

RESEARCH METHODS

Social science, today

A breezy, personal guide provides a road map to solid computational social science research

By David Lazer

Subatomic particle physics has CERN. Astronomy has the Hubble telescope. Social science has the Internet, smartphones, email, social media, satellites, and a myriad of other ways to follow human behavior. The gods of the information age have produced a whole panoply of technologies for social research along the journey to other destinations.

Generally, social scientists have been poorly equipped to deal with the 21st-century deluge of large-scale complex data. Computer scientists, well equipped to handle the data, are often ignorant of social theory and of foundational research methods in the social sciences. What is needed is an articulation of core principles of designing research that are accessible to multiple disciplines.

Into this breach steps Matthew Salganik. Salganik is one of the first natural-born computational social scientists, a sociologist whose doctoral work was one of the early landmark projects in the field (1). *Bit by Bit* is 90% textbook, 10% biography, putting into personal context issues that Salganik was among the first to wrestle with.

The volume fits solidly into a genre of textbooks in the social sciences, inaugurated in

1963 by Donald Campbell and Julian Stanley (2), that might be termed “practical epistemology.” That is, they address such questions as these: How do we create knowledge in the social sciences? How do we measure things? What’s the basis for statements like “20% of the population is X,” or “X causes Y”?

In a sense, this type of textbook is about the architecture of research, the structure that will be necessary to support potential assertions. However, the range of social science research activities has changed dramatically, with profound implications for potential design.

Much as a steel frame enables the construction of buildings that reach toward the heavens and transform city skylines, the pervasive instrumentation of human behavior should likewise transform social science research. *Bit by Bit* is the first—and quite worthy—successor to the Campbell volume, aimed at informing students how to design those blueprints in the emerging field of computational social science.

Other than the introduction and conclusion, there are five core chapters to *Bit by Bit*. The first, “Observing behavior,” is focused on the massive, passive data collection that occurs in everyday life, identifying key opportunities and challenges in using big data for research. “Asking questions” adapts lessons learned from survey meth-

odology to big data: Core concerns about representativeness and measurement are amplified when recycling big data collected for other purposes.

“Running experiments” discusses the scientific potential to run heretofore inconceivably large-scale experiments. The Internet, in particular, argues Salganik, enables the facilitation of large group experiments as well as the evaluation of heterogeneity of treatment effects. “Creating mass collaboration” discusses the harnessing of the small efforts of many people for large-scale scientific applications.

The final chapter, “Ethics,” is a thoughtful exposition on the core principles around ethical research generally, with a particular focus on the challenges that large-scale data collection poses. Privacy and security concerns, for example, are magnified with scale and consentless third-party data collection.

The text is clearly written—even breezy, in parts. It puts the reader in the shoes of the researcher: What decisions were made, why, and were those the best choices? It is suitable for an advanced undergraduate or graduate class in methodology, with a rigorous, mathematical appendix and a range of useful problems at the conclusion of each chapter.

This book is not the place to learn about cutting-edge computational techniques. However, if

you want to reflect on the potential value of, say, deep learning to understanding human behavioral data, there are relevant lessons. Despite the rapid evolution of the domain, this book will likely have staying power.

It is telling that my only complaint is that I would have liked to see more topics covered. How, for example, do we translate more quasi-experimental approaches to the big data world? How do we rethink the power of panel data when there may be thousands or millions of observations per individual? How do we manage the complex workflow of a computational social science project? How do we deal with the issue of replication with data that often cannot be shared?

It may be, however, that the field more generally must advance before these chapters can be written. In the interim, *Bit by Bit* will be required reading for my students. ■



Bit by Bit
Social Research in
the Digital Age
Matthew J. Salganik
Princeton University
Press, 2017. 445 pp.

The reviewer is at the Department of Political Science and the College of Computer and Information Science, Northeastern University, Boston, MA 02115, USA, and the Institute for Quantitative Social Science, Harvard University, Cambridge, MA 02138, USA. Email: d.lazer@northeastern.edu



A deluge of digital data is transforming the way we collect and interpret social research.

REFERENCES

1. M. J. Salganik, P. S. Dodds, D. J. Watts, *Science* **311**, 854 (2006).
2. D. Campbell, J. Stanley, *Experimental and Quasi-Experimental Designs for Research* (Houghton Mifflin, Boston, 1963).

10.1126/science.aag0679

sciencemag.org **SCIENCE**

PHOTO: TZIDOU/ISTOCKPHOTO