

# Book Review: Bit by Bit: Social Research in the Digital Age by Matthew J. Salganik

*In **Bit by Bit: Social Research in the Digital Age**, Matthew J. Salganik explores the process of undertaking social research in the digital era, examining a wide range of concepts while also offering teaching activities and materials. In bringing together the expertise of social and data scientists to the benefit of both, this is a comprehensive overview of new approaches to social research in our time, recommends **Marziyeh Ebrahimi**.*

*This review originally appeared on [LSE Review of Books](#) and is published under a [CC BY-NC-ND 2.0 UK](#) license.*

**Bit by Bit: Social Research in the Digital Age. Matthew J. Salganik. Princeton University Press. 2018.**

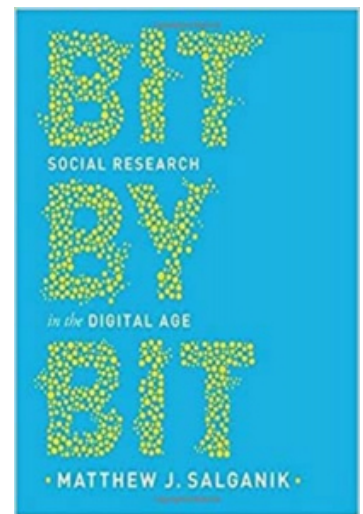
Find this book: 

*Bit by Bit: Social Research in the Digital Age*, written by [Matthew J. Salganik](#), Professor at Princeton University, gives clear and detailed information about carrying out social research in the digital era. The book has up-to-date content which covers a wide range of concepts related to online content and it also offers activities and solutions, which makes it a perfect text both for researchers working with big data and for university lecturers looking for a course book on this topic. The book is accessible on [its website](#) and Salganik has also provided an array of teaching materials, such as syllabuses and slides, for each chapter.

In the introduction, Salganik points out that this book has two key audiences that have a lot to learn from each other. On the one hand, it is for social scientists who have training and experience studying social behaviour; on the other hand, the book is also for data scientists who have received training in the computer sciences or engineering. This book attempts to bring these two communities together to produce something richer and more interesting than either community might produce individually.

*Bit by Bit* has seven chapters and progresses through four broad research designs: observing behaviour; asking questions; running experiments; and creating mass collaboration. At the end of each chapter it offers a wide range of activities, which are labelled by degree of difficulty and by the skills that are required to solve them, including maths, coding, or data collection.

In the analogue age, collecting data about behaviour – data related to who does what and when – was expensive and therefore relatively rare. Now, in the digital age, the behaviour of billions of people is recorded, stored, and analysable. Since this data is a by-product of people's everyday actions, they are often called "digital traces". The ever-increasing flood of big data means that we have moved from a world where behaviour data is scarce to one in which it is plentiful.



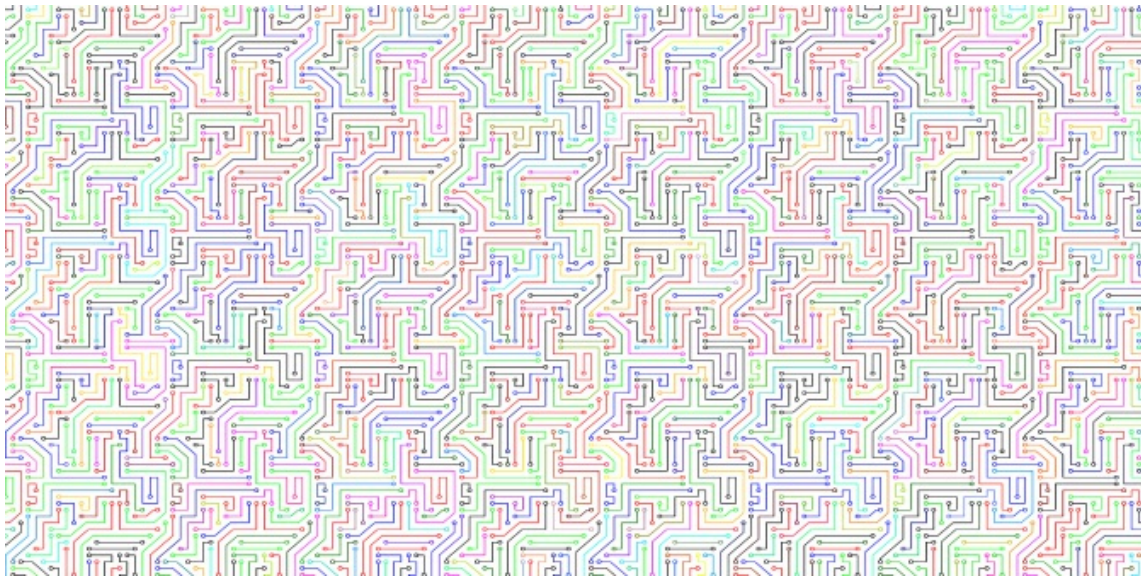


Image credit: [GDJ](#), via Pixabay (licensed under a [CC0 1.0](#) license).

Salganik considers that the first step to learning about big data is realising that it is a part of a broader category of data that has been used for social research for many years: observational data. Roughly speaking, this concerns any data that results from observing a social system without intervening in some way. Nonetheless, as he writes:

“Big data sources are everywhere, but using them for social research can be tricky. In my experience, there is something like a ‘no free lunch’ rule for data: if you don’t put in a lot of work collecting it, then you are probably going to have to put in a lot of work into thinking about it and analyzing it.”

Regarding big data sources, Salganik writes that many researchers immediately focus on online data created and collected by companies, such as search engine logs and social media posts. However, this leaves out two other important sources of big data. First, corporate big data sources that come from digital devices in the physical world; and, second, government administrative records. It’s true that these sources of information have been always [(un)ethically] used by social scientists, but what has changed is digitisation, which has made it dramatically easier for governments to collect, transmit, store, and analyse data.

Salganik introduces the ten common characteristics of big data: big; always on; nonreactive; incomplete; inaccessible; non-representative; drifting; algorithmically confounded; dirty; and sensitive. For example, “always-on” data collection enables researchers to study unexpected events in ways that would not otherwise be possible. For example, those studying the Occupy Gezi protests in Turkey in summer 2013 would typically focus on the behaviour of protesters during the event, since the data was publicly accessible online through social networks. Regarding “nonreactivity”, Salganik writes that:

“One challenge of social research is that people can change their behavior when they know that they are being observed by researchers. Social scientists generally call this reactivity. One aspect of big data that many researchers find promising is that participants are generally not aware that their data are being captured or they have become so accustomed to this data collection that it no longer changes their behavior. Because participants are nonreactive, therefore, many sources of big data can be used to study behavior that has not been amenable to accurate measurement previously.”

In addition, he identified three main strategies for learning from big data sources: counting things; forecasting things; and approximating experiments.

Amongst the digital research concepts that Salganik has developed throughout this book, creating mass collaboration, the title of Chapter Five, gives a unique perspective on the new forms of social research and collaboration in our time. He explains that:

“The digital age fortunately enables many novel forms of collaboration. As an example, the key to Wikipedia’s success was not new knowledge; rather, it was a new form of collaboration. Mass collaboration blends ideas from citizen science, which is about involving citizens in the scientific process, and crowd-sourcing, which concerns taking a problem ordinarily solved within an organisation and instead outsourcing it to a crowd and collective intelligence.”

However, the digital age has not only created new opportunities for collecting and analysing social data, but has also produced new ethical challenges. Salganik argues that four principles can guide researchers facing ethical uncertainty: respect for persons; beneficence; justice; and respect for law and public interests. However, research ethics involves struggling over decisions about what to do and what not to do. As a solution to these dilemmas, he suggests that the researchers put themselves in other people’s shoes: “Often researchers are so focused on the scientific aims of their work that they see the world only through that lens. This myopia can lead to bad ethical judgement”. Therefore, when you are thinking about your study, try to imagine how your participants, other relevant stakeholders, and even a journalist might react to the research. This perspective-taking is different from imagining how you would feel in each of these positions. Rather, it is trying to imagine how these other people will feel.

Ultimately, Salganik believes that social researchers are in the process of making a transition akin to the shift from photography to cinematography. The future of social research will be a combination of social science and data science, where there will be more participant-centred data collection, and ethics will move from being a peripheral to central concern and a topic of research in its own right. *Bit by Bit* gives a comprehensive overview about this new way of doing social research in our time, and is one of those books that all social academics should read or consider.

---

**Marziyeh Ebrahimi** started writing professionally as a journalist at the news agency, ISNA, in Iran when she was 17. She defended her doctoral dissertation on users behaviour analysis at Universidad de Navarra in Spain in 2016. She currently collaborates as a post-doc researcher with Universidad Panamericana in Mexico – Aguascalientes.

*Note: This review gives the views of the author, and not the position of the LSE Impact Blog, or of the London School of Economics.*