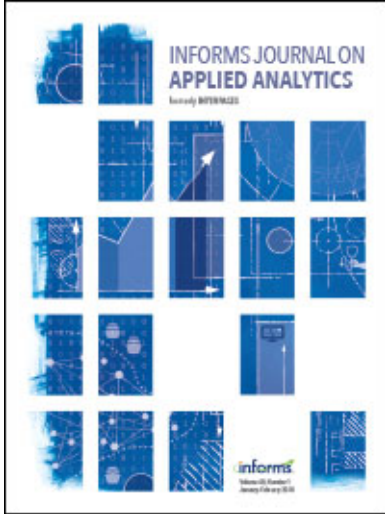




Interfaces

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Book Reviews

Wenjing Shen

To cite this article:

Wenjing Shen (2017) Book Reviews. Interfaces 47(3):273-277. <https://doi.org/10.1287/inte.2017.0895>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2017, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Book Reviews

Wenjing Shen^a

^a Department of Decision Sciences and MIS, LeBow College of Business, Drexel University, Philadelphia, Pennsylvania 19104

Contact: wenjings@drexel.edu (WS)

<https://doi.org/10.1287/inte.2017.0895>

Copyright: © 2017 INFORMS

Abstract. In *Book Reviews*, we review an extensive and diverse range of books. They cover theory and applications in operations research, statistics, management science, econometrics, mathematics, computers, and information systems. In addition, we include books in other fields that emphasize technical applications. The editor will be pleased to receive an email from those willing to review a book, with an indication of specific areas of interest. If you are aware of a specific book that you would like to review, or that you think should be reviewed, please contact the editor. This issue of *Interfaces*, 47(3), May–June 2017, includes two book reviews. The first, Themed Book Review: Big Data and Social Media Analytics, comprises three books: *Big Data at Work: Dispelling the Myths, Uncovering the Opportunities*, Thomas H. Davenport; *Human-Centered Social Media Analytics*, Yun Fu, ed.; and *Big Data Is Not a Monolith*, Cassidy R. Sugimoto, Hamid R. Ekbia, and Michael Mattioli, eds. The second reviews *Sustainability for a Warming Planet*, Humberto Llavador, John E. Roemer, and Joaquim Silvestre.

Themed Book Review: Big Data and Social Media Analytics

DAVENPORT, THOMAS H. 2014. *Big Data at Work: Dispelling the Myths, Uncovering the Opportunities*.

Harvard Business Publishing. 228 pp. \$35.00.

FU, YUN, ED. 2014. *Human-Centered Social Media Analytics*. Springer, 208 pp. \$109.00.

SUGIMOTO, CASSIDY R., HAMID R. EKBIA, MICHAEL MATTIOLI, EDS. 2016. *Big Data Is Not a Monolith*. MIT Press. 312 pp. \$60.00

Each day, we generate 2.5 quintillion bytes of data, 90 percent of which we have created in the past two years (IBM 2017). These big data are generally characterized by volume, variety, and velocity. Nevertheless, scholars are not awed by big data, but are asking important questions, such as: What new knowledge are we producing from big data? How can these data help to solve the wicked problems of our times? Can big data and social media analytics researchers ask important questions and develop novel solutions? This themed book review is based on three books on big data and social media analytics. The first book, *Big Data at Work: Dispelling the Myths, Uncovering the Opportunities*, addresses how organizations should employ and manage big data and the ensuing challenges. The

second book, *Human-Centered Social Media Analytics*, starts with a discussion of text analytics and then provides insightful chapters on image and video analytics. The third book, *Big Data Is Not a Monolith*, investigates big data problems and practices in various contexts that range from science to commerce. My objective in this book review is to introduce readers to emerging research questions and theoretical and methodological challenges. Because the three books complement each other, researchers should gain a broad understanding of the emerging big data and social media analytics research.

Big Data at Work: Dispelling the Myths, Uncovering the Opportunities

Big Data at Work: Dispelling the Myths, Uncovering the Opportunities explains big data as an exponential increase in data, including substantial amounts of unstructured data that are moving at a fast pace. Big data comes from numerous sources, such as posts on social media, digital pictures and videos, and the Internet of things (IoT), employed to collect environmental information, and smartphone applications. This book covers big data from an organizational perspective, including technical, consumer, and

management aspects. For the uninitiated, it provides a good overview of why and how organizations are employing big data, and explains the impact on both people and technologies.

In Chapter 2, the author explores the impact of big data on industries and organizations. The scenarios presented illustrate major topics that various organizations are exploring. Scholars who are new to big data research will benefit from understanding how big data are transforming firms. For example, this chapter describes a large plumbing-fixtures manufacturing company's use of big data to manage energy. This company employs an "analytics blanket" to manage all functions with the organization from its fleet of vehicles to its buildings. The chapter also includes a discussion of Google's innovation in self-driving cars, a big data project with substantial potential impact on the auto industry.

Big data technologies within organizations are the focus of Chapter 5. Big data cannot be handled with traditional database software because of the data's format, variety, and volume. As scholars study big data research questions, they will inevitably have to understand the technological context, especially the layers of technology called the big data stack that provide specialized, high-performance processing and storage. The data sources are both internal, such as operational systems, and external, such as social media data. The data variety includes structured, semistructured, and unstructured data.

Chapter 6 presents the updated data, enterprise, leadership, targets, and analysts (DELTA) framework, as applicable to big data and analytics. Chapters 7 and 8 illustrate how start-up, online, and large firms are employing big data technologies and the ensuing challenges. The author employs short cases that illustrate the role of big data in organizations. For example, LinkedIn's People You May Know predicts and presents recommendations for possible new connections based on a multifactor approach that includes shared schools, workplaces, connections, and geographies. The book takes a prescriptive approach and includes numerous examples of firms employing big data. Unfortunately, this approach is also its limitation because it does not delve deeply into the role of big data in society or industry.

Human-Centered Social Media Analytics

Social media analytics constitute an important part of big data scholarship. Presently, text-based analytics are the primary method employed to study social media data in management and social sciences. *Human-Centered Social Media Analytics* focuses on the novel social computational methodologies that are being developed to investigate social media data. Scholars working in this research domain have contributed to this book, which covers techniques that range from recognizing the social roles that people play in an event to the prediction and recognition of human attributes, such as social relationships. It comprises two parts: *Social Relationships in Human-Centered Media* (Part 1) and *Human Attributes in Social Media Analytics* (Part 2).

Chapter 1 augments traditional machine-learning techniques with socially aware algorithms. The authors have tweaked the latent Dirichlet allocation (LDA) model, a topic-modeling approach, to learn topics in real time from a social media stream; they refer to the topic model as online streaming LDA (OSLDA). This technique can be used to extract, learn, populate, update, and curate the topic space in real time with streaming tweets. As such, it might be helpful to scholars who are conducting complex text analytics.

The remainder of the chapters focus primarily on image and video analysis in a social context. Chapter 4 illustrates how social relationships can be analyzed in an event-based video, which was captured using pairwise interaction features with person-specific social descriptors. Chapter 5 proposes a random forest with a discriminative decision-tree algorithm to explore a dense sampling space for fine-grained image categorization focused on classifying human-object interaction activities. Chapter 6 takes people within a personal image collection and incorporates pairwise social relationships, such as mother and child, to represent the relationship. This model can potentially be adapted to examine the relationships between different object categories for other vision tasks, such as scene understanding.

The subsequent chapters build on these techniques to emphasize other social aspects of unstructured data. In Chapter 8, the authors develop an approach for facial age estimation from a data-representation perspective using personal and temporal characteristics

and a label-distribution learning algorithm. In Chapter 9, they develop a method to label faces in a photo collection for understanding kinship relationships. In Chapter 10, researchers propose a Bayesian network with the objective of recognizing occupations by viewing an image. These researchers integrate data on human body parts in social contexts into a probabilistic model to infer the occupations of multiple people in a photo.

In addition to text analytics, researchers must be aware of scholarship in image and video analytics. Because it is difficult to analyze unstructured image and video data, the authors of this chapter develop advanced techniques that will enable scholars to ask novel research questions. The challenge, however, is that because these methods are either new or experimental, their reliability and generalizability could be questioned. In the emerging discipline of big data and social media analytics, this is a risk researchers should be willing to take. I believe that this is a useful book; however, it would have been more valuable to readers if it included a section with resources that scholars could use to employ these new methods.

Big Data Is Not a Monolith

Big Data Is Not a Monolith posits that big data is ubiquitous but heterogeneous. Its chapters discuss challenges of big data research questions, methodologies, practices, and policies. For example, it engages in the big data debate—should the theory-driven approach to scientific knowledge be abandoned in favor of a data-driven approach? The book's contributors analyze big data's effects on individuals from mining to monitoring, on society from empowering to constraining, and on science from data governance to trust. Its chapters endeavor to present a more nuanced reality that emphasizes human participation, interpretation, and judgment in further developing big data.

The initial chapters discuss the monitoring and mining of individual behaviors using big data techniques for various economic, political, and surveillance purposes. As Chapter 1 discusses, in many cases, such as the sensors in more than 6.8 billion handheld phones worldwide, the question about choice in data collection is not truly available. It shows the re-identification of data at Google and Netflix. These firms released de-identified data sets only to have the

data re-identified within days by researchers that correlated them with other data sets. The authors suggest that the reliance on notice and consent at the time of collection is not sufficient and a more effective and scalable approach is needed. Machine learning, as discussed in Chapter 3, gives us new abilities to predict with remarkable accuracy; however, it also presents risk-management challenges (e.g., partially understood algorithms). Chapter 5 emphasizes the existence of data disconnects between how users think about their data and how data miners use and understand these data. Chapter 11 builds on the challenges discussed in Chapter 3 and explains the nonmonolithic context of organizational data use, data governance, and data management, which has also created obstacles in the implementation of corporate-data responsibility. Privacy and security researchers will benefit from understanding these emerging issues.

Chapter 6 elucidates a new data language, which can assist us in conceiving, instrumenting, and refiguring social interaction and social relationships. Chapter 9 emphasizes that the power of big data should not undermine the role of small and carefully selected data sets; Chapter 12 explains the creative use of big data to make strategic decisions. One example is Netflix whose acquisition of *House of Cards*, at a cost exceeding \$100 million, was based on detailed analysis of the original British show and the actors' and director's popularity.

This book provides readers with an understanding of big data heterogeneity from types of data users to big data analysis, which reveals shades of truth and power. Although its chapters cover a wide variety of topics, some would have benefited from detailed examples and evidence sourced from the authors' original research. Such coverage would have added further profundity to these chapters and thus benefited big data scholars.

Conclusions

As big data and social media analytics researchers study this phenomenon, they will encounter challenges of method, analysis, theory, and context.

- Method: How and where do we collect data to ensure that the results are valid? For example, are millions of observations, although sourced from one technology platform, sufficient to make generalizable conclusions?

- **Analysis:** What quantitative and qualitative data analysis techniques are appropriate for data being studied? For example, the popular quantitative techniques, such as structural equation modeling, will be challenged by an extremely large sample size; similarly, the major qualitative techniques could become extremely labor intensive and possibly impractical.

- **Theory:** How do we advance theory when we are overwhelmed with data? For example, which approach should we undertake if the data contradict most theories and their assumptions?

- **Context:** How can we completely understand a phenomenon without fully comprehending the context of the phenomenon? For example, is the IoT-based data used in studying energy in a social setting sufficient for us to comprehend energy usage within a community?

Although big data techniques and technologies can potentially enable us to make better decisions in situations ranging from health epidemics to natural disasters, we are still struggling to address key challenges of theory and practice. For example, in the event of a major hurricane, can we provide relief logistics in real time if we have the assistance of big data-based decision making, which integrates weather, terrain, social media, medical-expert, and relief-operations data? Quality real-time decision making and effective logistics in a natural disaster scenario are still extremely complex. Scholars, both new and established, should consider reading these three books to gain an understanding of the questions they should pursue and the challenges they must overcome as they strive to advance big data and social media analytics research.

References

IBM (2017) Bringing big data to the enterprise. Accessed March 10, 2017, <https://www-01.ibm.com/software/data/bigdata/what-is-big-data.html>.

Pratyush Bharati

Management Science and Information Systems,
University of Massachusetts, Boston, Massachusetts,
Pratyush.Bharati@umb.edu.

LLAVADOR, HUMBERTO, JOHN E. ROEMER, JOAQUIM SILVESTRE. 2015. *Sustainability for a Warming Planet*. Harvard University Press. 320 pp. \$45.00.

Since the publication of the prominent Brundtland Report (United Nations World Commission on

Environment and Development 1987), academics and policymakers have been interested in the twin concepts of sustainability and sustainable development. A second interest—studying the many facets of climate change or global warming—has also emerged. This issue is undeniably the most salient environmental problem confronting humankind today. Although economists have made several noteworthy contributions to the literature on sustainability and climate change, the authors of this book lament the fact that “sustainability appears to be of tangential concern in some of the most prominent work of economists who have concerned themselves with...man-made global warming” (p. 1). As such, the purpose of this book is to make sustainability considerations front and center in a formal study of climate change.

The proceedings in this book begin in earnest with a systematic discussion of sustainability and discounted utilitarianism. The authors sensibly contend that an ethical study of climate change and sustainability must have three components. First, each generation has a *right* to as much welfare as any other generation because when one is born is a matter of “brute luck.” Second, each generation *may* choose not to enforce its right, and hence may permit the growth of utility for subsequent generations. Third, when one models how humans weigh their own standard of living against future human development in a particular way, it is possible to compute the rate of growth of utility over time. They then go on to point out that despite the well-known problems associated with discounted utilitarianism, climate change economists have been excessively wedded to this approach. Therefore, a new and ethically sound approach is required. In Chapter 3, they discuss this approach.

This new approach involves the development, calibration, and optimization of a fully articulated model, which incorporates carbon dioxide emissions and a sector that produces knowledge. The basic objective is to analyze the nature of intergenerational equity in a world that is constrained to limit greenhouse gas emissions to keep the global temperature at an acceptably low level. A key distinguishing feature of the analysis undertaken is that the utility function does *not* depend exclusively on consumption, but also on education and knowledge, and on what the authors call an “undergraded biosphere.” This more elaborate analysis yields two interesting findings.

First, we know that in many dynamic models that utilize the so-called Ramsey formulation, the discounted utilitarian program converges to a solution for a large class of discount factors. This notwithstanding, the discounted utilitarian program associated with the model that the authors study diverges “for any reasonable discount factor . . .” (p. 155). They explain that this situation arises in part because the more traditional analyses based on the Ramsey formulation do *not* model the formation of human capital through education. Second, the authors note that instead of focusing on sustaining consumption, researchers should focus on sustaining utility. Why? Because when the focus is on utility and not on consumption, “the steady-state stock of knowledge is more than twice as large, and steady-state education is over four times as large” (p. 158). The authors provide an excellent discussion of their findings; however, because they assume that some of the mathematical results that underpin the analysis leading to these findings hold but do not show them to hold, there is some question about the reach of these findings.

In Chapter 5, the authors discuss the hitherto ignored problem of intragenerational welfare using a model that has two regions: the North (the United States) and the South (China). The key question they study is how to allocate the budget of total emissions for each generation to the two regions to ensure that it is optimal. The authors contend that the notion of egalitarianism is inconsistent with a perpetual welfare gap between the North and the South. Therefore, they focus on a scenario in which the per capita welfare in the North and South *eventually* converge. This raises the question of what “eventually” means in the model being studied. According to the authors, eventually means three generations or 75 years. They justify this selection by stating that their “motivation for this proposal is based on Thomas Schelling’s focal-point approach in bargaining . . .” (p. 211). However, they do *not* persuasively explain why convergence in three generations is a plausible focal point. Subject to this caveat, a noteworthy policy implication that emerges from the analysis conducted is that “successful agreement to curb global emissions can only be achieved if all large countries

agree to curb their growth factors to some extent and second, that the fair way to curb growth factors is to do so in a way that preserves the *relative* growth factors of GDP per capita in all regions . . .” (p. 230).

Chapter 6 contains a thought-provoking discussion of how to model the potential occurrence of catastrophes. Two points from this discussion deserve mention. First, unlike some influential work in the contemporary climate change economics literature, we are told that the generational probability of a catastrophe or the hazard rate ought *not* to be exogenous, but should “depend on the history of carbon emissions or on the concentration of carbon in the atmosphere” (p. 241). Second, the authors credibly claim that instead of viewing a catastrophe as an event that effectively wipes out the human race, it is more helpful to think of such an event as one that increases the immiseration of society.

Let me conclude this review with three points. First, this book employs sophisticated mathematical reasoning to challenge the notion of discounted utilitarianism and suggest a plausible alternative that is egalitarian in a very basic manner. Second, it points out that unlike the mathematics associated with utilitarian analyses of climate change, the maximization of what the authors support and call a “sustainabilitarian objective function” (p. 265) always yields a solution. Third, we learn that there is *no* politically feasible global solution to the climate change problem unless international negotiators and policymakers understand that there is a clear nexus between restricting emissions and curtailing growth. Therefore, readers who wish to learn more about how cutting-edge mathematical analyses can shed valuable light on the most pressing environmental concern in contemporary times will profit by perusing this book.

Reference

United Nations World Commission on Environment and Development (1987) Report of the World Commission on Environment and Development: Our common future. Accessed March 16, 2017, <http://www.un-documents.net/our-common-future.pdf>.

Amitrajeet A. Batabyal

Department of Economics, Rochester Institute of Technology,
Rochester, New York, aabgsh@rit.edu.