Manuscript

Copyright The Philosophy of Science Association 2018. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/697732

Sabina Leonelli, ***Data-Centric Biology: A Philosophical Study***. Chicago: University of Chicago Press (2016), 281 pp., $35.00 (paperback).

Beckett Sterner
School of Life Sciences
Arizona State University
E-mail: beckett.sterner@asu.edu

Introduction

The word "data" is everywhere in current discussions of science. How we store and share things people label "data" has become a central concern for the Open Science movement, for example, and the National Science Foundation and National Institutes of Health have invested billions of dollars in the creation of publicly accessible databases as a major new source of intellectual capital in science and industry. What is new here? Has there been a shift in what "data" means that is key to understanding the future of science?

Sabina Leonelli's new book, *Data-Centric Biology: A Philosophical Study*, argues for an important and fruitful answer outside the comfort zone of many philosophers. Along the way, she also delivers a range of valuable insights into the expanding efforts to standardize, automate, and communicate how scientists handle, share, reproduce, interpret, and store data. From the start, Leonelli rejects the idea that we can understand the significance of data for science in terms of intrinsic properties data possess as material traces of past processes. Similarly, the changes we're seeing in science aren't driven simply by revolutionary technologies or methods. "The real source of innovation in current biology is the attention paid to data handling and dissemination practices and the ways in which such practices mirror economic and political modes of interaction and decision making" (1). In other words, data have moved to the center of intense social, economic, and political negotiations over how science works and what responsibilities scientists have to the rest of society. "Data centrism… consists of a normative vision of how scientific knowledge should be produced in order for the research process to be efficient and trustworthy" (197).

One of Leonelli's central concerns is the potential for current efforts to regulate data to homogenize science in harmful ways, which she translates into recommendations about what philosophers and other scholars of science can contribute to the contemporary discussion. In particular, the situated and variable relation that data take to scientific evidence is a theme "that neither philosophers nor science scholars have investigated in depth, leaving a gaping hole in science studies literature about how to understand and research the role of data in science and technology" (198). Instead of trying to produce new and more general theories about the appropriate role of data in scientific research, "philosophers can unravel the epistemology of data-intensive science by identifying and analyzing the variety of situations in which the significance of data is evaluated by relevant communities" (189). The way that Leonelli

This content downloaded from 129.219.008.139 on April 04, 2018 11:47:10 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

Copyright The Philosophy of Science Association 2018. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/697732

addresses this gap moves outside analytical philosophy, though, and may therefore prove difficult for some philosophers to value. My aim in this review will therefore be a critical appreciation in the classical sense: arriving at a deeper understanding of a work's significance by connecting it to broader historical circumstances, even when these do not appear directly in the work itself.

Overview

Is adding more data always better? Is it possible to solve a host of problems all at the same time by just focusing on producing more data on a vastly larger scale? Optimists about the importance of big data for science tend to conceive of data as having intrinsic value that is insensitive to changes of context. For example, Mayer-Schönberg and Cukier argue that gathering enough data is typically sufficient to override noise or bias in statistical sampling (Mayer-Schönberger and Cukier 2013). In addition, some philosophical accounts conceptualize the scientific significance of data in purely physical terms, e.g. as marks made on a recording device, which can be characterized in general without mention of the situated aims or problems of scientists.[1] Leonelli aims to show these intrinsic views of data fail to capture the actual reasons that data do or do not have value for concrete scientific projects.

She begins by bringing into focus how data have the capacity to "move" across practical situations in science, industry, policy making, and public discourse. Chapter 1 sets out a new framework for tracking "data journeys," emphasizing how data undergo a process of decontextualization and then recontextualization in novel situations in order to become meaningful beyond their original context of production. Leonelli then shows how scientists put considerable work into enabling data journeys by "packaging" data into specific formats with associated information (metadata) about how they were made, such as experimental conditions, post-processing methods, and identifiers for biological samples.

The aim and value of making data travel establishes them as a central matter of concern for biology and other institutional actors more broadly. As material objects, data are simultaneously shared in common by individuals and social groups and interpreted as meaningful in multiple, sometimes conflicting ways. "What has propelled data into becoming protagonists of contemporary biomedicine is precisely their complex status as at once local and global, free commodities and strategic investments, common goods and grounds for competition, potential evidence and meaningless information" (66). Designing the standards under which data are produced and shared therefore becomes a social and political negotiation that establishes institutionalized relations among the modes of value that different groups bring to bear.

Recognizing the importance of data journeys and their role in constituting central matters of concern across science, however, "does not call for a new, overarching epistemology of science" (197–198), nor is the core novelty of data centrism to be found in some specific

[1] Bogen and Woodward, of course, believed that phenomena travelled across scientific contexts but not data (Leonelli 2009).

This content downloaded from 129.219.008.139 on April 04, 2018 11:47:10 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

Copyright The Philosophy of Science Association 2018. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/697732

reasoning pattern (191). Instead, in Chapter 3, Leonelli argues for a "relational" account of data that shifts our understanding of data to focus on how people interpret the products of research activity as potential evidence. A major achievement of her approach here is to bring the role of data as "tools for communication" within the purview of philosophical accounts of data, which have typically focused on the man-made characteristics of data in their original situation of production. Leonelli rejects views that claim the scientific value of data can be understood purely in terms of their status as representations or by the role data play in testing theories.

As a positive alternative, she brings a new level of sophistication to accounts of data within the "practice turn" of history and philosophy of science. Ian Hacking and Hans-Jörg Rheinberger both emphasize the material aspect of data as marks or traces, but neither fully articulates the full process that data undergo from initial production to packaging for travel to re-interpretation as evidence in new situations. When we start to track the journeys data take across this whole they generally don't maintain stable or immutable significance.

Leonelli's relational account incorporates this insight by adding human interpretation as a key element to determining what count as data. She characterizes data as material objects that "(1) are treated as potential evidence for one or more claims about phenomena and (2) are formatted and handled in ways that enable its circulation among individuals or groups for the purpose of analysis" (78). The status of objects as data is relational because it depends on whether someone takes them to be relevant to supporting or contesting a knowledge claim. "What counts as data depends on who uses them, how, and for which purposes" (78).

Chapters 4 and 5 leverage the relational account to reconsider the roles of experiment and theory in data-centric science. Leonelli argues, respectively, for the impossibility of separating experimental knowledge from its original situation of production and the underappreciated significance of descriptive classification as a form of theory. For the purposes of this review, I will focus on summarizing her arguments in Chapter 5. Here Leonelli considers "the organizing principles required to assemble and retrieve data in a database" (114) and argues that these principles deserve philosophical treatment as scientific classificatory theories. She examines the important development of computer ontologies as a basis for defining and expressing consensus knowledge as well as experimental measurements in biomedical science. The Gene Ontology (GO) was the first so-called "bio-ontology" to gain prominence and it remains a paradigm for the value of automated logical reasoning in biological data discovery and integration. Bio-ontologies are comprised of regulated vocabularies of classificatory terms and logical relationships between particular terms understood as classes or instances. As Leonelli puts it, "a philosophically interesting way to view bio-ontologies is to conceive of them as a series of descriptive propositions about biological entities and processes" (120). She goes on to argue that bio-ontologies should be treated as classificatory theories with major impacts on key epistemic activities in science such as generalization, unification, and explanation, despite the fact they don't fit the traditional form of theories philosophers have studied in physics.

The last two chapters reflect on the broader implications of data centrism for science and philosophy. Chapter 6 takes a new look at the topic of integration in terms of the opportunities and dangers that standardization and centralization pose to advancing science. Leonelli describes

This content downloaded from 129.219.008.139 on April 04, 2018 11:47:10 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

Copyright The Philosophy of Science Association 2018. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/697732

three forms of data integration using case studies from plant biology: (1) integration of data across levels of phenomena within a species, (2) integration of cross-species data for discovering molecular mechanisms, and (3) translational integration of data to address problems such as the response to plant parasite epidemics. These multiple modes of integration illustrate how data-centric biology encompasses data journeys that stretch and escape the limits of basic research on model organisms.

Finally, Chapter 7 discusses the underpinnings of Leonelli's work in John Dewey's concept of a pragmatic situation, in contrast to the common notion of context in philosophy of science. She points us away from analyzing "data" as a concept within a formalized system of language and toward studying how people interpret the material products of research activities as potential evidence. Reichenbach's old distinction between the contexts of discovery and justification implies a problematic hierarchical ordering of what matters in scientific research, which "continues to license philosophical disinterest in the relations between conceptual, material, and social practices" (179). In contrast, she argues, it is more fruitful to start by articulating the circumstances of pragmatic relevance to the concrete projects or actions of researchers. In the next section, I will pick up the question of what sort of methodological foundations are appropriate for studying human practices that are "situated" in this sense.

Sensitizing philosophers to situated meaning

From the start, Leonelli signals her intention to bracket certain philosophical discussions. Traditional analytic views of data are "grounded in the presupposition that data analysis follows logical rules that can be analyzed independently of the specific circumstances in which scientists produce data" (8). Although undercutting that assumption is one of her main aims, "relating my account to other types of philosophical scholarship, both within the analytic and continental traditions, would require a completely different book and it is not my ambition to fulfill this mandate in this text" (8). Leonelli is certainly right to get on with saying new and interesting things about data-centrism rather than rehashing the legacy of logical empiricism. In this context, however, it is worth stepping back and exploring what Leonelli is doing if she is not presenting a new analysis of how data figure in the logical rules of science.

A useful starting point, although not explicitly one of Leonelli's assumptions, is Ludwig Wittgenstein's famous dictum: "Don't think, but look!" (Wittgenstein 2010, 31). It warns against introspective thought alone as sufficient for knowing the meaning of words. "Consider for example the proceedings that we call 'games'. I mean board-games, card-games, ball-games, Olympic games, and so on. What is common to them all?—Don't say: 'There must be something common, or they would not be called "games"'—but look and see whether there is anything common to all" (Wittgenstein 2010, 31). One could similarly caution against the assumption that there *must* be something common to each use of the word "data." The point goes beyond Wittgenstein's particular notion of family resemblance to suggest that form follows function more broadly in conceptual analysis.

This content downloaded from 129.219.008.139 on April 04, 2018 11:47:10 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

Copyright The Philosophy of Science Association 2018. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/697732

More important than understanding Wittgenstein here is clarifying what version of this imperative Leonelli has adopted in her approach, because it matters for the appropriate standards we should bring to her work. Data-centrism on her view is a practical phenomenon, a matter of action, rather than a linguistic phenomenon where one can abstract meaning and argument from pragmatics. "Data centrism can be described not by reference to a specific method, technology, or reasoning pattern but rather as encompassing a particular *model of attention* within research, within which concerns around data handling take precedence over theoretical questions around the logical implications of [a] given axiom or stipulation." (178, emphasis added).

Leonelli's relational account of data expresses the importance she places on tracking actors' interpretations in specific situations. Recall that her intended contrast is with accounts that ascribe "intrinsic" meaning to data insofar as this meaning is independent of how people relate to the data. Any account that defines data as material traces representing the state of the world at some place and time therefore counts as intrinsic on Leonelli's terms even though representation is technically a relation.

One could worry whether allowing a role for interpretation unmoors the objectivity of data. Indeed, Leonelli notes that she has encountered some confusion about the precise aims of her account. "One question I was asked over and over again while presenting this work to academic audiences was whether this research should be seen as a contribution to philosophy, science studies, or science itself. Is my assessment of data centrism intended to document the concerns of biologists involved…? Is it rather a critical approach, aiming to place biological practices in a broader political, social, and historical context? Or is it a normative account of what I view as the conceptual, material and social foundations of this phenomenon?" (8). I believe that any concerns about the objectivity of data in Leonelli's account are ultimately misplaced, because her goal is to understand how scientists value data as potential evidence rather than reducing the meanings of "data" and "evidence" to subjective desires or social interests.

Nonetheless, there is a valid question here about how exactly Leonelli is positioning her account in relation to what scientists do. Her assessment of data centrism is based on a philosophical "framework" she proposes "through which the current emphasis on data within the life sciences, and its implications for science as a whole, can be studied and understood" (1; also see 9, 77–84, 178). However, she doesn't directly address what a framework is and how it differs from the traditional sort of "theory of data" (198) that she sees as unhelpful. I would suggest this distinction is in fact crucial, both for understanding how Leonelli's views relate to existing philosophical work and for evaluating it on its own terms. It's not that she refuses to provide any general definition of data, because this is exactly what the relational account does. However, she does identify the necessity of interpreting her definition in a situated way as essential to her framework. Does that mean the necessity of interpretation is somehow conceptually implied by or equivalent to a relational view of data? Is it possible to disentangle the distinction of relational and intrinsic accounts from the implied distinction between frameworks and theories? The answer, I believe, amounts to separating method from content in the study of data.

This content downloaded from 129.219.008.139 on April 04, 2018 11:47:10 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

Copyright The Philosophy of Science Association 2018. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/697732

In order to clarify the relationship between these two distinctions, it will help to bring in material that is external to Leonelli's discussion but which falls within her broader aims and tradition. In this regard, I believe the work of Herbert Blumer, an important inheritor of Dewey's legacy, can provide useful insights. Blumer is best known for two major contributions he made to the social sciences between roughly 1930 and 1980 (Blumer 1969a; Tucker 1988; Becker 1988; Hammersley 2010). First, he delivered scathing methodological critiques of major attempts to introduce quantitative rigor and precision to sociological theory, including various forms of positivism (Blumer 1940; Blumer 1954). Second, he launched symbolic interactionism, a broad research program that stresses the essential importance of interpretation and socially situated action in explaining the course of social processes (Blumer 1969b). Together, these contributions cemented Blumer's place as a co-founder of contemporary qualitative social science and are still relevant as a bridge between pragmatist philosophers such as Dewey and George H. Mead and later methodological innovations in sociology like grounded theory.

The main idea I want to draw from Blumer is the "sensitizing concept," a methodological tool he created for generating empirically-grounded theory about social phenomena. Sensitizing concepts serve to help stabilize the semantics of scientific terms for social phenomena that involve people's interpretations of each other's actions. Sensitizing concepts provide reference standards for identifying and analyzing instances of social phenomena without attempting to make an end run around the role of the sociologist as an interpreter. For example, sociologists sometimes draw sensitizing concepts from the existing language of their subjects, such as one study that describes how government officials and unemployed workers both came to use the term "social junk" to describe people on welfare (van den Hoonard 1996). Alternatively, sociologists also coin new terms based on external metaphors or analogies that provide insight into the deeper meaning of people's experiences.

Precisely how sensitizing concepts stabilize the semantics of qualitative sociology deserves more analysis, but the basic outline is clear enough. Selecting folk terms as sensitizing concepts provides an initial semantic anchor for researchers' collective theoretical language because one can refer the meanings of the terms to the embodied understandings of the actors involved. Researchers who can pass as members of the communities under study will also have approximately equivalent, embodied understandings of the terms. Folk terms, however, are rarely as standardized and precise in their meanings as rigorous scientific reasoning demands. Researchers therefore analyze the meanings of sensitizing concepts in order to formulate more explicit definitions, which in turn suggest new hypotheses about the terms' relationships to the social situations in which they occur. This is why Blumer called them *sensitizing* concepts: one selects a term that provides an initial semantic grasp of actors' meanings and then analyzes instances of its use in order to gain a more explicit and systematic understanding of its situated content.

I think this nicely illuminates the intended methodology behind Leonelli's relational account. Her emphasis on interpretation as essential to understanding the meanings of "data" in biology reflects her commitment to starting with the situated understandings of scientists as actors. In offering an explicit and general definition of the term "data", her aim is to establish a

This content downloaded from 129.219.008.139 on April 04, 2018 11:47:10 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

Copyright The Philosophy of Science Association 2018. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/697732

collective project of studying how scientists ascribe meaning to material objects as potential evidence and respond to the interpretations produced by others. The metaphor of a data journey also serves as a sensitizing concept by helping us notice the transformations that the material products of research activities undergo as scientists make use of them in new situations.

One way to express the difference between a framework and a theory, then, is in terms of the function that definitions play within each. Within a theory, the definition of a term expresses a stipulated and fixed meaning that must be followed in order to use the theory correctly. To paraphrase Blumer, definitions in theories tell us precisely what to look at *in advance*. In contrast, a sensitizing concept functions as a *provisional* means of access to social phenomena, and defining a sensitizing concept serves to make explicit important aspects of a phenomenon for the sake of further exploration and analysis. To put it another way, theoretical terms are supposed to have fixed semantics in order to enable strict empirical testing of the theory, while terms in a framework are expected to have some degree of open-ended semantics in order to expand what we can perceive and help formulate new claims about the phenomena. A framework, then, is a conceptual apparatus for theorizing, but it is not a theory in the sense that philosophers of science typically have in mind. Making and using frameworks are important parts of *doing* theory about social phenomena that nonetheless cannot be expressed *as* a theory (Griesemer 2012).

How does this relate, then, to Leonelli's claim that data centrism does not call for a new, overarching epistemology of science? If the goal is to understand scientific change, then her point is that intrinsic theories of data are blind to the most important thing going on right now: the emergence of novel visions for how scientific research should be organized through the coordination of multiple interpretations about what counts as scientific data. A lingering question, though, is whether Leonelli's relational framework offers the best or only way to understand data centrism as a phenomenon.

Inside and outside the frame

Having arrived at a more explicit understanding of the contrast between frameworks and theories, we can return to the matter of evaluating Leonelli's relational account and method. What sorts of flaws are possible in a framework? Can a framework be incorrect? Any framework carries implicit normative content in virtue directing our attention toward some things and making others harder to notice. One problem in this regard is if a sensitizing concept fails to orient us toward noticing elements or aspects of a situation that are important for our aims. Blumer, for example, talked of being able to "test" sensitizing concepts empirically by examining whether the cases they identify form generalizable relationships with other theoretical entities of interest.

One of Leonelli's key conclusions, for example, is that "in its most productive forms, the implementation of data journeys involves the use of computational tools to raise awareness of the conceptual, material, and institutional scaffolding required to package and interpret data, rather than hiding those aspects away" (171). However, the intended meaning of productivity is

This content downloaded from 129.219.008.139 on April 04, 2018 11:47:10 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

Copyright The Philosophy of Science Association 2018. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/697732

somewhat ambiguous: productive for whom, and in what way? Note that people working with intrinsic and relational understandings of data centrism may interpret the same facts about data journeys to have very different practical significance. One person might look at all the work needed to recontextualize a data package across multiple situations and conclude that this shows the value of a pluralist approach. Another person might look at the same work and conclude that we need a new set of more general and robust standards in order to reduce subjective variation in the dataset's use.

In this vein, one could argue that a number of leading data integration initiatives have succeeded precisely because they try to minimize support for explicit data interpretation in their computational infrastructure. The "realist methodology" of the Open Biomedical Ontology movement is one example (Smith and Ceusters 2010), and a second is the decision of biodiversity data aggregators, such as the Global Biodiversity Information Facility, to ignore expert disagreement about the meaning of taxonomic names (Franz and Sterner 2017). Leonelli tends to place greater weight on the merits of customizability. For example, "depending on the degree of reflexivity with which they are developed and applied, data-centric approaches can stimulate an increase in the accountability and rigor of research practices or instead be used to black-box data packaging activities and thus diminish critical participation" (174). There are important drawbacks to critical participation, especially in large and decentralized scientific projects, however. Depending on the depth and breadth of disagreements among stakeholders, for example, rigorous reflexivity can massively increase the amount of meta-level work necessary to agree on how to get any 'actual' work done (Gerson 2008). Allowing customizable interfaces to databases can also impose heavy costs on computational efficiency and require additional iterations of data curation to support alternative perspectives.

In more philosophical terms, we might worry that there are dialectical possibilities for intrinsic accounts of data that Leonelli has not yet fully ruled out. Perhaps the problem with existing intrinsic theories is that they aim to be universal in scope, making it hard for them to accommodate and explain data centrism. A contextualized intrinsic approach might then have a better shot at competing with Leonelli's relational account in specific scientific situations. Smith and Ceuster's realist methodology may be a genuine competitor in this regard. In a series of papers in the 2000s, they described a variety of cases where allowing ontologies to include terms that refer to mental constructs like conceptual meanings led to serious logical reasoning errors (e.g. Ceusters, Smith, and Goldberg 2005).

As an alternative, they developed an approach that starts by settling on a consensus view about the set of real things and their determinable physical properties in a domain. This is effectively a meta-theory about how to design classificatory theories for the sake of computer-assisted data integration. One could also read their approach as providing an intrinsic theory of what counts as genuine data that is contextualized to consensus scientific languages for specific domains of phenomena. Of course, not every area in science has a consensus language for describing reality, and surely scientists in those fields still use and make data! This point hints at the possibility of a synthetic perspective on data that would identify and explain the relative strengths and limitations of intrinsic and relational conceptions.

This content downloaded from 129.219.008.139 on April 04, 2018 11:47:10 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

Copyright The Philosophy of Science Association 2018. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/697732

The arguments I've given here are different ways of getting at the idea that Leonelli's framework, based on tracking data journeys, may in fact be more general and powerful if it is not predicated on taking a universally relational view of data. There is no question that the work of following data journeys across science provides an important counterbalance to views that consistently undervalue human interpretation in scientific practice. Still, the details of data journeys are not sufficient in themselves to settle the broader normative questions that are at stake with data centrism, and this is where I think it would help to bring the stances people take toward the value of human interpretation explicitly inside the framework.[2] If what we care about is understanding and explaining data journeys, then surely it is important that some actors hold (and will continue to hold) intrinsic conceptions of data while others hold relational conceptions. Simultaneously pursuing a general argument that we should all move to a relational view seems in tension with accurately accommodating the fact that this is not what many scientists mean by "data."

Normative arguments about the best way to design data infrastructure are most forceful when they are attuned to the specific mix of aims and social organization of the relevant communities. As a result, I would suggest that a framework for studying data centrism will be most valuable when it explicitly orients us to discovering circumstances where there is room for compromise or productive synthesis between people with intrinsic and relational conceptions of data. The intrinsic-relational distinction could then operate *inside* the framework rather than being used to position the framework as somehow *opposed* to intrinsic views of data. In other words, we can follow a methodological approach to philosophical theorizing based on analyzing the situated meanings of human actors without also requiring that specific terms such as "data" can *only* be defined as essentially involving human interpretation.

References

Becker, Howard S. 1988. "Herbert Blumer's Conceptual Impact." *Symbolic Interaction* 11 (1): 13–21.

Blumer, Herbert. 1940. "The Problem of the Concept in Social Psychology." *American Journal of Sociology* 45 (5): 707–19.

Blumer, Herbert. 1954. "What Is Wrong with Social Theory?" *American Sociological Review* 19 (1): 3–10.

Blumer, Herbert. 1969a. *Symbolic Interactionism*. Los Angeles: University of California Press.

Blumer, Herbert. 1969b. "The Methodological Position of Symbolic Interactionism." In *Symbolic Interactionism*, 1–60. Los Angeles: University of California Press.

Ceusters, Werner, Barry Smith, and Louis J Goldberg. 2005. "A Terminological and Ontological Analysis of the NCI Thesaurus." *Methods of Information in Medicine* 44 (4): 498–507.

Franz, Nico M, and Beckett W Sterner. 2018. "To Increase Trust, Change the Social Design

[2] My argument here parallels Matthew Sample's critique of the concept of scientific repertoires that Leonelli and Rachel Ankeny have developed (Sample 2017).

This content downloaded from 129.219.008.139 on April 04, 2018 11:47:10 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

Copyright The Philosophy of Science Association 2018. Preprint (not copyedited or formatted).
Please use DOI when citing or quoting. DOI: 10.1086/697732

Behind Aggregated Biodiversity Data." *Database* doi: 10.1093/database/bax100

Gerson, Elihu M. 2008. "Reach, Bracket, and the Limits of Rationalized Coordination: Some Challenges for CSCW." In *Resources, Co-Evolution and Artifacts: Theory in CSCW*, edited by Mark S Ackerman, Christine A Halverson, Thomas Erickson, and Wendy A Kellogg, 193–220. London: Springer.

Griesemer, James R. 2012. "Formalization and the Meaning of 'Theory' in the Inexact Biological Sciences." *Biological Theory* 7 (4): 298–310.

Hammersley, Martyn. 2010. "The Case of the Disappearing Dilemma: Herbert Blumer on Sociological Method." *History of the Human Sciences* 23 (5): 70–90.

Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. New York: Houghton Mifflin Harcourt.

Sample, Matthew. 2017. "Silent Performances: Are 'Repertoires' Really Post-Kuhnian?" *Studies in History and Philosophy of Science Part A* 61: 51–56.

Smith, Barry, and Werner Ceusters. 2010. "Ontological Realism: a Methodology for Coordinated Evolution of Scientific Ontologies." *Applied Ontology* 5 (3-4): 139–88.

Tucker, Charles W. 1988. "Herbert Blumer: A Pilgrimage with Pragmatism." *Symbolic Interaction* 11 (1): 99–124.

van den Hoonard, Will C. 1996. *Working with Sensitizing Concepts*. Thousand Oaks, CA: SAGE Publications.

Wittgenstein, Ludwig. 2010. *Philosophical Investigations*. Edited by P. M. S. Hacker and Joachim Schulte. 4 ed. Oxford: Wiley-Blackwell.

This content downloaded from 129.219.008.139 on April 04, 2018 11:47:10 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).