PLAN FOR A CO-OPERATIVE IAMSLIC MARINE SCIENCE BIBLIOGRAPHY

Charles A. McFadden
Virginia Institute of Marine Science
School of Marine Science
College of William & Mary
Gloucester Point, VA 23062

ABSTRACT: A plan is offered describing a way to form a general marine science bibliography of books and articles current and retrospective by pooling member OPAC records, local bibliographies and journal tables of contents. The bibliography would provide free one-stop shopping marine science bibliography, and would be owned and managed by IAMSLIC, though all the work would be done by VIMS and other interested parties. IAMSLIC, our professional organization, would in effect become an information producer and the world locus of marine science bibliography. The product is a file of records in uniform format. This access independence enables the bibliography to be presented in whatever the access du jour. Since a pooled, centralized bibliography appears to conflict with the current distributed model offered by the Web and the Z39.50 protocol, an argument will be given to show how it does not, but rather maximizes some of the Web’s overlooked potential. By returning to the bibliography business, librarians find their true niche as primary intermediaries between the published paper and the consumer.

An Archive of Pre-existent Records

Think of all the hundreds of electronic files in the scores of marine science libraries worldwide; think of the hundreds of thousands of records, ranging from OPAC records carefully maintained to local bibliographies prepared with love to grey literature available nowhere else. Think of the exhaustive marine science bibliography these would form if centralized and standardized. Since a large number of these records would be fielded with traditional citation fields, the task of standardizing the format and loading them into management software is lessened. The larger files, for instance, the OPAC records of three large libraries, could be loaded fairly quickly and would form the foundation, to which would be added in time smaller OPACs and more difficult files. This would take care of the books and some articles.

A comprehensive bibliography also has to cover journal articles. The best way to approach this is to identify appropriate journals and try to find electronic versions of their tables of contents. The first step would be to try to negotiate with CARL and other table of contents providers, and determine whether files could simply be copied. CARL offers its bibliography free to advertise its document delivery. Perhaps, if we included in our copy of their records a link to their document delivery service, we could extend their advertising and develop a quid pro quo. Another possible source for pre-existent table of contents files is the Web, where a number of publishers offer the service free. Perhaps they could be persuaded that our inclusion of their records in our bibliography would increase the demand for document delivery.
The original CARL proved that an association of librarians can create a table of contents bibliography. Perhaps we could imitate their methods and also try to maximize developments in scanning to create files of records for tables of contents of back issues not yet in electronic format. If we could discover an efficient way to create these files, perhaps we could offer these back files in exchange for CARL’s current files. A tremendous strength of IAMSLIC is that some member library has the table of contents page to every issue of every appropriate journal. Since the table of contents of a journal is not copyrighted, the pages could be put in electronic format.

Initially we would not fret about duplicate records. When the file collection phase neared an end, we would develop methods to dedupe the records. The fact that several libraries own the same book or that several local bibliographies include the same citation is useful information and would be retained in the 035 MARC subfield b by means of a three letter code. If a record has several of these codes, it denotes that this bibliographic record has a relationship to several files or libraries. Those relationships could be searched on, or displayed in a holdings-like arrangement. Value could further be added to the records by editing in links to the full text URL if that becomes available.

The Message is the Message; the Medium is Whatever

A weakness of many bibliographic projects is that they are medium dependent. They exist only in paper or only in microfiche, or can only be accessed electronically at one site. Our bibliography will be medium or access independent. It will be a file of records in uniform format, which can be searched at a Web site, but can also be delivered by FTP for local loading in part or in whole, or can be included on a CD-ROM indexed with search software, such as CDS/ISIS, or can be included with software on an external hard disk. There could be paper copies. The message will not be confused or identified with the medium.

IAMSLIC Proposes, VIMS Disposes

IAMSLIC would own the database, make strategic decisions for long term planning, and determine the method and conditions of distribution. VIMS and any other interested parties would do the work. VIMS has BRS/Search Full Text Indexing Software on a Sparc20 with a Web interface (http://www.vims.edu/library), in which are loaded ASFA 1988-, Water Resources Abstracts 1967-, VIMS OPAC records, VIMS journal records, the Chesapeake Bay Bibliography, the Chaetognath Reference Library, VIMS Contributions List, and an experimental Elsevier Table of Contents File. ASFA and WRA are excluded to users outside VIMS. Filter programs were written in Perl to convert the various formats to a standard format acceptable to BRS. This serves as a proof of concept for the IAMSLIC bibliography, blending several databases with varying formats into one concatenated database available on the Web.

There would be an endless supply of sub-projects for librarians interested in doing bibliography this old fashioned way: the pursuit of an efficient scanning method, identifying and soliciting appropriate files for inclusion, deduping, editing. With the Web interface it would be possible for librarians anywhere to work on the bibliography at VIMS.
But What about Z39.50?

A centralized bibliography is not “instead of” the Z39.50 databases distributed over the Web; it would be one of them. However, at a time fast approaching, when disintermediation is almost sure to happen, an organization such as IAMSLIC must choose whether it wants to be a host or a parasite. Disintermediation means the technology will remove the need for intermediaries between the producer of information and the ultimate consumer (Weingarten 1996). The traditional role of libraries as intermediaries will diminish on a Webbed planet where geography disappears (McKinzie 1996) and the client is not the person who walks through the door, but the consumer in front of his personal computer. The informational value of an organization will be the product it offers the Web audience.

Some librarians see their future role as catalogers of Web resources or collectors/organizers of URLs. But this is indeed risky business, placing them in direct competition with for-profit corporations, which smell huge profits in an easy-to-use Web and have correspondingly vast amounts of development capital. The incredible progress in Web search engines gives witness to this. Disintermediation will force librarians either to produce information or to consume it like the rest of humanity. Organized lists of URLs are useful at this point in time, but are essentially parasitic and present no dependable future for libraries.

The argument will be made that it is equally impossible for libraries or library organizations to become producers of information: how can we possibly compete with ASFA, NISC, CARL and OCLC? The answer is that before these bibliographies came to be, we were. Librarians are by definition bibliographers; it is our lowest common denominator and our highest common denominator, the daily job we all do and the noblest function we all perform. The answer is that CARL and OCLC were both created and developed by librarians. The answer is that the UNESCO funded ASFA contribution centers could be giving their bibliography directly to IAMSLIC as the minder of the marine science bibliography. The answer is that we even let NISC sell our own bibliographies back to us, when with a little more spirit we could be exchanging them with one another or centralizing them for access by all.

What entity is better suited to be an information producer than a professional organization of 250+ highly trained, subject specialized, strongly motivated, closely knit marine science bibliographers, from every corner of the globe, functionally united by the Internet? IAMSLIC could become the global locus for marine science bibliography simply by choosing to, by awakening from the cybernetic slumber induced by the exaggerated expectations of the Web search engines. A questionable assumption underlies these expectations: everything is already out there; it’s just a matter of developing the software to get to it. Structuring data is no longer necessary; sophisticated search engines will be so powerful that they will be able to find what you want from the whole universe of unstructured data.

This assumption is probably unprovable either way. I would argue that the preponderance of human experience says that it is more likely the data need structuring. A professor would not accept a student’s notes in place of a paper. It is too big a leap of faith to declare that the data need no more structuring, that the age of structuring data is past. If one queries one’s soul and the answer is: “More structure!” a world of opportunities opens up. If the answer is, “No need,” delete here.
If we opt for structure, if we believe that we need to return to and consummate the unfinished agenda of the Seventies and Eighties, such as shared OPACs, local bibliographies and union lists, we see the computer scene in a different light. Instead of regarding the Webbed computer as a TV set, carrying us far off to distant information sources, we view it as a file transfer machine, an incredible tool for gathering files and updates. Faster processors and huge hard disks are no longer seen as the hardware resources required by bloated software, but as the enablers of giant resident databases for the common man. Web search engines are viewed not only as a way for locating resources, but as the software development needed to make search engines in general cheaper and more powerful. Cheap computers already have major league speed and hard disks have major league space. What is needed is cheap, major league search and indexing software.

A professional organization of bibliographers, using the best modern computing tools to become an information provider on the Web, is not an unreasonable hypothesis to test.

What if Garfield was Right?

The information scientist, not the cat.

By recommitting ourselves to the bibliography business, we librarians are reintimidated rather than disintermediated. Although bibliographies are themselves information products, they really are the lowest level and most important intermediaries between the published paper and the consumer. Bibliographers are intermediaries but also information producers. This is our niche.

There is much unfinished business in bibliography. So much pre-1988 bibliography is not in electronic format. A more intriguing issue is whether there are implicit relationships between papers which need to be made explicit. These relationships can be subjective or objective.

An example of subjective relationships is Paula Wolfe’s Selenium bibliography. A person with expertise in a particular field pokes around among the vast universe of published papers and subjectively discovers the specific papers related to a topic, such as selenium. Since the papers are related to the topic, they are related to one another. The relationships among the papers would not be evident for other scholars unless the person/subject pointed them out. Radical Websters might say that there is no need to pull these records together since they are already “out there.” Bub bibliographies such as this one have always been recognized as useful since they add structure to the data.

Examples of objective bibliographic relationships are tables of contents bibliographies and citation indexes. This is territory explored by Eugene Garfield, who started Current Contents and Science Citation Index, modeled on the law citation indexes. These types of bibliographies exploit the objective relationships of papers to the journal issue they are in, and papers cited to the citing paper. In the latter case, search links are formed between citer and citee. They are bibliographies whose matter is not the universe of papers but other bibliographies, and which do not require subject expertise to exploit. They add a different type of structure to the data.

Value can be added to our bibliography by incorporating both the subjective and objective relationships among records. One field can include the subjective bibliographies which include this record. Another field can contain links to the papers cited by this paper. A third field can contain a link to the papers which cite this paper. A user could reassemble a paper’s bibliography
by activating the links or could view a particular journal issue's table of contents by searching on the issue. Or he/she could limit a search just to the papers in the Selenium database.

Either the Web search engines will eliminate the need for bibliography or they will not. If we suspect they will not, there is an infinite amount of interesting bibliography to do with the marvelous tools modern computing offers.

References
