Introduction to the Special Issue on Perspectives on Big Data

Kalyan Singhal
Merrick School of Business, University of Baltimore, ksinghal@ubalt.edu

Qi Feng
Krannert School of Management, Purdue University, annabellefeng@purdue.edu

Ram Ganeshan
Raymond A. Mason School of Business, The College of William and Mary, ram.ganeshan@mason.wm.edu

Nada R. Sanders
D'Amore-McKim School of Business, Northeastern University, n.sanders@northeastern.edu

J. George Shanthikumar
Krannert School of Management, Purdue University, shanthikumar@purdue.edu

Big data has the potential of offering valuable insights into the way organizations function, and it is changing the way organizations make decisions. Nine invited essays provide a wide range of perspectives on the role of big data in customer-driven supply chains, healthcare operations, retail operations, demand planning and manufacturing, environmental and social issues, humanitarian operations, agriculture supply chains, and service operations. Decision makers should have clean, valid, and reliable data, and they should have a thorough understanding of the contexts of applications. Big data shorten virtual distance to customers, and thus facilitate personalization of products and services. Successful implementation of big data applications requires sharing the data with appropriate stakeholders.

Key words: big data; data analytics; healthcare operations; operations management; service operations; supply-chain management

1. Introduction

Big data and associated technological developments (e.g., internet of things, industrial internet of things, cyber-physical systems) are dramatically changing the landscape of operations and supply-chain management. Researchers in our field, as in many others, are increasingly devoting significant effort to understand the emerging business models and changing business principles. Given many new phenomena, many unknowns have yet to be discovered, unknowns that will affect how the associated applications may evolve and how the related research should be shaped. We don't know whether big data is fundamentally changing the ways we conduct research, or whether we can just hit the new nails with the old hammers.

With these developments in mind, we invited some leading scholars in our field to share their thoughts on how big data is affecting research in POM. We have collected nine essays in this special issue. Each of these essays offers interesting views on how big data is reshaping the research across various domains. The authors provide examples of new business models enabled by big data.

2. Nine Essays

- *An evolutionary view*: Hau Lee adapts a three-stage framework for technology innovation to envision how big data may evolve and change the way we manage the supply chain. He highlights the shift toward managing the "long tail" and customer-driven supply chains.

- *Healthcare operations*: Wallace Hopp, Jun Li, and Guihua Wang explain the use of observational data from nontraditional sources to supplement the traditional experimental data from clinical studies and thus to transform the one-size-fit-all approach to precision medicine.

- *Retail operations*: Marshall Fisher and Ananth Raman share their knowledge of how data analytics lead to service innovations. They focus on the transformation from data to improved decisions concerning assortment, online dynamic pricing, online order fulfillment, and store closings.

- *Demand planning and manufacturing*: Qi Feng and George Shanthikumar focus on how big data is changing operational planning. They demonstrate how one should use prototype models in operations and production management.

- *An environmental and social perspective*: Charles Corbett describes examples of smart ways of using data to reduce the environmental footprint, to manage energy efficiency, and to plan weather-based service and production. He also stresses the challenges in using big data analysis, such as creating undesired social and ethical consequences.

- *Humanitarian operations*: Jayashankar Swaminathan discusses how descriptive, prescriptive and predictive analysis can enable rapid, impactful, sustained and efficient humanitarian decision making. He offers insights on how decision makers can use data to improve their identification of needy populations, optimize supplier bases and resolve production bottlenecks in the distribution of humanitarian relief products and services.

- *Agriculture supply chains*: Sripad Devalkar, Sridhar Seshadri, Chitrabhanu Ghosh, and Allen Mathias recommend innovative data collection applications and information dissemination applications. Using market price analytics as an example, they explain how algorithmic data analysis and decision making can help to improve the productivity of farmers.

- *Service operations*: Maxime Cohen highlights how the emergence of big data has led to the transformation from intuition-based decision making to evidence-based decision making. He also emphasizes the role of the marketplace in producing innovative offerings in financial services, transportation, hospitality, and online platforms.

- *An overview*: Samayita Guha and Subodha Kumar summarize existing research on various issues that arise with big data in research on information systems, operations, and healthcare.

3. Transformation from Data to Efficient Decisions

As a common theme, the essays in this issue identify opportunities and challenges in research using big data. Their authors emphasize the need for research and the value of research that enables the transformation from data to efficient decisions. For example, Devalkar, Seshadri, Ghosh, and Mathias write that while data access is provided to farmers, they lack ways of using the data to guide their planning. Guha and Kumar point to the need for research to evaluate the benefit of adopting technologies or systems that collect, process and analyze big data. Corbett argues that more data can lead to worse decisions if not aggregated and structured properly. Hopp, Li, and Wang point out that a pure machine learning approach, while shown to be effective for predicting medical outcomes, is not directly helpful to guide decisions for individual patients.

The authors particularly emphasized two aspects to ensure quality decision making processes. The first is the risk associated with the data itself. The veracity of big data requires careful design of data-acquisition and calibration strategies and of feature-extraction and selection strategies so that decision makers have clean, valid, and reliable inputs to use in making decisions. The second aspect lies in appropriate ways of using data, which requires a thorough understanding of the application contexts and a clever integration of data with structural knowledge.

4. Personalization

Many of the authors of the essays also recognize new research issues with the trend of personalization. Lee points out the shortened virtual distance to consumers and transformation toward customer-driven planning. Cohen discusses the mechanisms of real-time personalization and targeted promotions, and Feng and Shanthikumar devise prototype models to demonstrate data integration for personalized demand planning. Devalkar, Seshadri, Ghosh, and Mathias describe algorithmic analysis based on the diversity of individual farms. Fisher and Raman highlight the value of tracking the behaviors of individual retail customers, and Guha and Kumar point out its value for healthcare patients. Hopp, Li and Wang suggest that combination of data on patients' heterogeneous responses to treatment alternatives and data on individual patient preferences can enable medical decisions customized at the level of individual patients. Several authors also point out

the potential challenges of personalization that one should not ignore in one's research. These include *security and privacy* of personal data (Cohen; Guha and Kumar; Hopp, Li and Wang), *ethical* use of data (Corbett), and *fairness and discrimination* (Cohen; Swaminathan).

5. Data Sharing and Benefit Distribution

Another potential research avenue identified is related to data sharing and benefit distribution. Successful implementation of data analytics requires sharing the right data with appropriate stakeholders (Corbett). Feng and Shanthikumar discuss the need for data exchange and for a coordination mechanism in a manufacturing network to enhance overall capability, while avoiding a learning race. Devalkar, Seshadri, Ghosh, and Mathias stress the importance of sharing the right data with the right entity to ensure fair distribution of the benefit generated by big data without creating unbalanced power structures.

6. System Thinking

Finally, operations management scholars have long advocated system thinking in our research. The emergence of big data requires us to enlarge the scope of the system when we perform analyses and develop models. In the process, we also generate new research questions about the system. Data analysis without a system view, however, can lead to pitfalls. Cohen elaborates on how analysis of big data without a thorough understanding of the system can lead to "machine bias" and "spurious correlations." Corbett also points out the danger of "letting data availability drive priority" and advocates careful consideration of underlying phenomena.