Studies in History and Philosophy of Biological and Biomedical Sciences 55 (2016) 41–44

Contents lists available at ScienceDirect

Studies in History and Philosophy of Biological and Biomedical Sciences

journal homepage: www.elsevier.com/locate/shpsc

Introduction

Towards future archives and historiographies of 'big biology'

Christine Aicardi [a], Miguel García-Sancho [b]

[a] Department of Social Science, Health and Medicine, King's College London, Strand, London WC2R 2LS, United Kingdom
[b] Science, Technology and Innovation Studies, University of Edinburgh, Old Surgeons' Hall, High School Yards, Edinburgh EH1 1LZ, United Kingdom

When citing this paper, please use the full journal title *Studies in History and Philosophy of Biological and Biomedical Sciences*

The 20th-century has seen the progressive rise of 'big science', especially after 1945. In the last seventy years, research has increasingly been conducted by large, interdisciplinary teams spanning different countries and institutions, and attracting generous funding by both public and private actors (Galison & Hevly, 1992). The first paradigmatic example of this way of doing science was the Manhattan Project to build the atomic bomb that eventually ended World War II. After this, biological research gradually grew and acquired the dimensions of 'big science', to the extent that in the 1990s the Human Genome Project was called 'the Manhattan Project of biomedicine' (Kevles, 1997; Lenoir & Hays, 2000). Historians have been rightly sceptical about these alleged shifts in scale, pointing to the coexistence of both large and small-scale models of doing biology throughout the 20th century (Bud, 2016; Gaudillière, 2009; Pestre, 2003). However, recent biological research presents historically-specific contours that require attention at both the level of sources and of scholarly narration (Aronova, Baker, & Oreskes, 2010; Davies, Frow, & Leonelli, 2013; Hilgartner, 2013). In the life sciences, the rise of large-scale models has been driven less by the cost of gigantic experimental infrastructure than in the physical sciences. Rather, the driver here has been the complexity of biological systems, the study of which requires a diversity of practices, theoretical perspectives and experimental methods, at many different levels from the molecule to the organism.

This has resulted in the rise of the scale and disciplinary scope of biological research projects, involving large numbers of very diverse participants and generating huge masses of data. How to document these projects and write their history is a source of preoccupation for both historians and archivists. Humanities are being pushed towards scaling up small data to become a big data science, as data-driven research seems to hold promise for managing overabundant and heterogeneous evidence, especially within biology (Leonelli, 2012; Schreibman, Siemens, & Unsworth, 2004). Yet in the humanities, as in the social and natural sciences, critiques challenge data-driven strategies and warn against theoretical deficit (Carandini, 2015; de Chadarevian, 2009; Fisher, 2015; Scheinfeldt, 2012). The authors contributing to this special issue all took part in a symposium jointly organised in October 2013 by the Wellcome Library and the Department of Science and Technology Studies at University College London, entitled 'Making the History of the Postwar Life Sciences', the goal of which was to explore whether it is desirable to transform biomedical history into a data-driven endeavour and the alternatives to this course of action. The special issue came out of the desire to further develop some of the ideas that were fruitfully discussed during the meeting.

In their essays, the authors set out to explore the connections and interplay between historiographical and archival issues raised by the contemporary transformation of the life sciences into big science enterprises. Reflecting on narrative models, the nature and availability of sources, and the construction of archives, they challenge overly simplistic 'big data' strategies and propose a number of alternative methods for navigating 'big biology'. Despite the diversity of their objects and perspectives, the essays draw out a number of salient cross-cutting themes relevant to this overarching goal.

1. New sources for broader scientific narrations

The authors all highlight the importance of opening up the scope and remit of the scientific archive—of what counts as sources—for achieving a more inclusive definition of what 'doing science' encompasses. Broadening as an overall strategy may sound paradoxical coming from a joint effort that proposes alternatives to 'big data'. However, for reconstructing complex pictures, a selective focus on diverse small pieces may be more fruitful than the indiscriminate accumulation of masses of similar data (Secord, 1993;

E-mail addresses: christine.aicardi@kcl.ac.uk (C. Aicardi), Miguel.gsancho@ed.ac.uk (M. García-Sancho).

http://dx.doi.org/10.1016/j.shpsc.2015.09.009
1369-8486/

Subrahmanyam, 1997). In addition to questioning the adequacy of 'big science' as a specific post-World War II category, historians have criticised the view of recent science as characterised by an 'explosion of information.' Earlier periods also produced an overabundance of scientific records, many of which could not be indefinitely stored (Hughes, 1997; Müller-Wille & Charmantier, 2012). This suggests that the role of historians, and of humanities researchers more generally, may be working with a limited but meaningful set of evidence and focus their scholarship on the range and richness of the narratives that can be extracted from these records rather than on their number.

In this special issue, Susan Lindee addresses the problematic transition between microhistories and more comprehensive accounts to track the development of post-1945 human genetics and genomics. Her essay shows how microhistories can become the sources that build a scaffold for bigger picture narratives. She proposes risk as a category that emerged from her previous case studies and which allows the creation of a unified account of genetic research from the atomic bomb to the Human Genome Project. This account does not derive from the accumulation of 'big data', but rather from looking at pre-existing and manageable datasets with new eyes.

Miguel García-Sancho and Christine Aicardi address the problem of extracting meaning from historical datasets. Their essays propose narrative models for finding selective points of entry into big biology projects and navigating the multiple and dispersed records they produce. García-Sancho uses the administrative archives of big science projects as alternatives to individual scientists' papers. He singles out "the synthetic voice of the invisible administrator", as an actor who brokers between different epistemic communities and enables historians to harmonise disparate accounts. Using the life of the late Francis Crick as case study, Aicardi focuses on his role as an influencer, arguing that this was an integral part of his way of 'doing science'. She suggests that "'following the cross-worlds influencers' may be a fruitful heuristic for historians probing the rhizomic and genealogic entanglements of modern big bioscience."

Different sources and points of entry into science result in different understandings of scientific practice. Norberto Serpente argues against a narrow conception of what doing science entails and proposes an alternative approach where pedagogy and experimentation are not segregated. This broader understanding leads him to select new sources, historicising and documenting molecular images as vehicles of knowledge production in textbooks. Serpente shows that for historians of ideas, there is value in interrogating such vehicles for the production of scientific knowledge, as well as in taking into account multimedia archives that include images as well as text. These archives allow looking at science as an artisanal endeavour that is disseminated by artists in cooperation with scientists. Soraya de Chadarevian reflects more broadly on recent changes in scientific practice, historiographical trends and archival strategies and evaluates the possible place of diverse categories of sources—paper records, digital files, material artifacts and oral histories—in the archives of contemporary life sciences.

Where de Chadarevian's reflections are driven by her historian's experience, Jenny Shaw brings an archivist's perspective to the same question and shows how genomics as big collaborative science is challenging traditional archival theory. She argues for a move away from traditional archival approaches to science (generally focused on retired famous scientists) towards trying to identify and capture records of significance both at the daily routine level of scientific work and beyond the purely scientific sphere. Sara Peres broadens further the notion of what may count as an archive. Focussing on genebanking and how it has been envisioned as a strategy for 'genetic conservation' of plants, she defends the view that it is analogous to archiving, since it enables the preservation of diverse genotypes, embedded in seeds, for future use.

Opening up the scope of sources and archives to novel media raises many issues. De Chadarevian and Shaw both address the problems posed by digital archives: problems which are not specific to contemporary science but rather to contemporary archives and which require new strategies for the curation and conservation of digital material. They also consider the issue of the place of objects—things—in archives, which brings into view the increasing overlap between the concerns of archivists and curators (see also Robert Bud's commentary to this special issue). In Peres's essay, which analyses a different kind of material repository, the blurring between roles converges to the point that it becomes identification. She shows that the plant biologists who envisioned and designed the seed-storing genebanks that she has studied gradually became archivists.

Another category of sources discussed in the essays is that of oral histories and interviews. The value they present for contemporary history is implicitly accepted in the essays written by Aicardi, Lindee, Peres, and Serpente, whose historical studies have all involved interviews. De Chadarevian and García-Sancho reflexively interrogate the methodological issues that contemporary historians face when combining oral histories with conventional archival material, as well as the benefits this may bring.

Ironically, broadening the scope of the sources, actors and activities used for documenting contemporary 'big biology' can also underline how archives reveal many things but hide others, and more generally, the limits and constraints facing both historians and archivists. The historian is at the mercy of external factors, shaping what it is possible to research, which has a huge impact on historiography. Some factors relate to the tacit norms of what is socially and professionally acceptable to record in particular settings. For instance, Lindee's essay highlights the absence, in archival records, of emotionally charged interactions between clinicians and patients. Other factors are linked not to the creation of records but to their availability. The way in which García-Sancho approaches the limitations of individual collections amounts to political work: locating 'black zones' in terms of sources and working to open them up, or at least contour them. The hoped-for result is a greater plurality of sources, which implies plurality of historiography. There are, however, archival sections that will never open and de Chadarevian suggests that the expansion of secrecy motivated by private commercial interests and military classification could be a key problem in the historiography of recent bioscience. Confronting this issue from an archival perspective, Shaw suggests that the approach adopted for the Human Genome Archive Project could be appropriate for identifying and preserving records, even when their owners are unable or unwilling to deposit them in archives. However, these may still be subject to access restrictions, which themselves may affect research.

2. The informational worldview and genetics as history

The proliferation of born-digital sources[1] relate to dematerialisation and more broadly, to the rise since World War II of a pervasive informational worldview across all life sciences (Fox Keller, 1995, 2002; Kay, 2000). Taken together, the essays of de Chadarevian, Peres and Shaw demonstrate that framing the life sciences within an informational worldview has made possible specific forms of archiving while preventing others. Serpente reacts

[1] Born-digital sources are sources which original form is digital, in contrast to digitized sources.