Feature Endeavour Vol. 37 No. 3

Computerizing natural history collections

Mary E. Sunderland*

University of California, Office for History of Science and Technology, 543 Stephens Hall #2350, Berkeley, CA 94720-2350, USA

Computers are ubiquitous in the life sciences and are associated with many of the practical and conceptual changes that characterize biology's twentieth-century transformation. Yet comparatively little has been written about how scientists use computers. Despite this relative lack of scholarly attention, the claim that computers revolutionized the life sciences by making the impossible possible is widespread, and relatively unchallenged. How did the introduction of computers into research programs shape scientific practice? The Museum of Vertebrate Zoology (MVZ) at the University of California, Berkeley provides a tractable way into this under-examined question because it is possible to follow the computerization of data in the context of long-term research programs.

Building a database for scientific collections

Today it is difficult to untangle the life sciences from the information sciences. Bioinformatics is burgeoning and interconnected databases are increasingly central to the work of biologists. These developments are often attributed to advances in molecular biology but the establishment of electronic databases for scientific collections also anchored the life sciences to computers in important ways.[1] Natural history museums captured the interest of computer programmers and applied mathematicians. Museums held elaborate, well-characterized data that seemed ready to plug into computerized databases.[2] For the most part, the interest was mutual. Natural history museums were struggling with an image problem – their research was increasingly judged to be old-fashioned. The wave of new technologies that emerged in the life sciences during the mid-twentieth century contributed to the perception that natural history was not much more than stamp collecting.[3] Those working in natural history institutions were under pressure to modernize their methods.[4] Numerical taxonomists were convinced that computers would revolutionize taxonomy in the same way that microscopy had transformed biology in the nineteenth century. While numerical taxonomy promised to make systematics more rigorous, in part by employing the analytical powers of emerging computer technologies, collections managers and curators were more intrigued by the possibilities of computerized data banking.[5] Computers promised to update natural history, first and foremost, by improving curatorial procedures; research was only a secondary and passing concern.[6] Writing in the journal *Taxon* in 1974, Stanwyn Shetler labeled these promises as myths and warned that the computer 'has a greater ability to enslave than to liberate.'[7] How did the introduction of computer databases into natural history collections shape research? Tracing activities in Berkeley's Museum of Vertebrate Zoology (MVZ) offers some interesting and perhaps unexpected answers.

Looking at the MVZ's research and curatorial practices draws attention to the historical contingencies and institutional arrangements that made computerization possible. The MVZ story also reveals the richness of natural history databases in the early twentieth century. In fact, the early electronic data processing tools were not equipped to accurately represent the interrelated multimedia content of natural history collections. For example, although natural history museums had long been accumulating photographs, drawings, field notes, and correspondence that were associated with specimens, the early databases did not provide a means to store these objects nor offer a way of representing connections between them. As a result, some specimen data was computerized while

Corresponding author: Sunderland, M.E. (Mary.Sunderland@berkeley.edu, mary.sunderland@gmail.com)

Abbreviations: ASM: American Society of Mammalogists; GRP: Grinnell Resurvey Project; MVZ: Museum of Vertebrate Zoology; NSF: National Science Foundation; SELGEM: SELf-Generated Master; TAXIR: TAXonomic Information Retrieval.

*Tel.: +1 510 642 4581.

[1] Lily E. Kay, *Who Wrote the Book of Life. A History of the Genetic Code* (Stanford: Stanford University Press, 2000); Joseph November, *Digitizing Life: The Introduction of Computers to Biology and Medicine* (Ph.D. Dissertation, Princeton University, 2006); Bruno J. Strasser. 2010. Collection, Comparing, and Computing Sequences: The Making of Margaret O. Dayhoff's *Atlas of Protein Sequence and Structure*, 1954-1965. *Journal of the History of Biology* 43: 623-660; Bruno J. Strasser, 2011. The Experimenter's Museum: GenBank, Natural History, and the Moral Economy of Biomedicine. *Isis* 102: 60–90; Joseph November, *Biomedical Computing: Digitizing Life in the United States* (Baltimore: Johns Hopkins University Press, 2012); Miguel García-Sancho, *Biology, Computing, and the History of Molecular Sequencing* (Basingstoke: Palgrave Macmillan, 2012).

[2] Joel B. Hagen. 2001. The Introduction of Computers into Systematic Research in the United States during the 1960s. *Studies in the History and Philosophy of Biological and Biomedical Sciences* 32 (2): 291–314; Christine Hine, *Systematics as Cyberscience: Computers, Change, and Continuity in Science* (Cambridge: MIT Press, 2008).

Available online 7 May 2013

[3] Kristin Johnson. 2007. Natural History as Stamp Collecting: A Brief History. *Archives of Natural History* 34: 244–258.

[4] Toby Appel, Shaping Biology: The National Science Foundation and American Biological Research, 1945-1975 (Baltimore, MD: Johns Hopkins University Press, 2000); Hagen, 2001; Lynne K. Nyhart, Modern Nature: The Rise of the Biological Perspective in Germany (Chicago: University of Chicago Press, 2009); Mary E. Sunderland. 2012. Modernizing Natural History: Berkeley's Museum of Vertebrate Zoology in Transition. *Journal of the History of Biology*, doi: 10.1007/s10739-012-9339-3; Keith Vernon. 1993. Desperately Seeking Status: Evolutionary Systematics and the Taxonomists' Search for Respectability 1940-1960. *The British Journal for the History of Science* 26: 207–227.

[5] Keith Vernon. 1988. The Founding of Numerical Taxonomy. *The British Journal of the History of Science* 21: 143-159; Hagen, 2001.

[6] On research as a secondary and passing concern, see Mary Sunderland interview with James Patton, January 10, 2013, MVZ, Berkeley, CA. On the allure of computers to natural historians see David Sepkoski. 2012. Toward 'A Natural History of Data': Evolving Practices and Epistemologies of Data in Paleontology, 1900-2000. *Journal of the History of Biology* doi: 10.1007/s10739-012-9336-6 and Bruno J. Strasser. 2012. Data-driven sciences: From wonder cabinets to electronic databases. *Studies in the History and Philosophy of Biological and Biomedical Sciences* 43: 85–87.

[7] Stanwyn Shetler. 1974. Demythologizing Biological Data Banking. *Taxon* 23: 17-100, 71.

www.sciencedirect.com 0160-9327/$ – see front matter http://dx.doi.org/10.1016/j.endeavour.2013.04.001

other data remained embedded in photographs, field notes, and correspondence.

In the 1970s it wasn't necessarily practical to build computerized databases. Not only was the technology unable to capture the full scope of collections data, it also relied on mainframe computing. These complicated machines were very expensive to operate and usually required the establishment of distinct departments or centers with separate budgets and staff. As a result, there were significant barriers to their use, but many institutions pushed forward with electronic data processing initiatives. Even though the user community wasn't yet ready to fully embrace the computerized database as a research tool, the technology enabled subtle changes in the way work was done. Over time, these subtle differences cumulated to effect substantial change. It makes sense, therefore, to consider the effects of digital databases on different time scales and in different local contexts. Focusing on the efforts to digitize the MVZ's mammalogy collection during the late 1970s and early 1980s provides a window on the short-term effects of introducing electronic databases into natural history museums.

Demarcating short-term from long-term change is revealing when asking questions about how computers shaped research because it helps to unpack claims about whether or not computers were revolutionary.[8] Focusing on different time scales shows that although computerized databases initially had very little impact on day-to-day activities, at least for decades after they were built, their small influence effected substantive changes that ultimately ended up changing the character of the work. In the long-term, it is evident that computerization, more generally construed, fundamentally changed research. This large-scale change, however, was dependent on the coalescence of many different computerization efforts, including the digitization of scholarly journals, the emergence of the internet, the prominence of personal computing, as well as the computerization of databases. Looking at the MVZ's story reveals the short- and long-term costs and benefits of early electronic data processing. Although these early computer databases did not obviously impact research, the act of computerizing data contributed to a shift in how different kinds of data were valued. Because it was not possible to computerize photographs, images, and field notes, these types of data were no longer gathered or archived with the same kind of rigor. Although it is true that some investigators continued to take photographs and keep meticulous field notes, they were not obviously a part of the MVZ's database. Objects, such as photographs, and their affiliated data were located on the periphery and therefore became less visible to researchers, especially to those who might not physically visit the museum.

Now that the technology is available, the MVZ, along with other institutions, is making an effort to digitize its field notes and photographs while also imagining new ways to analyze photographs to extract ecological information. These recent developments are exciting, but they also draw attention to the potential long-term effects of excluding photographs from the earlier database. Finally, focusing on the effects of computer databases in the short term draws attention to the symbolic power of the MVZ's early computerization efforts. Computerization communicated the MVZ's commitment to engage in cutting-edge research and demonstrated its willingness to embrace new technologies – characteristics that attracted new researchers from across disciplinary and national borders.

In 1977 the MVZ submitted a proposal to the National Science Foundation (NSF) to initiate the 'computerization' of their collections. Transferring from a 'hand-written' to a 'computer organized' approach was motivated by the desire to improve curatorial practices and ultimately expand the ways in which the data could be used and disseminated. The director of the NSF's Biological Research Resources Program actively solicited the MVZ proposal because the institution had earned a reputation as being an effective early adopter of a variety of technological approaches to working with collections, such as electrophoresis. [9] The MVZ seemed like a good place for the NSF to grow its computerization efforts.[10]

Founded in 1908, the MVZ was designed as a research institution. Joseph Grinnell, the MVZ's founding director envisioned a place that could facilitate long-term studies of evolution. In fact, Grinnell predicted that the MVZ's collections would not realize their true value until 100 years had past. With the future in mind, Grinnell put in place a variety of standardized procedures to ensure that the MVZ's growing collections were scientifically valuable and well archived.[11] The result was a large database connecting objects and information deemed relevant to evolutionary questions (Figure 1).

According to Grinnell, the scientific value of each specimen depended on the information associated with it. Grinnell designed the MVZ's infrastructure to keep track of each specimen's detailed locality information in addition to any available ecological and behavioral information. Grinnell cared deeply about localities because he was interested in biogeography, speciation and subspeciation; he wanted to better understand the relationship between species formation and the kinds of geographical and ecological boundaries, such as lakes and mountains, which existed in the natural environment. This mattered to Grinnell because he was developing a theory of speciation. Although he never had a chance to write his big book on the topic, the

[8] Agar, Jon. 2006. What difference did computers make? *Social Studies of Science* 6: 869-907.

[9] Sunderland, 2012.

[10] William Z. Lidicker and James L. Patton. 26 April 1978. 'Operational Support of the regular collection of mammals in the Museum of Vertebrate Zoology.' National Science Foundation proposal, MVZ Miscellaneous Archives, MVZ, Berkeley, California, p. 2. The MVZ's solicitation to submit a proposal was discovered during Mary Sunderland's interview with James Patton, September 17, 2012, MVZ, Berkeley, CA; Toby Appel, *Shaping Biology: The National Science Foundation and American Biological Research, 1945-1975* (Baltimore: Johns Hopkins University Press).

[11] Joseph Grinnell. 1910. The Methods and Uses of a Research Museum. *The Popular Science Monthly* 77:163–169. For more on the early history of the MVZ see Elihu M. Gerson, *The American system of research: Evolutionary biology, 1890-1950* (Ph.D. Dissertation, University of Chicago, 1998); James R. Griesemer and Elihu M. Gerson. 1993. Collaboration in the Museum of Vertebrate Zoology. *Journal of the History of Biology* 26: 185-204; James R. Griesemer 1990. Modeling in the museum: On the role of remnant models in the work of Joseph Grinnell. *Biology and Philosophy* 5: 3-36; Barbara R. Stein, *On her own terms: Annie Montague Alexander and the rise of science in the American West* (Berkeley: University of California Press, 2001); Mary E. Sunderland. 2011. Teaching Natural History at the Museum of Vertebrate Zoology. *British Journal for the History of Science*, doi:10.1017/S0007087411000872.