The SAGE Handbook of Data and Society

General Introduction

Tommaso Venturini, Amelia Acker, Jean-Christophe Plantin, and Tone Walford

Venturini, T., Acker, A., Plantin, J. C., & Walford, T. (Eds.). (2024).
The SAGE Handbook of Data and Society (1st ed).
SAGE Publications Ltd.

Something has changed, or is changing, in the way we think of data and digital technologies. Since the dawn of the internet and maybe since the introduction of computers, digital technologies have raised great expectations and general optimism. While dystopian fiction and critical scholarship have long warned about the risks of digital developments, the general mood until quite recently has been bright and breezy. To be sure, many did call our attention to the possible *misuses* of digital technologies, but the increase in data and computational power was not considered, in itself, a menace. Moore's Law (which predicts the exponential growth of transistors in integrated circuits) has been around since the 1960s, but very few people considered it a threat to society or demanded interventions to restrain it.

The mood, however, has changed. At some point, or rather at many points in the last few years, the relentless digitalization and datafication of society has started to be perceived as a problem. A series of scandals about micro-targeting in political and commercial marketing, the increasingly oligopolistic structures of a few tech companies, new data protection regulations in Europe, India, and elsewhere, the implosion of many crypto-currencies, the startling power of deep-learning AIs, and many other small and big events have opened deep cracks in our digital optimism. None of these developments are entirely new, and many of them have roots that extend far back in the history of digital technologies, yet their accumulation is changing the ways in which we look at data and computation. Suddenly, 'more' has stopped rhyming with 'better', as a growing number of actors have started to call for a pause in technological development. Even Silicon Valley gurus and CEOs have begun to voice concerns about digital technologies, such as the recent moratorium proposal to 'pause Giant AI Experiments'. Even if these statements remain largely perfunctory (and while investments are rather pushing an even greater acceleration), they do signal that the public's perceptions and expectations have changed and that the complications of the digital have taken center stage.

The Sage Handbook of Data and Society speaks to this new perspective and offers a complete overview of the scholarship that for a long time – and increasingly in the last years – has investigated and critiqued the role of data in society.

Critique, to be sure, means more than just criticism. The scholarship gathered in this Handbook does not shy away from considering the far-reaching societal consequences of digital technologies but refuses to merely endorse a wholesale denunciation of their risks. Instead, it engages with a variety of specific datasets, cases, key concepts, and methods to dissect their subtlest mechanisms and tiniest details. In the words of Bruno Latour, it seeks not a blanket rejection of digital technologies but rather a 'critical proximity' with data in society.

The chapters of this Handbook all share an intimate knowledge of the technologies that they discuss. They carefully consider not just the risks and the possibilities of datafication but also the actual advantages and shortcomings that they bring, both in general and specific contexts. In different ways, the authors contributing to this book explore topics as different as digital labor, platform economy, online celebrity, internet subcultures, algorithmic policing, data doubles, gender, decolonization, and environmental justice, among many others. The Handbook features contributions focused on data and evidence collected on a plurality of social contexts but also on the algorithms used to exploit them; it is interested in the methods used both in academia and industry as well as their multiple combinations; it explores some advanced techniques of machine learning but also more mundane (but equally influential) everyday practices of social life today; it investigates not only single digital objects but also the platforms and the infrastructures that support them.

From an editorial perspective, this Introduction is maybe the only part in this book where concepts such as 'data', 'technologies', or 'digital' are used as general placeholders rather than addressing specific datasets, individual technologies, historical development, and particular forms of digitization. This allows the authors of this Handbook to avoid commonplace ideas, indiscriminate reproaches, or uncritical hype. More importantly, it allows some of the authors to imagine how data technologies could be *otherwise* developed and deployed. Other forms of datafication are possible and this Handbook offers many different visions of them.

This care for diversity is the *fil rouge* that inspired our editorial work, the goal that guided our choices for the line-up of authors, and for the curation and review process. We wanted to showcase that the ambiguity often associated with the etymology of the word 'data' is more than a flat truism: yes, data are always plural and no, they are never given.

Dedicated to *critical interdisciplinary scholarship on data and society*, this Handbook integrates perspectives from traditional disciplines such as sociology, history, and geography, but it also offers new contributions from emerging fields like digital methods, platform studies, critical algorithm studies, digital STS, and techno-anthropology among others. Across these disciplines, the authors of this book engage with a close study of datafication through a variety of approaches and methods. Chapter sections explore the influence of data on professional and personal practices, investigate how data entrench or

renew relations of power, how they empower surveillance and policing but also activism and resistance, how they redefine race and gender, how they affect the natural and built environment in which we live and the bodies that we inhabit, as well as how data transforms the way in which we conduct research.

This variety of subjects and line of questions are explored through 28 chapters organized in seven sections:

Infrastructures of Data. A central point of the investigation concerns the many infrastructures necessary at all the stages of the production and use of data. Considering infrastructure as socio-technical systems, the chapters of this section explore the key concepts of infrastructure studies (such as materiality, maintenance, temporality, and scale) to the study of contemporary data systems across various countries and contexts.

Labor of Data. While 'data' are often presented as 'given' and easily available, many forms of labor are involved in the collecting, processing, and publishing of data resources. Data systems are also used to reconfigure how we work and under what status. This section analyzes the complex transformations of work and digital labor due to data-driven applications. Authors critically investigate invisible and 'free' labor in digital ecosystems, the relations between data and the platform economy, and the mobilization of collective labor.

Power and Sovereignty of Data. The dominant discourse about data is often criticized for being falsely objective and apolitical. In this section we examine data as occasions for struggle from which new forms of political agency can emerge, but which also can re-inscribe old hierarchies of power. This section explores data as an ongoing, open ended, and dynamic site where meaning, material resources, and representation are debated. These chapters serve as a basis from which we can re-theorize data technologies from a focus on representation and power, to an exploration of data-saturated structured experiences of being governed, surveilled and policed by data technologies.

Data and Crises of Nature. Data practices and infrastructures of measurement and monitoring have always been the privileged methods of western scientific engagement with the environment. As critical scholars have pointed out, rather than simply represent or describe the earth system, these practices enact particular versions of nature. This critical move, however, is becoming more complex as the environment is increasingly understood to be under threat and in a constant state of crisis. These chapters examine the role of data and data practices in creating new forms of nature, of environmental politics, and of critical engagement with expert knowledge. They show how, over the past four decades, the environment has been increasingly datafied and modeled by experts; but it has also seen the re-emergence of citizen science and amateur engagement through data practices and the emergence of new markets for environmental data and AI.

Data and the Bodies. If in social theory 'the body' has been almost constantly theorized at the intersection of the biological and the social, how can this be extended to take account of the new forms of datafied bodies and the heterogeneous informational infrastructures and practices from which they emerge? This extension need not be seen as a rupture: information technologies and imaginaries have been at the heart of how the body has been conceptualized, in EuroAmerica at least, for a long time (the idea of genetic 'code', for example, or neural 'networks'). Contemporary data practices bring bodies into view in ways that are both historically continuous with past practices of quantification and control, inscribing forms of bodily violence such as race or gender, and that constitute transformations of bodies as sites of social, cultural, and political inscription. The contributions to this section investigate these dynamics in contexts as apparently disparate as genomics, self-tracking, public health, and policing.

The Sciences of Data. Social research (as all types of scientific research) has been deeply affected by the advent of digital data and computational techniques. The chapters in this section demonstrate that these technologies are not just passive tools at the service of social science but extremely active actors that have transformed the way in which we investigate social phenomena. They consider data practice from the point of view of social scientists and social actors who have integrated various techniques from data science in their practices and discovered that these techniques bring with them a variety of helpful but also harmful consequences.

Disciplines of Data. The relationship between digital data and social science disciplines goes two ways. On the one hand, the conceptual framework developed by different scholarly traditions can help us understand the social impacts of data. On the other hand, the same frameworks are profoundly renewed by their increasing digital operationalization and the role that data play in it. The contributions in this section illustrate both these transformations by unfolding the datafied version of four notions that are key to as many academic disciplines: space and geography, information and design, cultures and ethnography, and critique and critical studies.

The strength of this Handbook comes from the diversity of its contributions. As editors, we aimed for this Handbook to be a polyphonic project. We did not try to make a single argument but rather to offer a multiplicity of voices, various conceptual viewpoints, and a range of empirical research on a variety of phenomena related to data and society. Altogether, this Handbook contributes to the consolidation of the vibrant and interdisciplinary field that is now called 'critical data studies'. Its chapters show that critical data studies are neither a cacophony of discourses unified only by their object nor a unified theory of the digital but rather a disciplinary trading zone where different academic traditions can meet and exchange ideas. As this Handbook highlights, while the roots of this research field are sprawling and still green, they are strengthened by the multiple connections of the intellectual rhizome that they form.

For a long time critical data studies seemed somewhat *residual*. As long as the mood was one of data optimism, the kind of critique that runs through this Handbook appeared as nothing more than a set of useful guardrails in case of side-slips and side-effects. The questions raised and the problems emphasized by critical data studies seemed like tangible yet secondary *unintended consequences*. Not anymore. Now that digital data and technologies are omnipresent, so are their societal consequences. Their connections to race, gender, power, the natural and built environment, our bodies, and our scientific and everyday cultures look not like unfortunate byproducts but rather the beating heart of these technologies. Critical data studies, it turns out, are data studies *tout court*, and this Handbook hopefully will help to make this clear.