Preface

For being good to be used and reused, scientific data need people – scientists, technicians, and administrators – as embodied, cognizant, social, and cultural human beings. This book examines how scientific data not only represent information, but are also deeply implicated in social actions and accountabilities. Scientific data are the remarkable epistemic and social "stuff" that scientists use to learn about nature, instruct students, build connections, transmit knowledge, grow trust, and evaluate the work of other researchers. These uses often go together. What data are like shapes how these interactions and relations unfold, and vice versa. Drawing on an ethnographic study of research in astronomy, I focus on instruction, practical reasoning, and decision-making as lenses into scientific work with large datasets (but also with smaller ones).

I set out from the tenet that humans are fundamentally evaluative. We are concerned that others understand our intentions adequately, knowing that our actions are being evaluated, and we examine others' actions and intentions likewise. We pursue our goals while seeking to avoid blame and embarrassment. How ordinary members of various societies do this has long been of interest to anthropologists and sociologists. I adopt this perspective to examine the work of researchers, students, and technicians in contemporary scientific work. A second theme, resulting from the first, is my focus on what I call data-centric socialities: forms of social coordination unfolding in the making, use, and publication of data. My notion of data-centrism – that making, using, and publishing data causes people to interact and engage in specifiable ways – is inspired by sociologist Georg Simmel.

I begin with identifying digital data as media that afford certain uses and interactions and then examine successively how researchers learn from technicians about data production, instruct new data users, use diagrams to cultivate data, engage mundane reasoning to resolve disjunctive findings, organize collaborative work, manage normative issues in the open

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 20:04:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009686754.001

access to data, and seek to encode their collective knowledge in a dataset. If scientific data are epistemic and social stuff, then natural scientists are, by necessity, also social inquirers. As I approached research work as an ethnographer, I realized that scientists and technicians were, at times, doing a sort of ethnography themselves, using practices that are bound to become more important in the future.

These matters become visible only at the fine granularity of social interaction and the technical detail of a specific science. But their lessons are more general, because uses of social accounting practices are so widespread and common. Sciences other than astronomy may, of course, use other media, engage other disciplinary objects, and use other epistemic orders to organize their work. But they will need to train new members, use diagrams to interpret data, use forms of mundane reasoning, organize joint work, act normatively, transmit knowledge – and engage accounting practices along the way.

How Data Need People is published as high-volume data flows, cloud-based databases, machine learning, and generative artificial intelligence are transforming much scientific work. This is a time of rapid technological and conceptual change, but I am confident that attending to interactions and accounting practices will help when making sense of the future social fabric of science, data-rich or otherwise, whatever the participant or mediatory statuses of machines may be. I hope that this approach will also contribute to a rapprochement of the history and philosophy of science with the sociology of science, which some critics have reduced to Actor-Network Theory and the Sociology of Scientific Knowledge at the expense of alternatives, including the naturalistic and interactionist approach that I pursue here. Some philosophers' recent interest in epistemic agents (Chang 2022) and their application of social science methods (Veigl and Currie 2025) offer an opening. After all, epistemic agents are unavoidably also social agents, and, as such, users and subjects of accounting practices.

This book marks the end of what became, for me, a long journey. It began with an ethnographic study – inspired by Ludwik Fleck's (1979 [1935]) *Genesis and Development of a Scientific Fact* as well as Bruno Latour and Steve Woolgar's (1986 [1979]) *Laboratory Life: The Construction of Scientific Facts* – of how astronomers make, process, and use digital data in the construction of a scientific fact. But in doing so I was drawn to attend to scientists as people. I was also inspired by historical and epistemological studies, like Hans-Jörg Rheinberger's (1997) fascinating account of the discovery of protein synthesis, but missed more attention to human actors and their social commitments in them.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 20:04:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009686754.001

At first unsure where to start, I began with detailed examinations of extended sequences of work that I had witnessed, from the training of PhD students over two years (Chapter 3) to the collaborative fixation of a dataset over several months (Chapter 8). Doing so made me revisit my anthropological toolbox and learn about, and from, interactionist approaches like ethnomethodology. Along the way I continued my fieldwork with revisits, additional ethnography, interviews, and attendance at workshops and conferences. Thus equipped, I examined a series of topics that seemed to be of enduring importance for scientific research with large datasets (and smaller ones). Even though this book presents an account of research in astronomy, its conceptual work owes much to my ongoing study of research in paleoceanography as well.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 20:04:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009686754.001