

Grant Ramsey and Andreas De Block (eds), *The Dynamics of Science: Computational Frontiers in History and Philosophy of Science*

Pittsburgh: University of Pittsburgh Press, 2022. Pp. 308. ISBN 978-0-8229-4737-0. \$60.00 (hardcover).

Tim Boon

Science Museum, London

The use of computational methods in historical practice dates back as far as the *Annales* school, although it has always been a minority pursuit. But with the astonishing and recent growth of large language models and the statistical techniques of machine learning that no one can have missed, the question arises of how historical practice may be enhanced by their application. Across the humanities, the impact of AIs has been most pronounced in text-mining studies in English literature. In history and philosophy of science (HPS), take-up has, until the volume under review, been slight. Reasons for our discipline to fight shy of these methods may include scepticism about the anachronistic application of scientific methods to historical reality or aversion to the combination of grunt work on data and the current necessity to learn highly technical, mathematical, procedures. They may also include unwillingness to depart from the methodologies with which we have become comfortable.

The potential of computational methods is essentially to work at scale on huge numbers of machine-readable digitized texts. As the editors write, ‘each researcher can examine but a single drop in the ocean of [scientific] literature – their studies may involve depth, but they lack breadth’ (p. 8). But we always need techniques to help us to see the ‘Big Picture’, a concern of our discipline since the special issue of this journal with that title thirty years ago. Techniques that allow us to address large corpora of scientific and technological writing can act as a foil to the more familiar tight case studies and close reading; that is, they may offer a new route to valid generalization.

The Dynamics of Science is far from being a ‘how-to’ volume, but is openly a work of advocacy, inviting members of our disciplines to dive into the potential of new ways of doing HPS using large datasets and computational techniques. I come to it through working on Congruence Engine, an Arts and Humanities Research Council project using machine-learning techniques to link collections to enable historical research. As a historian of science working with these new techniques, I have good motivation to engage with the book’s subject matter. And so I am pleased to report that this is a valuable volume that has enhanced my understanding at the same time as it reveals a new field full of potential, albeit one at an inchoate stage.

The book is diverse in its contents, which are well set out in a helpful introduction. It even includes a chapter on good data management! Most of the individual chapters, the majority of which – true to the field – have multiple authors, have a philosophical rather than a historical bent. This is explicable given the potential to, for example, evaluate patterns in the ‘evolution’ of science. This application of the Darwinian metaphor to scientific texts is exemplified in the chapter by David Chavalarias *et al.* advocating phylomemias, a metaphorical extension of phylogeny in

evolutionary biology to study the filiation of words within scientific texts as a means to model scientific change.

Reviews in other journals will address the potential for philosophical insight. In these few words I will pursue the interest in the volume as a primer for historians curious about these new techniques. To maximize comprehensibility to researchers new to the approach, it is possible to take a graduated route through the essays. After the introduction, an ideal place to continue is Krist Vaesen's piece on topic modelling (a prominent technique within the volume's coverage) and supervised machine learning, because he takes particular care to make his account readily comprehensible. The chapter by Christophe Malaterre *et al.* is a good next stepping stone; it uses the same technique to examine the presence, then absence and reappearance of socially engaged themes within a an eighty-year corpus of the journal *Philosophy of Science*. Another good example is Charles H. Pence's data-driven analysis of the biometry–Mendelism debate, using the Named Entity Recognition technique on the 1,622 relevant articles published in *Nature* magazine. He shows that segmenting the network of discourse into six five-year periods reveals fascinating shifts of affiliation between the participants.

The editors warn in the introduction (p. 7) that 'many historians and philosophers simply lack sufficient training in mathematics and computer science for conducting computational HPS on their own'. Colin Allen and Jaime Murdock's chapter on topic modelling for HPS takes a case study to map the divergence of the text of Darwin's *Origin* from most, but not all, of his reading. Thereby they congenially propose to move away from word-centred towards document- and context-centred approaches to applying these techniques to the history of science. This reveals how, at this early stage, techniques are being imaginatively flexed to gain better insights. All the same, some of the paragraphs here do require an understanding of statistics that cannot readily be achieved by glancing at Wikipedia.

This kind of problem would seem to be at the heart of the dilemma of the current moment of the exploration of newly available textual corpora and the application of machine learning to history of science. It is our job to understand technicalities in the historical record, but we are not so much used to applying technicalities in our interpretive armoury. HPS is by nature an interdiscipline, and we pride ourselves on the catholicity of the disciplines we embrace. Computational HPS involves a new kind of interdisciplinarity that, in many cases, and for the moment, can only be achieved by multidisciplinary working, not so often by the uniting in single scholars of the wherewithal to conduct the solitary research that is still the dominant mode in our field. And yet this may turn out to be only a temporary state of affairs. There has been something of a Cambrian explosion of tools and techniques in the last few years. There will doubtless come a time when the field is less experimental, and user-friendly programs enable more of us to incorporate digital methods into the roster of techniques that we apply every day in our research. I hope so; the potential is vast, and there is no better place to witness it than the volume under review here.

doi:10.1017/S0007087423000493