Natures of Data: A Discussion Between Biology, History, and Philosophy of Science and Art. Philipp Fischer, Gabrielle Grammelsberger, Christoph Hoffman, Hans-Jorg Rheinberger, and Hannes Rickli. Diaphenes, 2020. 156 pages.

Adam Katz
Quinnipiac University

The existence of data is not something new, but what is new is that inquiry now tends to start with pre-existing databases, among which it must find its way and generate its questions, rather than subordinate data to hypothesis generation. This means that how data has been gathered, treated and handled, what Sabina Leonelli (2016) called its "journey," or the process of curatorship introduced between the "original" sources of data and any use, must now be accounted for in the inquiry itself. (Indeed, the curation of data is becoming a field of inquiry in its own right.)

The digital humanities have now been around for a couple of decades, but we are not yet speaking of student writing and other artifacts of student learning in any consistent way as "data." We will eventually have to because demands for quantified assessment cannot be met with a refusal to assess, but with a different model of assessment and therefore of data. The frustrations and absurdities of trying to reduce student work, in accord with a rubric, to numerical values should be apparent to anyone who has ever worked on assessment. Only qualitative assessment, involving detailed examination of student writing produced under very precise conditions, can teach us something about our pedagogies. This approach to assessment seems to cut against the grain of the now ubiquitous "Big Data" approach. Rather, it should be seen as the quintessential site of critical thinking, where the aesthetic intersects with the scientific.

The book presently under review emerged from a discussion group "formed in response to Hannes Rickli's first project, *Surplus: Videograms of Experimentation* (2007–2009), which reflected artistically and theoretically on a bundle of analog video recordings that were unsystematically compiled from laboratory contexts starting in the 1990s, presenting them as visual and simultaneously physical traces of the production of scientific facts" (p. 11). In the journey from analog to digital, it became clear that "the processes of research were withdrawing ever more from the human senses, thus making direct observation difficult and sometimes impossible" (p. 12). Rickli's project was to work against the grain of this vanishing of the perceptual and sensory in inquiry:

> The work of biology consists in isolating phenomena in confusing environments, applying complex methods of abstraction with digital programs in black boxes in order to cleanse them of the effects of particular and singular elements, and allowing them to be transported as distinct numbers or formulas. The work of art runs in the other direction: It is interested in the material conditions of this abstraction, reconstructing the

concrete circumstances in which biological data are collected, distributed and calculated. As such, it shifts the focus to the time spent, the concrete spaces and lighting conditions, the scientists' gestures, and, last, but not least, the animals at the focus of the biological knowledge interest. (p. 12)

It is in these contrasting directions of art and science that we might find a pedagogy focused on critical thinking. We want to isolate and abstract practices of reading, writing and thinking that we could identify, compare and rehearse across contexts and that might thereby appear "cleanse[d] . . . of the effects of particular and singular elements." We could have no way of speaking about pedagogy otherwise. But we also want to retain interest in and attention to the "material conditions," the "time spent," the "concrete spaces" and the "gestures" of students and teachers alike. In fact, it is in those gestures that we will find markers of learning when a given practice, inculcated in the controlled space of the classroom, interrupts, and in turn is disrupted by, the varied habits of literacy students bring with them to the classroom (and which prompt us to desire more data that might inform our pedagogical practices). We learn less from identifying "skills" students can or cannot perform than we do from hypothesizing, from the interference in the discourse students bring into the classroom generated in the course of novel, complex performances, what students take themselves to be attempting. In that case, we must be able to frame the abstract practices we seek to turn into habits in terms of the conditions, spaces, times and gestures that necessarily mark any performance. The more precisely we bring into view the interference of existing habits with new practices, the broader and more useful will be our speculations regarding the wider institutional and even social conditions informing events in the classroom.

The conversations in *Natures of Data* are divided into four sections, which we could see in terms of the increasing distance of data from direct observation and perception. First is "Data," which centers on the progressive intervention of more technologically advanced instruments of measurement, leading to more complex ways of thinking about data—for example, the precise way of registering and articulating sonic as opposed to visual recordings of fish. Second is "Software," which directs attention to the design of algorithms so as to remove humans from decisions regarding the recording of data while resituating human decision in the decisions made regarding the design of algorithms (while, of course, through machine learning, algorithms come to design themselves and other algorithms). Third is "Infrastructure," which brings into focus the immense, expensive, energy-consuming and distributed nature of the machinery required to make data increasingly invisible and mediated. And, finally is "*In Silico*," a term coined here to refer to a growing tendency to supplement empirical experimentation with the testing of hypotheses by drawing upon the vast stores of data in new ways so that the very formation of hypotheses can be seen to be heavily reliant on the means by which data is collected, stored and labeled. Throughout, the question of whether the object being studied eventually disappears in all this mediation—all participants insist the answer is "no"—is foregrounded.

The book has striking illustrations from Rickli's work, some of infrastructural apparatuses but mostly various visual representations of data. His interest in representing the paradox of increasingly present and imposing infrastructure and the never complete vanishing of the original sources of data doesn't get represented here (as far as I can tell), but there is the following description of one of his installations by Gabriele Gramelsberger:

What I found so beautiful about Hannes's panorama is that one sees how the fish is wrapped in something else, and this something else, for its part, is wrapped in infrastructure, and this infrastructure is wrapped in an even bigger infrastructure. Right, the wrapping. A tremendous number of calculations are running on the computer, but when the drilling tower stands still and the energy supply collapses, the cooling stops working. Then the computer beaks down, and the data may be lost, meaning that the research is lost: it dissolves into oblivion. The dependency becomes just as clear through Philipp's power outlet. Normally one does not think about it all here. This stupid seawater is so corrosive, it doesn't do what we want, while mankind generally has pretty much everything else under control. (p. 107)

We can see how the extension of the infrastructure into "everything else" creates vulnerabilities that not only threaten the so meticulously and expensively constructed means of collecting and analyzing data in ever more precise and variegated ways while demonstrating the dependence of experimentation on physical conditions that cannot be controlled but also, even though it is not explicitly noted here, the social conditions that allow for the sustained investments enabling such research. The fish (what, after all, we are interested in in the first place) almost disappears within the infrastructure while the infrastructure is itself in danger of a much more catastrophic disappearance.

If we think analogically, we can use this model to engage the metastasizing structures of assessment and oversight of our pedagogical practices—assessment and oversight which are surely in large part responsible for our pressing into service concepts like "critical thinking," which remains dependent upon instructors' assessment of what students aren't doing rather than what they are doing. We have our "data": student work, which we use to measure, assess and record student "capacities," according to whatever criteria we have established. One detail in a student paper will be important, and in a particular way, while other details, or patterns, will be less so, or in different ways. A collection of papers from a class is data used to address another set of questions regarding the efficacy of pedagogical practices or departmental agendas and policies. How do we determine the collection and use of such data?

Then we have "software," which we can think of as our assignments and feedback (which presumably refers back to, and is therefore an iteration of, the assignment). The assignment produces practices we can recognize as meaningful and thereby generates the data and brings into focus our own role in eliciting practices from the students. The data collected only makes sense in terms of the "software"—what students do is meaningful only in terms of what they are being asked to do. An assignment that does not generate the conditions for feedback, that is, for further data production, is more malware than software—what purpose could be served by asking students to do something that the instructor couldn't comment on in such a way as to further the transformation of the intellectual habit the assignment targets? An interesting paradox emerges here: one could imagine, on this model, the increasing standardization of assignments and feedbacks—who knows, perhaps even their automation—but, in fact, because of the uncontrollable "stupid" and "corrosive" learning conditions students have traveled through, the intersection of the students in a given class in a given institution with a given set of assignments designed by a

given structure is more impervious to reduction to computation the more precisely it follows permutations of concrete habits as they resist the demands of the more impactful practices of literacy. In other words, the "artistic" sense of the embeddedness of data in concrete conditions (even if conditions we can only reconstruct hypothetically, through the "metadata" through which we organize the data we produce) is becoming more essential for critical pedagogies of literacy.

And how problematic our own infrastructures have become! We face an ever growing array of formally institutionalized imperatives driven by economic and political forces and pressures on an increasingly expensive and necessary higher education to account for itself. It makes sense that universities, and therefore each of its individual units, are asked to explain, and justify with "data," what, exactly, they do for students. And, yet, there is no obvious reason to assume that, for example, the often amply infrastructured demands for equity will be in all cases completely consistent with demands to demonstrate that learning outcomes are optimal across the board. Certainly, no one has shown this to be the case or organized all of these institutional imperatives so as to make it the case. The often imperious and not necessarily well-informed demands of instructors in the disciplines that students come prepared to do the writing they would like to see in their classes, so they don't have to "teach writing," would already produce a sufficiently daunting infrastructure, eliciting from writing instructors challenging efforts at translation.

Like Rickli's fish wrapped within layer after layer of superstructure so that the fish itself is barely heard from, the instructor struggles to enter some carefully designed set of pedagogical practices into the data flow. So, we may have recourse to our own "*in silico*" practices, turning our classes into laboratories designed so as to enable us to hypothesize, along with students, and in ways that might never meet the demand for standardization required for large scale empirical testing, the virtually infinite ways in which habits of literacy confront concentrated practices of literacy. We can make less and less more and more meaningful, so that (I'm exaggerating a little) the composition of a single sentence might reveal data of multiple histories of literacy. Since such *in silico* work is resistant to conventional forms of "rubricization," being irreducibly singular except insofar as others might take the example and design practices aimed at generating the auto-intelligibility of those practices for practitioners insofar as they become practitioners, the data they produce must be carefully treated, collected, measured, and made ready to travel along unanticipated routes.

Reference

Leonelli, S. (2016). *Data-centric biology: A philosophical study*. University of Chicago Press.