June 2004

The URL Clearinghouse Offers Vendor URLs

Laura B. Cohen

University at Albany, LCohen@uamail.albany.edu

The URL Clearinghouse Offers Vendor URLs

by **Laura B. Cohen** (Network Services Librarian/Webmaster, LI 140 University at Albany, Albany, NY 12222; Phone: 518-442-3492; Fax: 518-442-3567) <LCohen@uamail.albany.edu>

The URL Clearinghouse is a set of instructions for creating URLs for licensed digital resources organized by vendor. It focuses on direct URLs for databases and e-journals on the title level. **The Clearinghouse** is freely available on the Web and located at *http://library.albany.edu/clearinghouse/*.

Why the need for a **URL Clearinghouse**? The availability of URL services in the form of URL lists such as **Serial Solutions**, or link resolvers such as **SFX**, hasn't done away with the need for librarians to create vendor URLs. Online catalogs, custom databases, subject pages, research guides, course Web pages, electronic reserves and other Web-based tools all employ librarian-maintained URLs that connect to licensed resources. Anyone who works with these URLs knows that figuring out how to create them poses multiple challenges. Vendors employ different types of URLs that have distinct requirements in their construction and behave differently when clicked on by a user's mouse. For example, some vendor URLs must be constructed according to a specific formula. Others may become transformed by scripts after a mouse click, rendering destination URLs that cannot be used to make a connection. These are just two of the issues that face librarians who do this work.

To be sure, there are vendor Websites that offer clear instructions on URL creation that are easy to find and to use. But this is not always the case. Instructions are sometimes difficult to locate, inaccurate, or unclear. Some vendor sites offer multiple hyperlinked title lists organized in various ways, but use different URL structures on these different lists. This leaves the librarian with an understandable confusion about which URLs to use. Even on the same list, URLs may be structured inconsistently. Those vendors that do use the recommended URLs on their sites rarely indicate this, and librarians are left to guess if these are the authorized ones or if some other structure is recommended. These and many other problems have created a state of uncertainty about doing this work. Contacting the vendor can help, but support staff isn't always fully in the know.

I had an interesting experience at this year's **Charleston Conference** that illustrates this point. I stopped by the **Project Muse** booth and asked the sales rep why all the URLS on the **Muse** site were incorrect. **Muse** uses directory URLs, yet all were missing the final trailing slash, for example *http://muse.jhu.edu/journals/social_ politics* rather than *http://muse.jhu.edu/journals/social_ politics/*. If this type of URL is missing the slash, it takes the browser and server four "conversations" to complete the transaction rather than the normal two. This causes unnecessary hits to the **Muse** servers and slows down retrieval for **Muse** customers. The rep listened with interest, and within a day she called the Webmaster to convey my question. She reported that the Webmaster had no idea why the slash would make anything better but she was willing to add it. All the URLs on the **Muse** site were corrected.

It is difficult for librarians to establish best practices in URL management under these conditions. This challenge is exacerbated by the fact that librarians tend to do this work in isolation. The library community has not pooled its knowledge about vendor URLs or provided a central repository where this knowledge can be easily obtained.

I have created such a repository. It has come out of my work of providing proxy access to licensed resources at the **University at Albany**. Determining URLs for proxy access has been enormously time-consuming and has involved a range of problem-solving techniques. Over the years I have been keeping records of how I found individual URLs and what were the rules of their construction. As my records became more extensive, I began to realize that colleagues in other libraries were working on the same issues. A centralized repository made much more sense. Out of this idea was born **The URL Clearinghouse**.

Each vendor record in **The URL Clearinghouse** provides step-by-step instructions for creating URLs. The site also includes information about the varieties of URLs employed by vendors. Librarians are invited to add vendors to the **Clearinghouse** to help make it a resource that has the potential to be widely useful. 🌳

Innovations Affecting Us — XML in Action

by **Norman Desmarais** (Acquisitions Librarian, Phillips Memorial Library, Providence College, Providence, RI 02918; Phone: 401-865-2241; Fax: 401-865-2823) <normd@providence.edu>

The software developers at **xrefer** envisioned the potential of **XML (eXtensible Markup Language)** as early as 1999. When they began to create **xreferplus**, they tagged the content with an abundance of metadata in **XML** compatible format. This metadata is fully searchable and interactive with the various components of the product. Because **xreferplus** uses **XML** compatible metadata, it is also compliant with **SFX** and similar link resolvers (see "For SFX See Librarian." *Against the Grain* 15:3 (June, 2003) pp. 102-103).

Xreferplus contains the full text of 147 books from twenty-seven publishers. Subject areas cover art, biography, business, geography, history, language, literature, law, medicine, music, philosophy & psychology, religion, science, social sciences, and technology. Technology and art, health, and science are the areas that saw the most growth in 2003. The library includes general encyclopedias like the *Crystal Reference Encyclopedia, The Columbia Encyclopedia, The Hutchinson Encyclopedia*, and *The Macmillan Encyclopedia*. The *Philip's Encyclopedia 2004* was added in the last week of January. The library also includes ten language dictionaries. There are about 76,000 audio pronunciations in MP3 format for the *American Heritage Dictionary, the Academic Press Dictionary of Science and Technology,* and *Dorland's Medical Dictionary*. Any images from the published sources are also included.

There is also a conversion engine that will quickly calculate English or metric equivalents for area, distance, energy, speed, temperature, fuel consumption, power, volume, and weight. This engine is fully integrated within **xreferplus**.

Many of the titles are from British publishers, especially **Peter Collin Publishing**, **Blackwell**, **Bloomsbury**, **Routledge**, and **Thames and Hudson**, showing the product's origins; but American titles have being added from publishers like **Houghton Mifflin**, **HarperCollins**, and **Wiley**. Content from American publishers will be the fastest growth area for 2004. **Harvard University Press** has recently signed up for a couple of music dictionaries. The individual titles are all richly indexed with **XML** compatible metadata that allows them to be fully cross-referenced. Unlike traditional cross-references which are tied to the framework of a single book, **xreferplus's** cross-references span the whole digital library to connect relevant information.

One can browse each book in the digital library, as in the print equivalent, to explore its content; but **xreferplus's** real power lies in the search engine. This engine, which resembles that for an Advanced **Google** search, permits **Boolean** searching, thematic searching, stem-