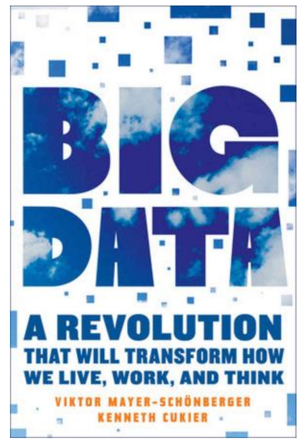
International Journal of Communication 7 (2013), Book Review 2727–2729 1932–8036/2013BKR0009

Viktor Mayer-Schonberger and Kenneth Cukier, **Big Data: A Revolution That Will Transform How We Live, Work and Think**, Canada: Eamon Dolan/Houghton Mifflin Harcourt, 2013, 242 pp., $27.00 (hardcover).

Reviewed by
Yadira Espinal
University of Southern California



Viktor Mayer-Schonberger and Kenneth Cukier's ***Big Data: A Revolution That Will Transform How We Live, Work and Think*** unveils the future possibilities of building on the analysis of vast amounts of data. Hypothetical correlation goes out the window, and the new methodology behind datafication opens new possibilities for companies and governments in the 21st century. The authors explain the importance of this data-driven approach, how messiness trumps exactitude, and the unlimited potential for many of data completeness. They also illustrate how searches on Google open new possibilities, but caution of the dark future of shattered privacy that could prevail as depicted in movies like *Minority Report*.

Each of the 10 chapters makes a main point about "Big Data" and its implications. The authors discuss the significance of sampling vast amounts of data, the prominence of datafication, the value of data, the risks of data, and the tools needed to control a data-driven society that is able to intrude on everyone's privacy. They suggest that Farecast, a company that tells online buyers if prices are more likely to rise or fall before their scheduled trips "is the epitome of a big-data company and an example of where the world is headed" (p. 5). If all goes well, data companies will generate greater growth and affluence globally.

The first three chapters of *Big Data* analyze the now, the more, and the messy of the current state of data analysis. Flight farecasting (p. 3) is an example of the "what" in data, which enumerates "what" is known about a flight (number of seats, location, time, etc.). Farecasting causes a ripple effect on the observations of data analysis and introduces the concepts of "the more and the messy." More data in a sample set increases its validity and helps decipher possibilities. For example, the 1880 U.S. Census took 10 years to compile, so the results,when reported, were obsolete (p. 22). Big Data allows the integrated analysis process of U.S Census Bureau data, replacing random sampling. Vast amounts of data can be compiled that encompass a population. Since the days of random sampling (a "messy" strategy), major advances have ensued. Methodologies of exactitude support the art of the comparison of multiple data sets. For example, British Petroleum supports a huge number of wired and wireless censors installed across their plants. Constantly measuring the stress in their pipes, instead of at intervals, proved that some crude oils are more destructive than are others (p. 41). These results came much more quickly than was previously possible. Adopting exactitude and precision of data allows industries across all platforms to implement necessary changes to expand their perspective markets.

The cultural shift of hypothetical correlation to datafication has helped large data-driven companies become more successful. Today, Amazon monitors the interests of its consumers, and Wal-Mart discovered that pop tart consumption rises prior to hurricanes. Such examples show how statistics can provide insights that can be used to boost sales. This ideology of statistical forecasting is called *predictive analytics*. It helps businesses "foresee events before they happen" (p. 79), allowing them to make informed and profitable decisions. But how do companies take search inquiries and turn them into data analytic inferences? Schonberger and Cukier explain how using datafication can make "words become data" (p. 83). Early on, Google digitized text by scanning millions of books and making them accessible on their platform. The search giant transformed text images into an algorithm text search, allowing consumers to track down the book of their choice using word searches. Text digitalization relied on various mediums, not just on the retrieval of numbers. Today, Google also keeps track of social media interactions and examines how datafication revolutionizes data contraction and ties business sales to search queries or topics of conversations trending on social networks. The revolution of SEO engines and their abilities to quantify the words of consumers engender sales for companies worldwide. Schonberger and Cukier itemize this cultural shift by dissecting search inquiries and unveil the profits arising from predictive analysis.

Datafication helps monetization. Mining information about social media as it integrates yields a gold mine for businesses and will structure business sales' strategies going forward. "Data exhaust is the mechanism behind voice recognition, spam filters, language translation and much more. When users indicate to a voice recognition program that it has misunderstood what they said, they, in effect, train their system to get better" (p. 114). Isolating the flaws of a firm's strategy is as important to devising more effective strategies, and data exhaust provides a filter that aggregates the flaws of data analysis in efforts to shift business strategies. Data exhaust is used to monitor all consumer behavior. Monitoring how long people discuss a trending topic, commenting on family photos, and liking certain artists' profile pages creates a pool of data that companies can tap into when targeting consumers. Datafication provides access to the innate interests of consumers by tracking their daily rituals. Such innovations improve the ability to forecast the success of a product and enhance or prevent shortcomings that companies face.

Data exhaust leads to the methodology of data implication and the conclusion that data-driven companies can leverage this information. *Big Data* explains the transformation of MasterCard from a firm governed by payment processing to one where data processing is paramount. "MasterCard Advisors aggregate and analyze 65 billion transactions and 1.5 billion cardholders in 210 countries in order to divine consumer trends. Then it sells that information to others" (p. 127). Datafication is now a priority area for MasterCard, allowing it to shift into new service and profit in different sectors. The ability to access such data is feasible and fast developing. Companies other than Google and Amazon are swimming in data wealth. Firms in all sectors want to use new big data analysis techniques to follow consumer trends and build their profits.

The profitability of data raises ominous risks that threaten to undermine the control presented by Big Data. Datafication risks arise when crossing the lines of consumer analysis to invasion of privacy. The authors see the threat of "penalties based on propensity" (p. 151) or the ability to use data predictions to punish people for crimes before they commit them. They question this presumption, quoting the line of dialogue from *Minority Report*: "I am placing you under arrest for the future murder of Sarah Marks that was to take place today" (p. 158). We risk this kind of "probable cause" if we lose control of data consumption. Access to data can create a fixation for power and cloud the judgment of data strategists who might be tempted to intrude. If access to personal information is wide open, societal sanctions and quarantines based on Google searches become more likely.

Fortunately, Schonberger and Cukier suggest how to minimize such risks: Make data- driven companies responsible for auditing their information and avoiding breeches of confidentiality. Data control would lie in the hands of "algorithmists" (p. 180) who would review data analysis and take a vow of confidentiality to impede government intrusion. If this somehow worked in practice, algorithmists would provide adequate data sustainability without the acquisitive inputs of data-driven enterprises.

Ultimately, *Big Data: A Revolution That Will Transform How We Live, Work and Think* is for forward thinkers who crave and fear new data techniques that could spur profits and understanding, but who also recognize the dangers that Big Data might pose to privacy and individual liberty. Schonberger and Cukier's work transforms business models to data-driven focus strategies while acknowledging the risk factors that data access can create. "What is great about human beings is what the algorithms and silicon chips don't reveal . . . the sparks of invention becomes what data does not say" (p. 196). Appreciating the value of data, but leaving room for intuition and natural inventiveness to take its course is at the heart of *Big Data.*