Ber. Wissenschaftsgesch. 45 (2022): 508 – 516 www.bwg.wiley-vch.de

doi.org/10.1002/bewi.202200035

The Politics of Sources Meets the Practices of the Librarian: An Interview with Esther Chen**

Esther Chen,* Lara Keuck,* and Kärin Nickelsen*

Summary: [I] want to single out one phenomenon that could be called the 'politics of sources'. It points to the extent to which the histories that both scientists and historians can write are artifacts of the available sources. The Rockefeller Foundation not only opened its archives very early on for historical work but also invested a lot in making the archives readily available for historical exploration. During the 1980s, many young historians took advantage of this opportunity. Thus, in a relatively early phase of the professional historiography of molecular biology, one could have gained the impression that the development of the new biology as a whole was a bio-politically directed enterprise of the Rockefeller Foundation sustained by the vision that social processes could ultimately be controlled by biological processes.[1]

Keywords: Library, Library Catalog, Collection Building, Dewey Decimal Classification, Digitization, History of History of Science, Max Planck Institute for the History of Science

The library at the Max Planck Institute for the History of Science (MPIWG), Hans-Jörg Rheinberger's professional home since 1996, has played an important role since its founding and remains indispensable to the functioning of the

E. Chen
Max Planck Institute for the History of Science, Berlin
E-mail: echen@mpiwg-berlin.mpg.de

L. Keuck
History and Philosophy of Medicine, Bielefeld University
Max Planck Institute for the History of Science, Berlin
E-mail: lkeuck@mpiwg-berlin.mpg.de

K. Nickelsen
History of Science, LMU Munich
E-mail: K.Nickelsen@lmu.de

** This interview is part of a special issue, *On Epistemic Times: Writing History 25 Years after Synthesizing Proteins in the Test Tube*. We would like to thank Elizabeth Hughes for translating the edited German interview into English.

institute today. The library's first director, Urs Schoepflin, made important decisions regarding the development and arrangement of the collection and catalog of services; from its start, the library has been a significant actor in the institute's digitization strategy. However, just as archives, whose "politics of sources" Rheinberger reflected on,[2] libraries have always been much more than just service and infrastructure. Libraries collect, preserve, and organize knowledge resources. They are places of serendipity but they silently steer our fortuitous discoveries by arranging the books on the shelves into distinct clusters and intellectual neighborhoods. They can make books accessible through inclusion in the collection, but can also render them invisible through exclusion. Their classification systems reflect and reify the categories of analysis, but at the same time, at least to some extent, they shape and influence analytical perspectives. All of this is especially true of reference libraries that invite on-site use. On 24 January 2022, Lara Keuck and Kärin Nickelsen met with Esther Chen, the head of the MPIWG library since 2015, to learn more about how this dynamic plays out at the Institute. The interview delves into the complex nature of libraries, the productive tensions between spaces—physical and virtual—and the people who are creating, developing, and changing these spaces, librarians and scholars alike. It offers insight into the practices of the institute's librarians both in the physical and digital spheres, the constraints that they are subjected to, and the impact they have on historical research.

Lara Keuck/Kärin Nickelsen (LK/KN): Over the past decades, the MPIWG, along with the history of science at large, has enormously expanded its subject area, from a Western history of science to a global history of knowledge. Since the MPIWG library's acquisitions have always been aligned to research projects, the holdings effectively document this development of the Institute—perhaps even of the entire field. How, if at all, is this reflected in the library's classification systems and/or physical arrangement?

Esther Chen (EC): These are two different things, the systems and the arrangement. The classification system, in our case the Dewey Decimal Classification, reflects this development insofar as we continue to expand the system in step with the expansion of research at the institute—which has been enormous, there is no question! As to the arrangement, this is largely determined by the system but there is also a peculiarity, namely, the distinction between sources and research literature (*Figure 1*). The library's center is dedicated to what we call sources and these titles are ordered according to their authors. This core, so to speak, is surrounded by research literature, which is arranged according to subject. This has been a controversial decision from the start, but so far it still holds, despite all other changes.

LK/KN: This distinction between sources and literature is extremely interesting. Maybe we can test out some examples? Let us take, for instance, the cases of William Whewell and Pierre Duhem. Both have written books in

[2] See quotation above.

Figure 1. Library of the Max Planck Institute for the History of Science (photograph provided by the library, 2018).

history and philosophy of science, but both in their own historical times. Where do they fall in the grid: are they sources or literature?

EC: These two are in the sources section.

LK/KN: What about Thomas Kuhn? Sources or literature?

EC: Also mostly in the sources.

LK/KN: Hans-Jörg Rheinberger?

EC: Research literature!

LK/KN: Ok, the last one seems uncontroversial, but the others could very well have been sorted differently. How do you decide whether a book is part of the sources or research literature? There seems to be a temporal element, that is, literature from the 1990s is not considered part of the sources—or, at the very least, not yet. Is this a moving boundary? In other words, will all the research literature books eventually be reclassified as sources?

EC: Now, the definition of coherent criteria is clearly the challenge with this setup. It will always remain open to criticism. Perhaps it ought to be a moving boundary, but the sorting is not only a question of age but also of research focus: what is considered part of the literature in one project may be source material in another. Accordingly, holdings in the library would have to be moved continuously from one shelf to another. This of course is not practicable.

There are other problems as well. Take the example of the history of science from a certain region: Do we classify, and physically arrange, these studies according to their geography, or do we prioritize their systematic focus? We may want to avoid regional isolates; but how can we make our increasingly

global approach visible in a classification system that is based on traditional European subjects? The answer will always be unsatisfactory. However, we need a setup that is in tune with the needs of everyday use, and we cannot regularly rearrange it, even if some of the system's implications are controversial. Ultimately, I think, the problem of Eurocentrism must be solved in research and debate, not on the bookshelf.

LK/KN: This sounds as if these questions have been debated for some time already. Is it particularly difficult to work with historians of science, who are trained to constantly question systems of classification and the ordering of knowledge?

EC: Yes, I would definitely say so. You do not have these discussions elsewhere, that is absolutely true. Elsewhere it might be—why isn't my new book there yet, or if it is there, why isn't it up-front in the display case? These complaints are very common. But this sensitivity to order and classification systems, and what they might imply, is particularly challenging here. I think it is good and important! You just have to keep in mind that for us it is not only about categories but also about how, and how often, we would have to move almost 70,000 volumes from left to right. In intellectual disputes, you can turn things constantly back and forth. But in the library, we cannot remain ambivalent; we have to decide whether to place a book here or there, even if both options would make sense in their own ways.

LK/KN: If there is a difference in sensitivity, is there also a difference in practice? How do other libraries solve the question of classification?

EC: Basically, similar to what we do here. There are several systems, and there is often a distinction between the system for entries in the catalog and the placement on the shelves. The principles of classification are very similar. Most academic libraries classify by subject first, so agriculture, psychology, or medicine, and then more specific categories follow. In medicine, there is perhaps surgery on the next level, and so forth, into ever more specialized fields of knowledge.

LK/KN: Hence, these classification systems reflect what was understood as an academic subject at a certain point in time, and how it was divided into sub-branches—right?

EC: Exactly! Typical history of science question! Anyway, two of the major systems used worldwide are the Dewey Decimal Classification (DCC) and the Library of Congress Classification (LCC). We use the DCC, which was the standard for academic libraries in the U.S. for a long time, until the first half of the 20th century when the Library of Congress started printing catalog cards and sending them to libraries. Most librarians then said—okay, if we can buy these cards, it saves us a real lot of work; so we're switching over. In effect, all academic libraries in the U.S. now use the LCC classification, not because it is better but for purely pragmatic reasons, because the libraries no longer had to create the catalog cards themselves.

There is a wonderful project in Switzerland, in the Sitterwerk near St. Gallen, which has made the dream of a permanently dynamic arrangement come true. You can take a book off the shelf and put it back elsewhere, in a completely new bookish environment, so to speak. The shelves are con-

tinuously scanned by a robot to record the current locations of the books and document them in the catalog. By this means, one can always find the books, but maybe each day somewhere else entirely, so that users will encounter new titles on the left and on the right. Tables in the reading room are also equipped with a scanning function. So, if a scholar works there for, say, six months, the system automatically assembles a bibliography of their research project. I would very much like to try something similar here, and would be very excited to see the outcome.

LK/KN: Fascinating! This sounds like an infrastructure that creates serendipity, which ought to be a contradiction in terms. But the question of the catalog brings us to another aspect of the ordering of knowledge. In the past, keywords have been crucial for literature research, but this, presumably, has changed over the past decades?

EC: This is so important, the whole subject of indexing and keywords, which of course has its origins in the era of card catalogs or even earlier. In Germany, this is still especially tedious, because only persons of a certain rank are authorized to do the indexing—which means that it can take a very long time before a book is out on the shelf. Personally, I think that subject indexing has dramatically lost its importance with the rise of e-books and digital collections. Search engine technologies use algorithms that access the full text of the searched items, so keywords matter less.

This, of course, implies that the one hundred percent control that librarians once had over their catalog entries has gone. Furthermore, we are often not even entirely sure of the body of titles. Our own catalog system (MPIWG Search), does not only access our own book collection, but also the databases that we either have in-house or to which we are connected via consortia. This is a very dynamic entity, and it can be challenging to keep track. And in some sense, it is also unnecessary, as long as it serves the users' needs. However, some of my colleagues in the field still mourn the loss of control.

This network of entangled catalogs saves a lot of work, as there is only one single catalog entry for each item, which all the other libraries in the network will copy— another loss of control. Not for us, though, because we at the MPIWG are often the first in our network to incorporate new books; our workflows are so much faster than at large institutions. This means that it is our team that creates the one data set that will be used throughout Germany, which of course comes with a high responsibility. Yet this only concerns the printed collection. In the case of digital books, much of the metadata is provided by the publisher and imported into our catalog. It is simply impossible to do this manually, for all 650,000 digital books that are available in our system. But, yes, these entries are not of the same quality as those for the 67,000 printed volumes in the library.

LK/KN: If we understand correctly, you are describing a very fluid situation where many things are changing in parallel in our new digital library universe. There is a loss of control and a shift of competences for librarians, while there are new possibilities and new uncertainties for users. What about the texts themselves? With all the hyperlinks and cross-references, a digital item no longer has the same boundaries as a physical book. How does this change

the notion of *a work* or *authorship*—categories that, of course, have been debated for a long time already. Have we finally arrived in the semioticians' dream—or nightmare?

EC: I don't know what the semioticians would think, but the changing of the text is a very important point. In the digital world, we are increasingly dealing with different units of information and new forms of publishing, and there is more radical change yet to come. We are still in a transitional period, between analog and digital. In many cases, the digital space tries to imitate some of the analog forms. We are still largely working with PDF, for instance. One of the main reasons is the publisher's commercial interest and their wish to stay in control.

But, as I said, this will change very soon. There is, for example, a steady increase in data publications in the humanities, including micro-pieces in blog format and so on. There are many unresolved issues, e. g., the question of authorship and credit. The respective scientific communities still need to recognize these new formats as achievements, but I think it is only a question of time. I expect that in the future we will see a much greater variety of publication formats, beyond the monograph and the article.

LK/KN: These new publication formats bring us once again to the question of ordering criteria and classification, and to the challenge of interweaving analog and digital holdings, in the catalog and on the shelves.

EC: Yes, this is a huge challenge. Currently, our users are basically in two separate libraries, depending on whether they are walking along the shelves or reading the e-book collections; we share this problem with libraries around the world. How can we make the vast amounts of digital holdings visible in the library space, how can we remind digital users of items that are only available on the shelf? Some have tried putting QR codes of digital books on library shelves, others have tried virtual or augmented reality approaches—but I personally find the idea creepy to run through the library with special glasses. I hope that we will find better solutions here.

On the question of sorting digital holdings: Since they do not require a physical set-up, the ordering is done exclusively via the catalog. However, given the possibilities of full text search practices, we are so much more flexible in the digital space than with the printed collection. This is why I think it is a particularly urgent task for libraries and librarians to radically question the classic catalog, because its structure is still firmly rooted in the era of printed journals, monographs, and catalog cards, and fails to integrate the infinite potential of digital collections.

At the MPIWG, we have now created a digital research infrastructure that follows a completely different approach. We are trying to not only include all so-called classic library holdings but also link them with pertinent research data, for example from web pages. I cannot go into details here, but with the use of powerful ontologies and what are called stable identifiers, sources and research data are now accessible both at the same time—and many, many,

many of them.[3] A full universe of data opens up—with the caveat that direct access is only possible with the appropriate programming skills. In order to be useful for others, an interface is needed that is based on certain predefined paths through the data universe. The infinity, which was so dazzling a moment ago, is thus curtailed again. These paths naturally have to be chosen with great care, because they determine what will be easier to find later and what may be more difficult.

LK/KN: It seems that, in a sense, we have come full circle back to the ordering of knowledge and the balance between intellectual adequacy and the challenges of pragmatics. The criteria for defining these paths and, in particular, the nature of the stable identifiers, which suggest unequivocal clarity, immediately raise our interest and suspicion, but we will leave it there, for the moment. We have one final concern, which is, however, a big one. If we consider the totality of the new digital orders, keywords, links, identifiers: What is their impact beyond specific search practices? How does the digital library shape our research perspectives, possibly even our understanding of history?

EC: Puh. Let me dodge the metaphysical question for now. However, the practical changes are already enormous. There are so many more sources almost immediately at hand now, and they can be compared in a completely new way in direct synopsis. In the past, you had to travel all over the world to do this (and first find the money to do so!), while now you can do it on screen, in the large portals like Europeana.eu. And there are also completely new technical tools to compare and examine the sources—just take the possibility of zooming in on individual image sections. These are huge practical changes, in so many different ways.

On the other hand, we must not forget that while some things are now easily accessible, so much more is excluded at the same time, simply because it is not available in digital form. We should always keep in mind that only a very small part of our cultural heritage is digitized. Despite the wealth of material that is accessible on screen, there is still far more to discover in the analog world! Paradoxically, the more we focus on digital accessibility, the more sources and information we will miss.

I see today's librarians less as guardians than as pilots, navigating users through the collections, also in technical terms. My impression is that research today does not need less support from the library and IT, but more! We are much more involved in research processes, in dealing with research data, classification systems, access paths, and so forth, all of which comes with a lot of responsibilities.

LK/KN: Speaking of responsibility. Let us summarize our insight into the impressive multitude of elements that come together in the MPIWG library. We always suspected that the library was important, but it seems that we still underestimated its role. There is the arrangement of books in proximity to some and at a distance to others, which steers our serendipitous discoveries on

[3] A detailed description of these approaches can be found in: Kräutli et al. 2022; Chen and Kräutli 2020.

the shelf. Then, there is the classification system and the catalog entries: apparently, the MPIWG library often creates the standard data that many other libraries will simply copy. In the digital space, stable identifiers are established, as well as paths determined and interfaces created, and the librarian acts as pilot to help us navigate the system. Now, considering that a large part of the global history of science community visit the MPIWG at least once in the course of their careers, and conduct significant parts of their research here, quite substantially supported by the library with incalculable consequences: Is Esther Chen the éminence grise of our field?

EC: Um. (*Laughs.*) No, definitely not. Our job is to do the basic work that enables research, but it is nice that our contribution is being so well recognized. It is, after all, a statement to put the library in the middle of the institute—at its heart, so to speak. But precisely because we are dealing with historians of science, our influence should not be overestimated. We have already talked about the fact that everything is being questioned here: How did the source item get into the archive in the first place, how did it reach us in the library, and what context is it placed into here, physically, systematically, and digitally? What are the alternatives, what are the implications? Everything will be interrogated, and although it is sometimes exhausting, it also makes our job particularly interesting. Our users are not passively supplied, they bring so much to the table! I rather like the image of a port, where ships are constantly arriving and departing, perhaps taking something on board, but at the same time also unloading goods. It is not a one-way street, it is an exchange, and that is very important.

LK/KN: But in the port, of course, there are also only certain wharfs, the goods are begin loaded and unloaded according to clear specifications, there is harbor police and pilot boats, with whose help the large tankers are towed in. So, there is exchange between those who arrive and those who are in charge of the port, but not quite on equal footing. At the same time, a port is not eternally the same, there is always remodeling going on, and new anchorages may be developed for new types of ships.

EC: Very nice image. My brother was a sailor for a long time, he will like this a lot.

Acknowledgements

Open Access funding enabled and organized by Projekt DEAL.

References

Chen, Esther, and Florian Kräutli, "Digital Sources and Research Data: Linked and Usable," *Europeana Tech Insight* 15 (Mai 2020), online: http://hdl.handle.net/21.11116/0000-0006-6FBF-8 (accessed 19 July 2022).

Kräutli, Florian, Esther Chen, and Matteo Valleriani, "Linked Data Strategies for Conserving Digital Research Outputs: The Shelf Life of Digital Humanities," in *Information and Knowledge Organisation in Digital Humanities*, ed. Koralijka Golub and Ying-Hsang Liu (London and

New York: Taylor & Francis Group, 2022), 206–224, online: DOI 10.4324/9781003131816-10.

Rheinberger, Hans-Jörg, "What Happened to Molecular Biology?", *BioSocieties* 3 (2008): 303–310.