Searching and Understanding Institutional Repositories (IRs)

raditional academic publishing has been in upheaval for years over its distribution process, which detractors would describe as scholars' libraries paying large sums to buy back work that was produced for free (or produced at government expense) at that same institution. While serials budgets in libraries have remained static or declined, scholarly journal prices have increased over 7% per year for the past 25 years, according to a 2013 study from Allen Press. Academics and scholars have spent years trying to come up with effective ways of distributing free or low-cost "open access" journals. There are many varieties of open access, and many ways of implementing it. But if you want to share your research for free online, you need a place to put it and you need people to find it. You can put it on your own website, but almost no one will ever know it's there, and it will never be used.

Libraries then built or purchased "institutional repositories," or IRs: organized and curated web servers where the institution's authors and researchers could store and share all kinds of content: data sets that underlie their published work, early versions of the work ("pre-prints"), work they didn't want to distribute through for-profit publishers, or authors' reprints from open access journals that allow redistribution through other channels. Much of this work is peer-reviewed, scholarly, valuable content; when used by others, it will add to the historical or scientific record. But a discovery problem still remains: how do you know which, from among thousands of extant IRs, holds information of interest to you?

Librarians came up with several solutions for that problem. One was a standard for sharing descriptive metadata (the type of information you might find in a library catalogue entry, such as the type of media, author and title, and where the item is located) about the contents in their IRs. Using this standard, called **OAI-PMH**, information about many IRs can be searched at once. Multiple sites search this metadata; the most common is **OAIster**, available at http://oaister.worldcat.org/. OAIster

searches the metadata provided from thousands of IRs, and numerous types of content: digitized books, images, theses and dissertations, data sets, audio and video files, and much more. One significant benefit of IRs is that they become a place where authors can share their articles outside of the journals in which they were published. This allows readers without access to the journal itself to access and read the author's work. Of course, this only occurs if the journal allows it, and if the author has shared a copy with the IR managers. But it is becoming more common, and can be a good way of very easily getting at content not otherwise available.

While OAIster searches just the metadata, many tools nowadays allow for searching the full text of a collection. OpenDOAR, at http://opendoar.org, is the Directory of Open Access Repositories. It allows one to both search for repositories, and search through repositories. In the first case, one might find repositories that focus on a specific subject, event, person, location, or other topic, via http://opendoar. org/find.php. But one can also search through the full contents of these collections, using a specialized Google index of the databases listed, at http://opendoar.org/search.php. A sample search for "schooner" there returns great images from Louisiana State University and the Metropolitan Museum of Art, plus a history of the Amistad insurrection from UNC-Chapel Hill, part of a 1903 MIT thesis on the strength of the schooner John B. Prescott, an 1863 article about the loss of the schooner Alice Webb to Confederates off the North Carolina coast, and more.

Searching IRs remains a challenge, and there are many opportunities for improvement in this area. For folks without access to institutionally-provided subscription databases, however, they are a great way to find reliable, useful, and scholarly content.

Suggestions for other sites worth mentioning are welcome at **peter@shipindex.org**. See **http://shipindex.org** for a free compilation of over 150,000 ship names from indexes to dozens of books and journals.

\$\ddot\text{\$\text{\$}}\$



