

The *Isis* Current Bibliography of the History of Science: Working on a Bibliography That Is Global in Scope

by Stephen P. Weldon, Ph.D.
Associate Editor, *Isis*
Dept. of History of Science
University of Oklahoma

The following paper was written to provide participants of the World History of Science Online project (WHSO) an understanding of the history of the Isis bibliography of the history of science and to provide insights into two areas directly relevant to the project, namely subject authorities and international collaboration. My history of the Isis Current Bibliography (Isis CB) includes a description of the History of Science, Technology, and Medicine online database (HSTM) hosted by the Research Library Group (RLG). In my discussion of authorities, I discuss the Isis classification systems that have been used for the various versions of the bibliography for readers interested in classification; however, I also note that the online version of the Isis data does not include the classification data that are used to organize the citations in the printed bibliographies. I go on to explain how I have derived my subject thesaurus that is used in the HSTM database, and I suggest that WHSO might be able to use the Isis thesaurus as the core of a standardized authority list for all WHSO bibliographies. Finally, I explain how I have created some simple and effective ways of interacting with contributors who have supplied entries to me or who have agreed to help proofread foreign-language citations. I hope that this discussion will help others who are creating their own bibliographies for WHSO.

The History of Science Society in the United States publishes the journal, *Isis*, subtitled “An International Review Devoted to the History of Science and Its Cultural Influences.” A print copy of the *Isis* bibliography is produced annually and mailed along with the December issue of *Isis*. Each issue ranges in size from 2,500 to 4,000 classified bibliographic citations, and it also includes hundreds of book review citations. In addition, the data that is published in print is also uploaded into the Research Library Group’s online History of Science, Technology, and Medicine database.

Half of my time as a faculty member at the University of Oklahoma is devoted to the production of the bibliography, and I have the assistance of two graduate students in the History of Science Department.

I want to draw your attention to the subtitle of *Isis* quoted above. Notice that the journal aspires to be international, and it also aspires to be broad in scope: its subject matter is not just science, but science and its cultural influences. The founder of *Isis*, the Belgian émigré scholar George Sarton, had a global vision of history in which science played a dominant role. The scope of the *Isis* bibliography continues to be driven by Sarton’s vision, including works on the history of science in all of its manifestations across the globe and from prehistoric times to the present.

This wide subject matter presents both great advantages as well as a number of challenges. Since taking the position of *Isis* bibliographer three years ago, I have sought to implement practices that address the most difficult of these challenges. I’m hoping that my experiences as bibliographer for the History of Science Society in the United States will be useful for the larger

community of bibliographers and archivists, as we think about the World History if Science Online project.

History of the *Isis* Bibliography

The *Isis* bibliography originated with the first issue of *Isis* that George Sarton produced in 1913. Sarton decided early in his life that the history of science ought to be promoted as a distinct discipline of its own and that the most important way to begin was to establish a journal in the field, a journal which would not only publish scholarly articles in the history of science, but would also be a resource that would bring together scholarly work that otherwise would not be found. (Joy Harvey, in her Introduction to the 2001 *Current Bibliography* [*Isis* 2001 {92}: vi-vii], provides an excellent short history of the bibliography from its inception, and much of my discussion is taken from this.)

Because Sarton's vision of the history of science was so expansive, he had to develop a rather complex classification system for the various works that he cited. (See Appendix 1.) In addition to compiling and categorizing the entries, however, Sarton also annotated them; his bibliography was not a simple classified list of bibliographic sources, it was what he termed a "critical bibliography." And the flavor of his scholarship is evident in the annotations that he made throughout the bibliography.

Sarton conceived of *Isis* in Belgium, and this is where he published the first issues, but with the advent of the First World War, Sarton was forced to leave home. He ended up in the United States where he was able to obtain a professorship at Harvard University. Until Sarton retired in 1952, the work on *Isis* was done by him and a few assistants, mostly graduate students. With Sarton's retirement the editorship of *Isis* was taken over by one of his students at Harvard, I. B. Cohen, and the bibliography was given over to a group of scholars around the world under the direction of Henry Guerlac.

The annual bibliography had entries each year. The committee made one very significant change in the bibliography, by modifying Sarton's classification system, getting rid of his very broad civilization categories and tending to focus the bibliography more tightly on established scientific disciplines. (See Appendix 2.)

The committee continued to publish an annual bibliography in the pages of *Isis* until the History of Science Society appointed John Neu, a librarian at the University of Wisconsin—Madison, as editor in 1967. Neu edited the bibliography from then until his retirement in 1999. Two things changed when Neu took over editorship. First of all, he quickly decided that it was easier for him to work alone than with a committee of scholars. Since the University of Wisconsin had one of the largest university library systems in the country, he found he was able to gain access to most sources that he needed to include in the bibliography. Second, and more important, Neu changed the nature of the bibliography because, not a historian of science himself, he felt that he was unable to produce a critical bibliography in the way that Sarton had done. The name of the publication changed from "critical bibliography" to "current bibliography." Neu was much less selective than his predecessors, and the bibliography grew large enough to publish as a separate issue annually. By the time of his retirement, Neu was able to produce annual bibliographies that included about 4,000 main entries.

The bibliography moved to the University of Oklahoma in 2000 where the historian of science Joy Harvey acted as interim editor for two issues before I was hired to assume editorship in 2002. The first three years of my tenure were extremely busy. I found it necessary to implement a number of important changes in the bibliography itself as well as in the way that the bibliography was produced. Entirely new software had to be created to replace the database that John Neu had been using, which was quite outdated by the time he retired. In addition, I found that after fifty years, the classification system that had been created by Guerlac's committee was showing its age as well. New scholarship being produced was not easily placed in this classification system. After a series of discussions with historians of science from different fields, I made some significant changes to the classification system—changing terminology, adding new categories, and rearranging older categories—at the same time, however, I retained the system's two-level classification structure that gives priority to chronology and cultural/geographic contexts. In addition, because the new software was more powerful than the old software, it was possible to add a full subject index, which had included only proper names in Neu's bibliographies.

The citations from the annual bibliographies have also been published in two other formats: as printed, cumulative bibliographies and as records in the online History of Science, Technology, and Medicine (HSTM) database.

There have been four print cumulations. The first collected all of the works from 1913 to 1965, and was edited by Magda Whitrow. In order to bring all of the citations for those fifty years together into one collective bibliography, Whitrow created a new classification scheme that included both subject and aspect terms and a complex system of notation that provided quite fine subdivisions for locating works on particular subjects. John Neu continued the cumulations at ten-year intervals, publishing one in 1975, 1985, and 1995. With the publication of the last of these cumulations, the History of Science Society decided to no longer publish any further cumulations in print because by that time, the bibliographic data was accessible to scholars in universities around the United States and all members of the HSS through the HSTM database. It was thought that this online database would take the place of the print cumulations.

The HSTM database is hosted by RLG, the Research Library Group. It is a subscription service that all members of the history of science society are able to access for free. The RLG database was established in the 1980s, and now includes citations from the History of Science Society, the Society for the History of Technology, and *Bibliografia Italiana di Storia della Scienza*, and the Wellcome Institute for the History and Understanding of Medicine. Protocols for the data submission of bibliographic records are based on MARC standards, a widely used electronic cataloguing standard developed by the Library of Congress. (A slightly revised version of this protocol has been drawn up as a standard for WHSO records.) The data from *Isis* goes back to 1976. Book reviews, unfortunately, were not included until 2003.

The current mission of the *Isis CB* is clear. It publishes in two formats, an annual print format and an electronic, on-line data format. It maintains its original global vision. The wide disciplinary and topical scope have probably increased over the years as new fields of the sciences have arisen and as historians of science have become more interested in traditional knowledge systems. The scholarly approaches have also increased so that the bibliography covers history, philosophy, and sociology of science in their diversity. The one thing that has not changed, however, is the nature of the sources. It still indexes only books, chapters of edited collections, journal articles, and book reviews. Although some of these sources may be

published only in digital form, the criterion is that they must have been published. No manuscript or archival material is indexed unless it is compiled as a recent publication. Likewise, no primary sources are included unless they are old works published for their historical value in special editions.

* * *

Having set forth the background of the *Isis CB* I want to consider two major aspects of the project that have direct relevance for other online bibliographies.

Classification and Indexing

For printed bibliographies, classification of entries is extremely important. In the *Isis* bibliographies, there have been several classification schemes over the years. All of the classification schemes that have been used have made discipline, chronology, and cultural context central to their organizations, though they have done so in different ways and used different terminologies. In each of the classification schemes, the preferred place to locate an entry is within a chronological period under a subcategory related to a specific scientific discipline: thus, “Nineteenth century—Astronomy” or “Prehistory—Technology.” In each of the bibliographies there are also categories that admit items that do not fall within those period-discipline categories: an ethnographical or cultural context (“East Asia”), for example, or a disciplinary category that does not fall under a chronological period (“Geology”), or a topical or methodological category (“Science and art” or “Historiography of science”). The number and arrangement of the categories differs from one classification system to another, but the basic principle of classification holds for all of these bibliographies. The interested reader can compare the examples of classification given here in Appendixes 1-3 to get an idea of how the various classification schemes work.

No linear, hierarchical classification scheme is adequate to the task of arranging the multiplicity of all the data for all users; compromises must be made, and certain schemes that work better for some users will necessarily be much less manageable for other users. Print sources are by their nature bound by the constraints of linear arrangement of data. Indexing can partially alleviate the problem, and to that end, the annual printed *Isis CBs* after 2002 have extensive subject indexes. These new subject indexes allow users to come to the data in non-linear ways, and find items that might otherwise be missed in scanning the hierarchical arrangement.

Computer databases extend the search capabilities of subject indexes. With their greater flexibility, they allow users to make searches that more closely approximate their own interests. Databases thus break apart the linear arrangement of the classification scheme and make it less relevant, which is one reason that I don’t include the classification information with the *Isis* citation data when I submit it to the online database. There are disadvantages, however to indexing as well. Many people complain about the non-hierarchical nature of databases, arguing that they appreciate having items in well ordered linear lists because there is an advantage to simply browsing through the list. Serendipitous discoveries, for example, can sometimes be among the most useful, and these will not be as common when users adopt the more precise terminology of indexes. At the same time, there is also a false sense of precision with databases. Terms are not always applied in precisely the same manner, nor can they be, since subject terms are as unclear as broader category terms. The advantage of subject index terms lies in the fact that they can be clustered in a variety of ways that linear category terms cannot.

The importance of finding index terms that will be consistent and that will work across databases is the main obstacle. When I took over the bibliography, I changed the way that items were indexed. John Neu had been utilizing the more detailed, but still linear category classification scheme used in the cumulative bibliographies, but this proved to have some serious drawbacks when adopted as a subject search term for entries. The main problems were with the imprecision of the category terms. Thus, one of the cumulative categories was “mechanism; vitalism; reductionism; materialism”; in order to make subject searching more precise, I have separated the four terms so that I can use them independent of one another. This also has given me the flexibility to add new terminology where there is no clear fit with older terms.

Adding new terminology that is not already part of the Neu subject data, however, is sometimes rather difficult because I must search for terms that may already be in use. First, I look for terminology already present in the *Isis* data. Second, if there is not an appropriate subject term already in use, I will check the HSTM database to see if other bibliographies have used a term that I need. If so, I will borrow that term and add it to my thesaurus. If, however, I find nothing suitable, then I do a third search for an approved authority in the Library of Congress. If I determine that the concept is significant enough for our discipline, even after failing to find an appropriate term in the above three places, I will then create a term based on my understanding of the concepts importance and use in our discipline.

This work has resulted in a thesaurus that now has over 17,000 terms in it. It includes discipline- and topic-specific subject terms, it contains personal and institutional names (which constitute the majority of entries), and it has hundreds of geographical locations and chronological periods. In addition, I have numerous cross references and a list of terms that have been used by Henry Lowood in the Technology and Culture database. I have proposed that this list be used as the core of a subject authority list for the WHSO database, and the thesaurus is now available online in a very simple set of linked web pages: <http://www.ou.edu/cas/hsci/isis/th/thesaurus.html>. Eventually, I would like this thesaurus to be made interactive so that bibliographers who are contributing to the WHSO project will be able to add new terms and cross-references, as well as comment on the usage of terms.

Communicating Internationally

Because the bibliography is international in scope, I've found it necessary to utilize the help of scholars around the world to both locate sources to which I have no access and to proofread material that I have entered into the bibliography. To this end, I have established a board of contributors who send me a modest collection of entries each year which I enter into the bibliography and send back to them to be proofread. Currently, my board is composed of Duan Yaoyang (China), Chu Ping-yi (Taiwan), Moon Joong-Yang (Korea), and Slava Gerovitch (USA, who helps with Russian sources). The reason that this is possible is that I have been able to devise a way to rapidly produce a proof sheet in PDF format that can be read with a free Adobe Acrobat Reader. This relatively rapid two-way communication is essential for international collaboration of this sort.

As I was developing the bibliographic software necessary for producing the *Isis* bibliography, I have had to make some key decisions relating to language use. First of all, I have decided that I will only use Latin-based languages until I can develop a clear and rapid way of including non-Latin scripts. Even using Latin-based languages presents problems in dealing with less common

diacritical marks because the database software that I use, FileMaker, doesn't always support the full extended character sets. Where that is the case—and even in cases where I need to enter rather common diacritics directly from the keyboard—I have had to develop a series of codes that will be translated by the typesetting software into the proper Unicode character for the PDF. Thus, an acute accent over an e, for example, is coded as <e'>, and an umlaut over an o is <o''>. My encodings are expanding, but currently they are as follows. (See Appendix 4.)

Finally, there is the question of the medium of communication. I've already mentioned that I rely extensively on electronic communication, mainly email, with my board of contributors. But I also value regular mail because of the great advantage that hard-copy sources have. If I can get the front matter and table of contents from a book or journal, I find myself much more certain of my source than an electronic citation from a database like OCLC WorldCat or a publisher's web page, where errors are relatively common. At the same time, I do recognize the need to obtain information about sources from the many scholars that produce them and use them, so in addition to publishing my email address, I have developed a web-based form that allows users to submit information about citations. The form is simple to fill out, and the submitted data is then compiled as an email message and sent directly to me.

Conclusions

There are a number of things to be learned in this project. First, the nature of international projects requires careful thought and standardization of authorities insofar as possible. To that end, I have been creating a thesaurus of names, institutions, and subjects that is based on the subject terminology used in the RLG database. Where that is impossible, I use the Library of Congress. Occasionally, I will create a subject authority that is not in either of these databases but is a "term of art" in our discipline.

Second, international projects require extensive communication among individuals and the capability to work in different languages. To date, my work continues to be done only in Latin-based scripts; all other languages are transliterated into a Latin script. But for any serious international project non-Latin languages will need to be accepted as well, and software based on Unicode will most likely be required. In addition, people will need to be able to communicate both electronically and with hardcopy, so their protocols will need to be flexible enough to deal with both.

The effort to produce an international bibliography will be great, but the tools available for this are at hand, and it is now possible to produce a web-based internet resource based on collaborative bibliographic production that was not possible even fifteen years ago. My hope is that the *Isis* bibliography will be able to provide help and guidance for people who are in need of assistance in contributing to the World History if Science Online project.

[Appendix 1. Isis Categories 1911-1953--Handout.doc](#)

[Appendix 2. Isis Categories 1953-2001--Handout.doc](#)

[Appendix 3. Isis Categories 2002-present--Handout.doc](#)

[Appendix 4. Special symbols .doc](#)