The Evolving virtual library II : practical and philosophical perspectives, p61-83. ISBN 1-57387-070-6

Library-Generated Databases

Tore Brattli

Academic Librarian
University Library of Tromsø, Norway

The development of the Internet and the World Wide Web has given libraries many new opportunities to disseminate organized information about internal and external collections to users. One of these possibilities is to make separate databases for information or services not sufficiently covered by the online public access catalog (OPAC) or other available databases. What's new is that librarians can now create and maintain these databases and make them user-friendly. Library-generated databases can be available to users at home and at the office. In addition, these databases can become powerful information services by linking to other databases, services, or information on the Internet.

This chapter will focus on the organization of information that can be done by librarians locally in each library. The emphasis is primarily on bibliographical references. As I examine why local organization is still important and why internal databases are the best tools for much of this work, you will also see how internal databases can be created and made accessible on the Web. In addition, several cases will be presented. The aim is to become aware of some of the useful things a local library can do with the new information tools that have arrived. The view is from an academic librarian at the University library of Tromsø in northern Norway.

Organizing Information in Libraries

These are the main tasks of academic libraries:

- Have collections of documents locally available
- Organize documents so that it is possible to be aware of and retrieve the relevant ones
- Educate and help users so that they can find the information they need
- Archive documents so that they do not get lost for the future.

Traditionally, much of the organization of documents has been done locally. Librarians have systematically arranged local documents on the shelves, maintained a primary catalog, and created bibliographies and specialized catalogs for specific purposes. It's my impression, based on what I have observed in libraries with which I am familiar, that this work has been declining in the latest years, and there are at least two reasons for this. More and more of the large-scale organization of references is done externally. For journal articles, this has been the main rule for a long time with indexing and abstracting services on paper, CD-ROMs and online databases. For books, it has changed more recently with libraries sharing OPACs and downloading ready cataloged and indexed records from external sources. On the Internet, search engines and subject indexes are dominating, and most librarians have not entered this arena yet. Generally, it seems that the organization of references has shifted from local to national, and even to international.

The other reason is that it has been difficult to compete locally with the technology of large networked reference databases and OPACs, even if librarians were to have something

interesting to offer. Users are not very keen on using paper bibliographies or catalogs anymore, and so far the tools for creating and publishing user-friendly databases have been too difficult for the average librarian to use. Large scale organizing is efficient and gives access to large amounts of information, but unfortunately for many librarians, it has reduced much of the organizational responsibility to just routine registration.

The time has now come to turn this trend around. For libraries, the task of organizing information also means supplying different groups of users with tailor-made relevant bibliographies and special catalogs. As we will see later, it's impossible or at least very difficult for large external databases to take care of every need for information in each and every library. For libraries that want to offer a first class service, some organization still has to be done locally. With the Web, librarians finally have the information tools that make it easy to create powerful, user-friendly, accessible, and tailor-made databases that can be important supplements to the large reference databases and OPACs.

The Future

In the library of the future, the virtual (digital) networked library, there are strong indications that collection and archiving will be less important, at least in each local library. With networked information, there is less need to store local copies, and there is a shift away from collections and toward access. This might imply fewer libraries and less information in each library. Of course, other interesting opportunities exist for a library to collect documents. Some suggest that collecting and publishing local unpublished information may be an important future role of libraries (Webb, 1997). Others see the library as the publisher of local research material, bypassing today's publishers. However, this will not produce the same volume of documents as today. When libraries do not collect as much, there is less need for local archiving. The task of archiving will probably be assumed by regional and national archives or libraries in each country.

When it comes to organizing information and educating users, everything depends on the levels of motivation and skills available in the library for doing such tasks. One thing is quite sure: both tasks—organization and education—will become more important in the future. The amount of information is increasing very quickly, and users must be able to determine and locate relevant information for themselves. With smaller local collections, there will always be more limited organization of documents even though identification of their locations will be a part of the large-scale external organizational efforts. What is left is the task of serving our primary users with tailor-made information about documents relevant to their work. This important task will become even more crucial in the future. The tools for accomplishing this are databases with Web access. We must learn to create such databases and reaffirm our professional reputations for excellence at small-scale organization skills. In addition, contact with the users is important. If the users don't come to us, we have to visit them.

Publishing Library-Related Information on the Web

The World Wide Web is the most recent piece of a large technological puzzle, and it will revolutionize the library's opportunities to serve its users with information. Already many libraries offer a significant quantity of information on the Web. So far what is available is mostly information about information (i.e. metadata or references) and general information about the library and its collections. However, the revolution has just started. Digital journals are emerging in increasing numbers, and digital books are probably not that far behind. On the Web, information is published as Web documents (or pages). There are basically two different methods for generating Web documents, and both are equally appropriate for libraries to use. We may call them "static" and "dynamic" documents.

Static vs. Dynamic Web Documents

A static Web document is what most Web users would regard as the "normal" Web document. It consists of text, graphics, and HTML code—all stored as ready-to-use files on the Web server. When a user requests a static document, the files are sent to the user's Web browser and displayed. Static documents are normally best suited to information which is not easily listed and which does not change very often. Typical information to publish as static documents would be general library guidelines, information about the collection, hours of operation, user education, lending/borrowing rules. Of course, full-text documents are static documents as well.

Dynamic documents are the opposite of static documents. These can be generated on the fly in response to a request from a Web user. Generally this means that dynamic Web documents did not exist before being requested. Dynamic documents are often the result of a database search, made by a computer program, based on a data file and a search string. Such dynamic documents are what I envision as the appropriate format best suited to publish local library databases.

External vs. Internal Databases

Library databases can also be divided into external and internal (or local) databases. The distinction between the two concepts is not always quite clear, and what seems external to one person may seem internal to another. It actually depends more on the user's purposes than on the intrinsic structure or information in the database.

External Databases

External databases are databases that are externally operated in a way that the local library or librarians have little influence on content or functionality. Typical external databases (independent of medium or location) are the large reference databases for journal articles like Medline or Science Citation Index. OPACs are often regarded as external databases, but this of course depends on the actual solution. For libraries collaborating with other libraries about the OPAC or for libraries using commercial OPAC systems, this may often be the case. External databases—whether specific by program or content--are often more library independent than internal databases.

BIBSYS (http://www.bibsys.no/) the Norwegian national OPAC for universities and colleges, was actually one of the first in the world available on the Web (in November 1993). Since then, many other OPACs and larger reference databases have become available on the Web. Many libraries have developed reference databases that are integrated with their OPACs for holdings and online article ordering. Linking to full-text articles has also begun. Together with static documents these Web services can offer users nearly all the information they need about the local library and its collections, in addition to information about books and journals in general.

Internal Databases

Most libraries have so far been publishing information as static Web documents and as elements of external databases. In addition, practice often shows that there is a need for a third format: library-generated or internal databases. Internal databases are databases made within the library in a way that the library has more control over content and functionality. Typical internal databases may be catalogs of the library's journals or video collections, or bibliographies of famous local authors. Internal databases are often more library dependent or more tailored to a specific group than external databases. They are better suited to cover local specialties. For most libraries, internal databases will be an important supplement to external databases. Yet as a tool, internal databases for the library are a bit like bibliographic management software such as ProCite or Reference Manager for scientists.

What Is Special About Internal Databases Now?

Of course, internal databases are not new (Biggs, 1995; Brudvig, 1991; Raeder, 1989; Smith-Cohen, 1993). Librarians are trained to realize that objects can be described in terms of their structural commonalities. The best way to organize these objects is to make bibliographies or catalogs (i.e. "databases") of the descriptions. Therefore librarians have made lists, i.e., bibliographies and catalogs for books, journals, and articles and then made these lists available for users. This has always been one of the more important tasks for librarians, and still is.

With the introduction of computers, however, it seems that librarians have been inhibited, or at least limited, from taking part in the latest bibliography and catalog developments by difficult-to-use-technology. This has been especially noticeable in the cases of smaller or tailor-made applications, requiring help from computer experts.

The news now is that along with the Web there have arrived information tools which make it fairly easy for librarians to create internal databases and for users to access them. The databases can be made more user-friendly, and together with the possibilities for integration with other information, increase the chances that the new databases are really used. Of the four improvements mentioned, the design and use functions have been more evolutionary while the database access and integration through the Web have been truly revolutionary.

Easy to Create, to Access, and to Use

We no longer need to be computer experts to create user-friendly Internet-accessible databases. With minimal training most librarians can now create and maintain relatively advanced Web databases. This allows them to focus more on content than technology and to create databases for particular purposes. This is especially important for smaller libraries and even for librarians in larger libraries who do not have access to support from computer consultants or management. The creativity can finally grow again!

With the worldwide Internet, it is easy to make the databases accessible wherever the users are, whether at home or at the office, and twenty-four hours a day. The databases are now user-friendly in every way—in the installation of search software, in the searching, and in the presentation of information.

Integration with Other Information and Services

Last but not least, internal databases can function as powerful information services by integrating them with information and/or services on the Internet. Records in an internal database can easily be linked to records, searches, or services in other databases or to external information or full-text documents on the Internet.

At the local level, an internal book database entry for a specific title may be linked to its associated records in the OPAC, or to the publisher's pages about the book. This gives the users bibliographic information, location, loan status, loan order, table of contents, and description of the book. In a database of journal entries, titles can be linked to the OPAC for bibliographic information and article ordering, to the publisher's pages for table of contents, or to related reference databases. A locally generated Shakespeare bibliography can be linked to external full-text materials and to other information on the Web. A thesaurus or a classification table (databases of classes or subject words) can be linked to subject searching in reference databases and can be an alternative search interface to both internal and external databases.

Adaptability Is Essential

In many cases the databases that are desired locally would have been best realized as an available OPAC function. The information is often already within the system, but it is not

possible to access or present. The problem is often lack of functionality in OPAC software along with too little interest or financial incentive to improve it. An example of this in the case of BIBSYS, our Norwegian national OPAC for universities and colleges, is the system's inability to search and present a list of recently purchased books. All the necessary information is present but it isn't searchable. Instead, more and more local libraries must spend time and effort making their own internal databases in order to obtain this frequently required information. For instance, we wanted from BIBSYS a sorted subject list of our library's journals. Yet this need was not met, if indeed it was even recognized as a legitimate need.

The solution, of course, is that external databases should be made more flexible when it comes to searching for and presenting information. Still, it's not possible for an external database to cover every need in every library. One problem with large databases is that they cannot do everything wanted (Spore, 1991; Gates, 1989). In order to satisfy many users and libraries they often end up as the lowest common denominator of all needs. Only the most basic functionality and information fields are offered. Even if they contain many records, there are often limitations in searching, presentation, and information that each record contains. The result is disarranged quantities of valuable information, which cannot be suitably extracted.

Some of the problems with large databases are as follows:

- High flexibility in searching, sorting, and presentation demands many indexes which in turn are processing and storage intensive.
- Functionality and data structure is difficult to change because of the size.
- Many options can actually be user-unfriendly (at least for the average user).
- Large databases tend to both hide relevant documents and to retrieve many irrelevant documents, especially for subject searching.
- It's not possible to reflect all the different aspects of each document in one database.

Because of the needs and problems mentioned, it is not possible, practical, or economical to satisfy every need in one database. In addition, large databases are often externally controlled and operated. Experience often show that this makes it difficult even to convince the database producer that your library or users actually have a "need" at all. The adaptation of external databases to local needs is difficult because the databases often are large, complex, and externally managed.

Compensate for Limitations in Other Accessible Databases

The main reason for the library to want to build its own internal databases is to compensate for limitations or defects in other accessible databases. For example, suppose it was simple and flexibly easy to download information from external databases into internal databases. This would better take care of local needs. With powerful PCs and fast networks, even regular downloading of whole portions of the OPAC should be a possible option. An example of such a large scale downloading is the Computer Science Library at the University of Oslo, which every week receives a copy of all of its records in BIBSYS in order to update its internal database (Hegna, 1994).

Another compensatory local adaptation which we can all appreciate is the management of references. The best tool to make collections available is databases. Static Web documents are more suited to general textual descriptions. For references, the work of creating and updating presentations is demanding, and searching must be limited. The users are not a homogeneous group, and the need for information will vary with different groups of users. The challenge for the library is to make all kinds of collection options available for all different groups of users. Again, adaptability is essential.

The best tool to register, organize, update, and disseminate detailed information about collections is a reference database. In a reference database the documents are represented by surrogates (references) which only contain key information about content and location, and therefore take little space. This makes it feasible to create databases covering most documents published in a subject area or contained in libraries. Reference databases make it easier to become aware of relevant documents without having to read all of them. This is, of course, especially important for external documents. With the Web, references are becoming increasingly more important because of the possibility of making links to the documents they are representing, wherever they may be. Most academic libraries today are fairly well equipped with reference databases for journal articles and an OPAC which includes books and journals.

The most cost-effective way to create reference databases is to make them large in terms of number of records, and they should have a general functionality. This makes them useful for many users, and many libraries can share the work and/or operational cost. Examples of this are large reference databases like Medline, or OPACs like BIBSYS. These databases are important and powerful tools to extract relevant documents from the enormous amount of irrelevant material at a fairly inexpensive price. We can divide the limitations of external standardized databases into (1) information not possible to include, (2) selections of records not possible to search for, or (3) information not possible to present as desired. These three categories are not quite independent of each other.

Information Not Possible to Include

There may not be a commercially available or free database that already includes the particular information that is desirable for local needs. Though most established databases contain records and fields developed according to standard criteria, this also means that there are references to documents and information about documents that are not useful at the local level, although they meet the criteria. Two examples from BIBSYS that I have found involve searches of Web documents (records) and abstracts (fields).

Information Not Possible to Search For

Generally speaking, when you are searching, the database returns the selection of records containing the information specified in the search string. If this information isn't included, indexed, or possible to combine with searches in other indexes, you might well have an unsuccessful search. An example for me from BIBSYS is my attempt to limit the search to new journal issues received last week. Even though the date information is present and searchable, it may be difficult to make the database return the relevant selection of records exactly. For other instances of subject searching this is often the case, as well.

Information Not Presented Well Enough

Many databases offer very limited ways to present the retrieved selection of records. What is needed is a report generator, like those provided in most modern PC-database programs, where the user has nearly unlimited options and can:

- Choose which information to present from each record (fields, parts of fields or combinations of fields).
- Create layouts where the fields can be placed.
- Supply headings and extra text.
- Control typography.
- Control sorting.
- Send output to screen, printer or disc.

Most external reference databases unfortunately offer just a few of these features. If the limitations are serious enough, and there is no simple way to solve them, the solution may be to create internal databases to improve the situation. If you do not get what you want, you can now more easily make it yourself! Together, static Web documents, external databases, and internal databases will complement each other so that the library can offer a complete information service on the Web.

Advantages with Internal Databases

I have attempted to demonstrate that one of the advantages of internal databases is the opportunity to serve specific users, groups of users, other librarians, and yourself. The needs of every science project or class with tailor-made expectations can be met with detailed information and functionality. This means that if there is a need for an alphabetical list of the library's journals and specific issue holdings, you can search out just that. With internal databases, it is possible to have more control of information, searching, and presentation. The use may be simpler, the noise (irrelevant information) lower, and the benefits higher.

Internal databases are also useful for promotion of specialties that often disappear in large databases. If your library have some interesting old books or some new and "hot" titles, it is difficult (and not acceptable) to tag them in the OPAC as "old and interesting" or "new and hot." With an internal database containing only old and interesting books, this is no problem. This is one of the reasons why many scientists have their own personal reference databases.

Each year libraries spend a lot of money on literature, but too much of this is seldom or never used because it had a tendency to disappear into the larger library collection. Through special promotion of different aspects of this literature—like famous authors, interesting newcomers, classics or subjects of current interest—there is a great chance of increased use of such material. Such promotion is well documented for displays in the library (Baker & Lancaster, 1991), and perhaps it should be as effective using the Internet.

Also, collections other than books or journals can be organized with the help of databases. For some collections internal databases are the only option. Consider these possibilities:

- A database of pre-formulated searches for literature relevant to the most popular term-paper subjects in a primary course. For example, each search can be linked to the OPAC by subject, which then would always provide updated lists of books. This may save time and effort both for students and librarians.
- A database of (UDC, LC or Dewey) classes and descriptors linked to classification searches of the OPAC. This may offer an alternative search interface and access to something that users usually find difficult to work with.
- A database of FAQs (Frequently Asked Questions) about the library and services for users who want to learn about the library.

Benefits for Librarians

In addition to making collections available for the users, internal databases are great tools for academic librarians to manage their collections when the OPAC doesn't support this sufficiently. In fact, for several of the internal databases made at the University library of Tromsø, we feel that this collection support is reason enough for developing annotated databases. The extra added fields for comments, markings, prices, dates, use, etc. can be searched and presented in nearly any way needed to make the informed decisions. Of course, much of this information can be downloaded each time from the OPAC and reformatted with a word processor to make the desired lists, but often such processing is more labor intensive than creating and updating a duplicate database.

Each year, an academic librarian at our library may purchase about a thousand books and subscribe to some five hundred journals. Because of large variations in ordering time for books, journal prices, currency, budget and feedback from the accounts department, the task of managing the budget can be a real challenge. Our internal databases for journals (with prices) and acquisitions of books (with order dates) are of great help in this work.

For collection development, internal databases can be used to divide acquisitions as fairly and professionally as possible among the different subjects and scientists. For books, we can more easily check how many are purchased in each subject. For journals, the database can produce user-friendly lists (sorted alphabetically, by price, or by subject) in order to evaluate the start of new and cancellation of existing subscriptions. With these lists, it's also easier to take advice from the scientist. The database can also prove important for operations like binding and discarding. The main advantage for journals is that a database can be used as a searchable notepad for assessments where the decisions often take much time.

A side effect of having internal databases is that the data can easily be examined or used for library statistics. It's also possible to register more information about each object than the main purpose of the database demands, in order to derive more interesting results. Many OPACs have surprisingly few and limited statistical features. Examples of "research" information available from our databases are journal prices, book processing time, use, and the effect of library displays.

Internal Databases as an Information Tool

More and more reference sources and documents are now available in digital form on the Internet. This is most noticeable for bibliographic reference databases, but journal access is now increasing as well. We already see online reference databases linked to online journals. For books it is more difficult to tell because most people have an affection for books in their traditional form. As I see it, the success of digital books is highly dependent on screen technology. With cheaper, lightweight flat screens which are readable like books and use battery and mobile communication, much may change. It's likely that the number of digital books will increase in the future.

Another aspect of this situation is that much information is bypassing the library on the way to the users. For the library an important question is how this information should now be organized when it may never be present in the library. Databases should play an important role in this organization. Reference databases will be the main tool for organizing information about documents, in addition to becoming the virtual "bookshelves" for the same documents. Where the documents are located physically is of less importance as long as they are accessible. Large external databases will act as the general collections and internal databases as the tailor-made or special collections. The key to virtual libraries are reference databases (Morgan, 1998) and access to documents. This should imply that creating databases will be a substantial task for the librarian of the future. It is therefore important to learn the skills needed to organize information with the new information tools that have arrived. One way of doing that is to start creating internal databases.

How to Create Internal Databases

For librarians with a little interest in the Internet, there are several ways to create and publish their own internal databases. In this chapter a database is defined as a collection of related data, together with tools that can retrieve them effectively. It is important to choose the right tools and techniques to solve the problem. Simple solutions that many computer experts would call nonprofessional may still be the best. There is nothing wrong with that! There are big

differences between large OPACs or other reference databases, and most of the internal databases that are useful to libraries. The former may consist of more than a million extensive records, while the latter range from a hundred to ten thousand brief records. These differences have major implications. Smaller databases do not necessarily need indexes, advanced relational structures, complex record formats like MARC, or large thesauri to do what they are supposed to do. Searching can be done sequentially, and there is no need to worry about storage capacity or redundancy of information. Even search fields are not necessarily needed. User-friendly searching can be done effectively by choosing (browsing) from menus with pre-formulated searches, instead of using search fields with more or less complex search languages.

Content and Structure

The first decision is which information to include in the database, and how to structure it. This depends on the desired functionality and the possibilities for utilizing external information. Quantities of highly structured information allow greater search options at the cost of more updating and higher computer capacity. The easiest way to include information (at least virtually) is to link the database to external sources. This reduces the work of creating and updating the information to a simple update of its address. However, it is also necessary to have some information locally. The database must at least contain the information needed for local searching, sorting, presentation, and linking. A basic but sufficiently operational internal database may contain as little as a title, subject, and a link field.

Linking

There are three different ways of linking internal databases to external information. First, we might link to static Web documents, database records, and database searches. The last two, of course, are basically the same (i.e. links to searches). A link consists of a link text, a link address (URL), and some HTML formatting code. All three must be included, if the database is going to be able to produce a Web document with a real link. It's only necessary to store the variable part of the link text and URL in each record. The rest can be stored as a global constant. From this, the database can generate the complete link by merging the variable and the constant. Sometimes the variable is already included (e.g., title linked to OPAC records using ISBN); otherwise, it can be included in a separate (link) field. A second option is to link to Web documents by storing the complete URL. For records, it is often best to use unique identifiers like ISSN/ISBN or external record-ID. Links to searches give interesting opportunities and can be made from the actual search string or other already included fields. An example of the latter is to link each record of an internal journal database to search for articles in a reference database using ISSN. Unfortunately some information providers do not allow you to take this "backdoor" to their information (i.e. bookmark it). They want you to use their "main entrance," "shelves," and organization. This makes local adaptation difficult and reduces the availability of the information if your users will not be able to initiate fully the indirect route to the information on their own, using your path directions. Every library should demand the right to bookmark information they otherwise have access to, at least for purchased material.

Loading and Updating

In most cases, some of the information in internal databases will be a duplicate of information in OPACs or other reference databases. To avoid updating several places and/or to reduce the work, the best solution is to use a program that more or less automatically can extract information from external sources and update the internal database (Knudson et al., 1997).

One possibility is to use traditional downloading of a text-file which can be imported to the internal database. For a medium-sized library with modern PCs it is even possible to download the whole OPAC regularly (100,000 records at 2Kb each; is only 200Mb), if there is an option

for this. Another way is the principle used by Web search engines, which regularly fetch Web documents (one at a time) from which they extract information. A more simple solution is to fetch records manually and copy (and paste) the relevant information, field by field, into the database. If the information is not available in digital form, typing or scanning may be the only solution.

In practice, a combination is often necessary. The degree of automation will usually depend on the possibilities to extract information from the original source, the possibilities to import this information into the internal database, the local programming skills, the size of the database, and last but not least, the rate of change in the information. There is no need for an advanced updating system for a small database with static information.

Searching and Presentation

The advantage with smaller internal databases is that they can be presented in a YAHOO-like way, using menus instead of, or in addition to, search fields. This is often simplest for subject searching. Using this principle, each Web document produced by the database is actually the result of a search and presents the list of hits (often titles) together with a list of other related pre-formulated searches. The searches should be chosen so that it is possible to retrieve each record, without the number of hits for each search being too high. For larger databases, this means many pre-formulated searches, which are then best organized in a hierarchical way. The clue is, of course, to index the records with the pre-formulated searches, which probably have turned into a thesaurus. In addition to menus, it is useful to have a search field. For quick lookup, the list of titles can be presented with holdings or call numbers. If more information is needed, the titles can be linked to local or external bibliographic information, full-text, more information from the publisher, etc.

PC Databases and HTML Files

As I have already stated, there are many ways to create internal databases. The solution depends on the size, functionality, computer skills, external help, financing, etc. One of the simplest ways to create and publish a database is to make a static Web document with indexes and a list of records, just like the good old bibliographic reference indexes or the phone book. Such a database has no search possibilities beyond the Web browser's "find in page" function, the indexes, or the order of the records. This is a perfectly satisfying solution for a smaller number of records, maybe up to a total of one thousand or a few hundred per Web document.

The best tool to administer, maintain, and format this kind of information is a PC database program. It provides user-friendliness and flexibility for the librarians operating the database. When some of the information is updated, the program can automatically find the right selection of records, sort them, format the right fields from each record with HTML code, supply adequate headers and footers, and export this to the Web server as ready-to-use Web documents (Delfino, 1996). This is the same principle used by producers of bibliographic reference books after the introduction of the computer, but now both the tools for making them and the medium for transmitting them are more effective and user friendly.

PC Database, Text File, and Perl Script

A more advanced solution is exporting the content of the PC database to the Web server as a presorted text file and using a CGI script to generate Web documents from this file. A typical text file format consists of records separated with paragraph marks and fields separated with the tab character. CGI scripts are small programs on the Web server (server extension programs), often programmed in Perl (Perl script), C or shell scripts, that users can start up and transmit information to. Perl is especially suitable because it is a high level interpreted programming language with powerful abilities to search for and manipulate text. A Perl script can receive and

interpret the search string, search the text file for matching records, format them with HTML-code, build a Web document, and return it to the users. Depending on the size of the records and the number of users, this kind of database works very well for up to 10,000 records. With a little programming skill, this is a relative simple and flexible solution (Knudson et al., 1997; Zollman & Zollman, 1997).

PC Database on Web Server

Maybe the best method, considering its simplicity, is to get a multi-user PC database program that can execute at the Web server. With a program like this, it is easy for us computer amateurs to create and maintain the database program, user interfaces, data structures, and automated routines. The information can be simultaneously updated by several librarians, and users can have direct and immediate access to the information. Several programs with these properties are available, and for some you can even define your own PC as the Web server. FileMaker Pro 4 (Langer, 1998) is an example of a general purpose, user-friendly PC database program which works very well for all the three methods mentioned so far. Recently personal bibliographic management software like ProCite and Reference Manager (see http://www.risinc.com/) has arrived with options for database publishing on the Web. This will be of interest to librarians who wants to create internal databases with Web access.

Professional Solutions

For large databases or projects with sufficient financial and technical resources it is possible, of course, to use more advanced database tools to develop more sophisticated databases. There are several possible ways to create these kinds of databases and to connect them to the World Wide Web (Lang & Chow, 1996; Ehmayer, Kappel & Reich, 1997). The typical way to connect SQL databases, which are among the most common of professional databases, is by using a server extension program on the Web server. The program works as an interpreter between the Web browser and the database server. Extension programs can be homemade CGI scripts (Cox, 1998) or commercial products (Beiser, 1997; Perez, 1998) like Cold Fusion or HotSQL. The advantage of professional tools include greater flexibility and better performance for large quantities of data. The major disadvantage of professional solutions for the computer amateur, which most librarians are, is greater dependence on computer consultants. A more practical solution is simply to rely on external databases when your database goals become too complex. Other possibilities also exist. Internal databases can be hosted by commercial OPAC software (Morgan, 1998; Notess, 1993) or by utilizing programs for building Web indexes like ROADS (see http://www.ilrt.bris.ac.uk/roads/).

Two Cases from Our Library

The University library of Tromsø consists of two main libraries. One for science, medicine, and health, and the other for humanities, social studies, and law. In addition, we have some smaller scattered libraries serving a few of the off-campus institutes. The library for science, medicine, and health is the result of a centralization in 1991, when several institute libraries merged and moved into a new building. It is a typical research library with about 80% of the budget spent on journals.

Centralized libraries have both advantages and disadvantages. Before the centralization, the users from some of the institutes had to walk just a few meters indoors to visit the library and they did so nearly every day, if only to read the papers. They knew the librarians and the collections well, and they were aware of everything new and relevant that arrived. After the centralization this changed. They now had to walk several hundred meters outdoors to a much larger library with thirty librarians and large collections that are difficult to survey. Our

location, Tromsø, is situated in the same latitude as Point Barrow, Alaska, and outdoor life can be a real endurance test, especially during the winter. We found that many users did not come to visit at all, and the rest visited the library infrequently and only when they had no alternative. Because of this, they missed many of the new books and journal issues that migrated into the large collections and thus these were rarely used. In both science and medicine, the latest books and articles are the most relevant because of the fast progress in research. At that time we didn't offer any networked services, and everything was located in the library.

Something had to be done to change this situation. Moving or changing our new library was, of course, considered next to impossible. We had to give our users compelling reasons for visiting the library. We were not lacking in new and relevant books and journals—the problem was that the users did not know about them since they did not visit the library. Our solution was to reach out to them and give them some irresistible reasons to visit us. We decided to create a kind of a virtual library that allowed them to easily check the latest week's new books and journal issues from their own offices. We hoped that if they discovered something interesting, and knew that they could go to the library and get it, they would do it.

The Journal Database

We started work in 1993 by improving user awareness of our journals. We started with journals because we mainly were a journal library and because the general situation for journals was unsatisfactory. At a minimum, our users needed answers to the following three questions:

- Which relevant journals does the library subscribe to?
- Does the library have this specific journal issue?
- Which relevant new journal issues have arrived lately?

The only information we could offer was the physical collection with 4,500 journals arranged alphabetically by title and an incomplete alphabetical list in the library with only the most basic bibliographic information. The situation improved in 1995 when journals were included in BIBSYS, but even today the functionality is far from adequate.

In June 1993, we created a FileMaker database through a combination of downloading and keying in records. This database could now print out complete and updated alphabetical and subject lists that we circulated among the scientists. Later, in June 1994, when we got our first Web server, the database could immediately export the same lists as Web documents and make them available on the Internet. In October 1994, we changed the list into a Web database using a Perl script and a text-file. The database (Brattli, 1998) has since been improved further with a Yahoo!-like interface (http://www.ub.uit.no/cgi-bin/tid4.pl/T/Generell?lenke=lokal).

The core of the journal database includes four FileMaker databases, one for each of the participating academic librarians. Together they cover all the libraries' journals and all updates are done here. Each week the content from these databases is copied and merged into a complete journal database. From this database the content is copied to the Web server as a text file or sometimes printed as a paper catalog available in the library. On the Web server, a Perl script generates Web documents from the text file dependent on requests from the user.

The four FileMaker databases are not similar but are tailor-made by each participating librarian. However, we have defined 23 common fields and formats because of the Web database and paper catalog. For these, each librarian has to supply data which includes: local ID, title, holdings, location, subject word, classification, and subscription status. Other common fields include title change, ISSN, record ID in OPAC, content, latest issue, and external URLs. Fields only in FileMaker databases include price, budget, and comment fields for binding, cancellation, and discarding.

The Web database is built around a Yahoo!-like subject tree. For each subject an alphabetical list of journals with holdings is presented. The users can choose between titles linked to OPAC or local bibliographic information, including non-subscribed titles, and including titles from subordinated subjects. For some subjects it is also possible to choose a chronological list indicating the latest issues to arrive. Local bibliographic records give more information about each journal. They are in addition linked to the OPAC for article ordering and to the publisher for general information and table of contents.

With so many fields one would imagine that updating each database is labor intensive, but fortunately most information about journals seldom changes. The updating of new issues actually takes little time, maybe fifteen minutes a week.

Recently Purchased Books (New Books)

The situation in 1993 was different for books than for journals. We had our excellent BIBSYS OPAC covering nearly all our books available in the library. The situation improved even more, especially for the users, from the end of 1993 when BIBSYS became openly available on the Web. Unfortunately for us and for the users, it was impossible to search for recently purchased books because the date fields are not indexed. Of course we could search for year of publication, but that is not the same. For us it was important to be able to tell the users that certain books arrived this week, others arrived last week, etc.

Our new book database at http://www.ub.uit.no/cgibin/bok4.pl/B17/Informatikk?liste=utvidet&periode=92 (Brattli, 1997) is built on the same principles and has gone through the same stages as the journal database. So far, it only covers half of the library's subjects. Because of the large number of new books, we soon realized that it was best to keep the local bibliographic information to a minimum and instead link each title to its record in BIBSYS. For each subject the database presents a chronological sorted list of titles of acquisitions over the last 12 months. The list is divided into three parts: books ready for lending, received but not ready, and books in order. The reason for including books long before they can be borrowed, is that the users can make reservations at an early stage. This triggers rush processing, and the requester is notified when the book is ready. The linking to BIBSYS supplies bibliographic information, location, loan status, and possibilities for making reservations.

Updating is done up to three times for each book. When the book is ordered, title and record-ID (for linking) are manually copied from BIBSYS to the FileMaker database. In addition, subject is chosen from a menu (thesaurus) and the order date is generated. When the book arrives in the library, we just have to click on the record to generate the "date arrived" and the corresponding is done when the book is ready for loan, to generate the "ready for loan date." With this database we have also kept statistics on processing time, use of books, and the importance of the new book display with respect to use.

Benefits and Use

An interesting observation about these databases today is that the reasons for keeping them have changed. When they first were created the main purpose was to inform the users about our books and journals. This is still an important use, but they have also become indispensable tools for us as academic librarians managing our collections. In fact, most of us feel that this is reason enough for keeping them. BIBSYS does not supply all of the data we need as the database is presently constructed.

We do not know whether the databases have led to an increase in the use of books and journals, but our impression is that they have. The importance is difficult to estimate, especially since so much has happened both with our OPAC as well as with the large external databases. We can envision our next research project to determine whether the internal databases lead to increases

in use. At present we can only say that the use of the databases is satisfying to us. We need to publicize these services more to our users. Many users do not enter the Web at all, and thus we need to advertise the Web as well. Both for us and for our users change takes time! So, did the databases solve our problems? The answer is, yes, some of them.

References

Baker, S. L. & Lancaster, F. (1991). *The measurement and evaluation of library services*. 2nd ed. Arlington, VA: Information Resources Press.

Beiser, K. (1997). Database-driven Web sites: Cold fusion for Web publishing. *Database 20*(6), 48-50, 52.

Berners-Lee, T., Cailliau, R., Luotonen, A., Nielsen, H., & Secret, A. (1994). The World Wide Web. *Communications of the ACM 37*(8), 76-82.

Biggs, D. R. (Ed.) (1995). *ProCite in libraries: Applications in bibliographic database management*. Medford, NJ: Learned Information, Inc.

Brattli, T. (1997). Bruk av lokale databaser til informasjonsformidling og bibliotekforskning. *Norsk tidsskrift for bibliotekforskning 10*, 51-74.

Brattli, T. (1998). Tidsskriftdatabasen ved Universitetsbiblioteket i Tromsø. *Ravnetrykk 15*.

Brudvig, G. L. (1991). Tailoring a journal article database to local needs: Planning and management issues. *Journal of Library Administration 15*(3/4), 85-100.

Cox, T. (1998). Using Perl with databases. *BYTE 23*(5), 57-58.

Delfino, E. (1996). Automatic HTML-part 2: Creating HTML from a database program. *Online 20*(6), 96-98.

Ehmayer, G., Kappel, G. & Reich, S. (1997). Connecting databases to the Web: A taxonomy of gateways. *Lecture Notes in Computer Science 1308*, 1-15.

Gates, R. (1989, October 17). Downloading info for local processing. PACS-L item 690. (PACS-L@LISTSERV.UH.EDU).

Hegna, K. (1994). Bidrag til idéen on det elektroniske bibliotek. *Bok og bibliotek 61*(8), 25-30.

Knudson, F. L., Sprague, N. R., Chafe, D. A., Martinez, M. L. B., Brackbill, I. M., Musgrave, V. A., & Pratt, K. A. (1997). Creating electronic journal Web pages from OPAC records. *Issues in Science & Technology Librarianship 15*.

Lang, C. & Chow, J. (1996) *Database publishing on the Web & intranets*. Scottsdale, AZ: Coriolis Group Books.

Langer, M. (1998). *Database Publishing with FileMaker Pro on the Web*. Berkeley, CA: Peachpit Press.

Morgan, E. (1998). We love databases! *Computers in Libraries 18*(2), 38-39.

Notess, G. R. (1993, June). Offspring of OPACs: Local databases on the net. *Database 16*(3), 108-10

Perez, E. (1998). Supercharge your Web site: Cheap and easy Web-database apps! *Library Software Review 17*(1), 24-30.

Raeder, A. & Tung, S. (1989). Downloading and converting bibliographic records from mainframe to micro using dBase. *Microcomputers for Information Management 6*(1), 11-32.

Smith-Cohen, D. (1993). Developing an inhouse database from online sources. *Special Libraries 84*(1), 9-17.

Spore. S. (1991). Downloading from the OPAC: The Innovative Interfaces environment. *Library Hi-Tech 9*(2), 69-79.

Stokes, R. (1982). *The function of bibliography*. London: Gower.

Webb, T. D. & Zhang, B. (1997). Information dropshipping; *Library Hi-Tech 15*(1/2), 145-149.

Zollman, K. & Zollman, D. (1997). Creating a simple searchable database on the Web. *Computers in Physics 11*(3), 225-228.