Hallam Stevens: Life Out of Sequence: A Data-Driven History of Bioinformatics. University of Chicago Press: Chicago, Illinois, 2013. 294 pages.

Life Out of Sequence investigates intersections of biology, physics, and computer science to offer an account of the historically recent emergence of bioinformatics as a scientific discipline. Stevens draws from his field work at the Broad Institute in Cambridge, Massachusetts as well as interviews and archival research to investigate the dynamic relationship between biology and computing technologies, both the epistemological space which computers responded to, and how knowledge paradigms shifted once computers began to be integrated in the laboratory. Stevens makes it clear from the outset that he is not offering a technologically determinist analysis of computers in biology. His argument, rather, focuses less on the machines that go "ding" and more on the types of research questions and knowledge production mechanisms that these machines both afford and constrain. He looks at how biology *shaped* and is *shaped by* computing technologies.

Life Out of Sequence is organised by a concern for the movements of different types of objects (including data, laboratory workers, "wet ware," and other laboratory technologies) and the spaces through which these objects move (physical and virtual). This creates an engaging organisation that mirrors how knowledge circulates and is produced and reproduced in these spaces. Oscillating between ethnographic accounts and archival research, we learn about the physical organisation of laboratories,

especially at the Broad Institute. Readers see how this physical organisation of laboratory space reproduces divisions of labour, centralising and privileging some types of work as "real science" and casting others as more menial tasks. Similarly, we learn about ongoing difficulties in balancing the need for bigger experiments, bigger laboratories, more interdisciplinary teams, and the perceived need to defend traditional knowledge making forms. This physical organisation and growing needs of the biological laboratory translates into interdisciplinary tensions, where traditional biologists seem to carry a burden to defend their traditional forms of knowledge production as "real science." As a result, the "real scientists," we are told, tend to control knowledge production in these interdisciplinary laboratories. The tension within interdisciplinary teams goes beyond interpersonal communication; it stems from differences in what is viewed as legitimate means of knowledge production. To illustrate the stark differences in knowledge production, Stevens offers us a compelling direct comparison of two projects interested in alternative splicing: one conducted by a biologist, and the other a computer scientist. In two short anecdotes telling of the work of graduate students in these intersecting fields, and we learn how they would proceed with their investigations. The difference between the two visions of knowledge production lies in their approach to data: the biologist is more concerned with "wet ware" and specific

cases, whereas the computer scientist looks for ways to crunch as much data as possible, as quickly as possible.

In describing how spatial organisation relates to this restructuring of biological knowledge production, Stevens focuses on ethnographic field notes taken at the Broad Institute in Cambridge. He uses front/back and central/periphery orientations to describe how laboratory space at the Broad is organised in order to present an image of "real" biology. In addition, this configuration of scientific labour brings in issues of control and surveillance. Using before-and-after diagrams and scenarios, readers are shown how the laboratories at the Broad were run using notions of lean management, which was borrowed from industrial management. This type of management places values on speed and efficiency, much like factory production lines. In this model, teamwork and productivity are favoured over individual intelligence and innovation.

Following this description of the order of physical laboratory space, three chapters offer a description of the organisation of virtual spaces. Stevens' primary critique in this section focuses on the pipeline metaphor of computation, which presents the movement of information as passive flows from genome to hard drive. This pipeline metaphor, he argues, glosses over the effects of human choice in informational systems, tools, annotations, and gene ontologies, "flattening" messy data into "universal" data. More specifically, the movement of data into virtual spaces creates a linguistic problem; ontologies applied to this data create a particular way of viewing biology, and constraining ways of talking and acting within the biological sciences. A change in the language used to describe the data results in a change in what one can do with the data.

To historicise the development of genomic databases, Stevens offers a side-by-side discussion of Margaret O. Dayhoff's development of the *Atlas of Protein Sequence and Structure* (Dayhoff & Richard, 1968) and Walter Goad's collection efforts at Los Alamos. Dayhoff and Goad were the two primary candidates for the creating a genomic database for the National Institute of Health (NIH), with Goad ultimately receiving the funding. Stevens offers a compelling argument on how Dayhoff may have been less favoured than Goad because her efforts were understood as "mere collection and compilation" and not as a real contribution to the systematisation of biological thought. Goad's GenBank, on the other hand, proposed a "flat file" structure for the database that appealed greatly to the NIH. Using the GenBank flat file as an example, Stevens ultimately argues in these chapters that the flattening of biological information into easily transportable entities not only obscures human judgement that is part of the digitisation process, but also creates a particular landscape for particular kinds of biological action. In short, the movement of biological information pre-determines what is considered legitimate biological knowledge. Stevens does not, however, go as far as explaining exactly how gene ontologies and gene databases structure biological knowledge.

Using ethnographic and archival research of physical and virtual lab spaces, Stevens offers a way of seeing computers in the laboratory as they directly influence organisation, labour, surveillance, data collection, and knowledge production in the name of biology. Throughout this text, Stevens explains the technical concepts necessary to follow the arguments he puts forth, making the text very accessible for readers with different levels of familiarity with bioinformatics. However, despite

Stevens' clear and explicitly stated intentions of not following a technologically determinist thinking, Stevens seems to uncritically distinguish between "the digital" and "the biological," or "data" and "nature." To a certain degree, this distinction seems necessary in order to support Stevens' arguments, however he does not provide any lengthy account of how this distinction is made. This single criticism aside, Stevens follows in the tradition of Latour and Woolgar's (1979) *Laboratory Life* to bring an updated account of the circulation of knowledge in biological spaces. Stevens' work provides a compelling and insightful analysis of the changing role of data in biology, and shifting ways of knowing with increasingly interdisciplinary work centred around computing technologies. It is a crucial new read for STS researchers interested in engaging in interdisciplinary research on emerging science and technologies.

References

Dayhoff, M.O. & V.E. Richard (1968) Atlas of Protein Sequence and Structure, 1967-68 (Silver Spring, MD: National Biomedical Research Foundation).

Latour, B. & S. Woolgar (1979) Laboratory Life. (London and Beverley Hills, CA: Sage).

Molly Hartzog
Communication, Rhetoric, and Digital Media
Genetic Engineering and Society
North Carolina State University
Ricks Hall, Box #7579
Raleigh, NC 27695
mshartzo@ncsu.edu