Database History

Chapter 5

ERIC

ERIC, the Education Resources Information Center, is another database with a long history of excellence, but it has gone through some significant changes in its lifetime. ERIC was a “born digital” effort: in 1966, the Department of Education established 16 clearinghouses, each with its own educational subject specialty, to “gather, index, and input bibliographic information” (Tenopir 2004) for journal articles and documents known as grey literature. Grey literature is material that has not been commercially published, but that has scholarly value: papers presented at conferences, progress reports, working papers, technical reports, lab notebooks, student papers, curriculum guides, etc. One of the great contributions of the ERIC database was the amount of effort put into gathering, evaluating, indexing, inputting, and making available (through a microfiche distribution program, and now directly in PDF format) a huge amount of otherwise unpublished and inaccessible, but valuable, grey literature. At more than 1 million records, the ERIC database is now the largest and foremost subject resource for education.

But what the government creates, the government can also change. Starting with the Education Sciences Reform Act of 2002, the ERIC effort began to change. In December 2003, the 16 clearinghouses were closed, “as part of an effort to revamp, streamline, and centralize” the system (Viadero 2004). The Education Department’s goals were laudable: one centralized, easy-to-use resource with more full text. What followed, however, was a period of interregnum: the database and the materials already indexed continued to be available, but no new material was being added, even though production continued apace. In March 2004, the Department of Education awarded the contract for their new vision of the ERIC system to Computer Sciences Corporation (CSC), and by June 2005 the loading of new materials had resumed. The library community closely followed all of these changes. Members of the Education and Behavioral Sciences Section of the Association of College Research Libraries/American Library Association and the Education Division of the Special Libraries Association continue to monitor what is happening with the database and to communicate with CSC, providing a useful outside oversight for this important resource.

Luckily, even when there were 16 clearinghouses offering various services in support of their educational topic area, there was still only one ERIC database to which they all contributed records. A variety of vendors contract with the CSC to obtain the ERIC data, and each vendor then offers the database through its own search interface. In addition to these choices, the Department of Education also provides free access to ERIC on the Web, and it is this version we will work with here.
PsycINFO

PsycINFO is the online version of the American Psychological Association’s venerable Psychological Abstracts, a print abstracting and indexing service dating from 1927. At more than 3 million records as of May 2011, PsycINFO is the largest, and most well-known index of the literature of psychology and the behavioral sciences. In addition to journal articles, document types indexed in the database include books and book chapters, dissertations, and electronic collections. The American Psychological Association (APA) draws on international sources, with journals in more than 29 languages, and English-language books and book chapters from all over the world. The full list of facts and statistics for PsycINFO is very impressive (American Psychological Association 2011). The Thesaurus terms (subject headings) used to index entries in the PsycINFO database are developed and applied by APA indexers.

One of the truly remarkable features of PsycINFO is the range of time it covers. PsycINFO has abstracts of books and journal articles dating from as early as 1806, although most journals are from the 1880s to the present, and most of the books are from 1987 to the present. The database is updated weekly, and the latest updated week is reflected in the search interface. Also notable is the profoundly scholarly nature of the material indexed: as noted in the APA’s PsycINFO Facts (2011), of the “more than 2,450” journals covered, “99%” are peer reviewed.

Like ERIC, the PsycINFO online database is available from several different vendors. Depending on which vendor and what backfile option your institution has chosen to subscribe to, you may have access to the complete file in one database, or the file may be divided in various ways, and you may have access to all or only to the most recent file.

Extensive Set of Search Fields

Another aspect of the PsycINFO database that demonstrates what an expertly crafted product it is (and it is, literally, crafted by experts in the field) is the set search fields used in the record structure. As of May 2010, the APA Field Guide included 39 fields, from Unique Identifier to a field for Correction Update Record (to hold the date a record was corrected, if correction was necessary). The APA appears to be fearless about making changes in the field structure, field names, and field content values, something that, as we’ve indicated in this book, is not a task to be taken lightly. The list of fields changed significantly from the first to the second edition of this book; now rather than trying to reproduce a list here, I feel the best thing is to refer you to the online Field Guide at the APA Web site—it contains the most up-to-date and complete

©2015 ABC-CLIO, LLC
information: http://www.apa.org/pubs/databases/training/field-guide.aspx (as of December 2011). The list includes searchable and display-only fields (indicated with an asterisk), and it provides helpful information about what the field is used for: content or possible values.

Some of the fields listed in the online table deserve to be highlighted here:

- Cited references (the article's bibliography; this is very important in tracing the development of research).
- Auxiliary material (indicates that there is additional material, such as audio, video, Web sites, tables, data sets, etc., separate from the source document).
- DOI (Digital Object Identifier, to help users link to full text and provide a stable URL for the item).
- Methodology (a field indicating what kind of study, e.g., clinical, field, qualitative, quantitative, is discussed. This would be of great interest to psychological researchers).
MEDLINE

MEDLINE (Medical Literature Analysis and Retrieval System Online) is a product of the NLM, which describes it as their “premier bibliographic database” containing “over 18 million references to journal articles” in the life sciences, with a focus on biomedicine. Date coverage for the MEDLINE database as offered through commercial vendors is generally from 1966 to the present, but PubMed includes material back to 1945 and even older. MEDLINE is a journal citation database (many citations offer links to full text, but the database itself does not contain full text), and it is international in scope: it selectively covers “approximately 5,516” journals from all over the world, in 39 languages. The NLM enhances records from non–English-language journals by supplying translated article titles, and “about 83%” of those records include abstracts in English. The database is updated daily, from Tuesday through Saturday, with the addition of “2,000–4,000” new citations each day. (Additions “each day” does not mean, however, that the contents of a journal published yesterday will be indexed and added today. The kind of detailed processing required to create full MEDLINE records takes time.) In 2010 alone, “nearly 700,000” references were added to the database. This phenomenal effort is the result of a distribution of labor among the NLM, its international partners, and collaborating institutions (U.S. National Library of Medicine 2011b).

A key distinguishing feature of MEDLINE is its carefully crafted thesaurus, known as MeSH, for Medical Subject Headings. (Think of it as the Library of Congress Subject Headings for the medical world.) This highly developed, hierarchical system of subject headings is also the work of the NLM, which has a whole branch devoted to continuous maintenance, revision, and updating of the MeSH vocabulary. The impression one gets is that this is a dynamic system, continually growing and changing in response to developments in scientific and medical research and practice. Statistics about MeSH are impressive: the 2011 MeSH contained 26,142 subject headings (“descriptors”) augmented with more than 177,000 other “entry terms” (U.S. National Library of Medicine 2011a). An entry term is like a see reference: if you search the MeSH database for Vitamin C, it brings up the record for Ascorbic Acid, the official heading. Vitamin C is just one of seven such “entry terms” for this common chemical, which helps to explain the huge number of entry terms versus actual subject headings.

As with any thesaurus, of course, these terms are then employed by highly trained indexers to describe and provide entry points for each article. As a user of this database, one important thing for you to remember is that indexers are instructed always to choose the most specific terms available. Although the MeSH system is hierarchical (to 11 levels of specificity . . .), with such broad terms as “Digestive System Diseases” at the most general level, if an article is discussing a specific malady, such as Crohn’s Disease, only the more specific term will be applied.
as a heading. All the layers of headings between Digestive System Diseases and Crohn’s Disease (e.g., Gastrointestinal Diseases, Gastroenteritis, and Inflammatory Bowel Diseases) would not appear as additional subject headings. So you have full permission to enter very specific terms here.

**Note**

1. In June 2011, the MeSH home page (http://www.nlm.nih.gov/mesh/meshhome.html) was quite “homey”—they were still celebrating the 50th birthday of MeSH; there was a photo of 11 members of the MeSH Staff, with a link to their biographies; and vocabulary suggestions were actively invited. Indeed, everything you could possibly want to know about MeSH is there—including a link to download an electronic copy of the whole database!
WorldCat.org

OCLC didn’t immediately produce the current WorldCat.org: starting in December 2004, OCLC first experimented with a program dubbed Open WorldCat, which only inserted “Find in a Library” pages into the results from Google and other search engines. Open WorldCat also let the search engines have access only to certain subsets of the total WorldCat database, and it operated under a number of other constraints as well (Hane 2006). Although limited in many ways, Open WorldCat evidently convinced OCLC that they should go further, and the appearance in August 2006 of WorldCat.org, a “destination website,” was greeted very positively by librarians (Flagg 2006; Hane 2006). At last, anyone who could get online could search the entire WorldCat database, using a simple, friendly interface. Each search result is linked to a “Find in a Library” information page, where the user can enter his or her zip code and receive a list of nearby WorldCat-participating libraries that own the item. Users can also link right to a library’s online catalog record to check circulation status or access electronic content directly (“OCLC Launches” 2006).

Developments and Change

Since its launch, WorldCat.org has been enthusiastically adding both technical and social features to appeal to all sorts of audiences. Since 2006, it has been possible to add a WorldCat.org search box to your personal or institutional Web pages or download Web toolbars or plug-ins for your favorite browser. RSS feeds from WorldCat.org were also part of the launch, as was the ability to add reviews and notes on individual items (“OCLC Launches” 2006). By early 2007, WorldCat.org records offered links out to related “Web Resources,” faceted browsing, a Chinese-language interface, and automatic geographic location sensing based on IP address (“WorldCat.org Adds” 2007). By August 2007, WorldCat.org was offering personalization with “My WorldCat” accounts, allowing users to build and share lists of materials as well as a personal profile (“WorldCat.org Adds List-Building” 2007). Since then, OCLC has added the ability to limit results by format and to be kept up-to-date on the contents of other people’s (public) lists by means of RSS feeds. To aid students and other writers in their quest for easier bibliography creation, WorldCat.org can format citations in any of the five common styles (APA, Chicago, Harvard, MLA, Turabian), and lets you export them to RefWorks, EndNote, or other bibliographic management software. (Want to learn how to do it? The tutorial is a video, only 1:47 minutes long, on YouTube.) And what would modern life be without Facebook and Google? To further its intention to be where the users are, OCLC
provided code to Facebook to allow users to add WorldCat.org as an application there, so that you can have a WorldCat.org search box on your Facebook page. OCLC’s relations with Google Books, a strong emphasis circa 2008, appear to have cooled somewhat: WorldCat.org results limited to Format eBook appear to provide links to Google Books only when there is no other eBook source. On the Google Books end, a quiet link to “Find in a Library,” without the WorldCat logo, appears below all the purchasing options in the “Get this book” list. WorldCat.org has its own blog, where you can keep up on the latest features, among other things. There are apps so you can use WorldCat.org on your mobile phone. You can tag WorldCat records, follow it on Twitter, or watch tutorials or “Why I love WorldCat” videos on its YouTube channel. Where the trends go, WorldCat.org is right there with them.

References

