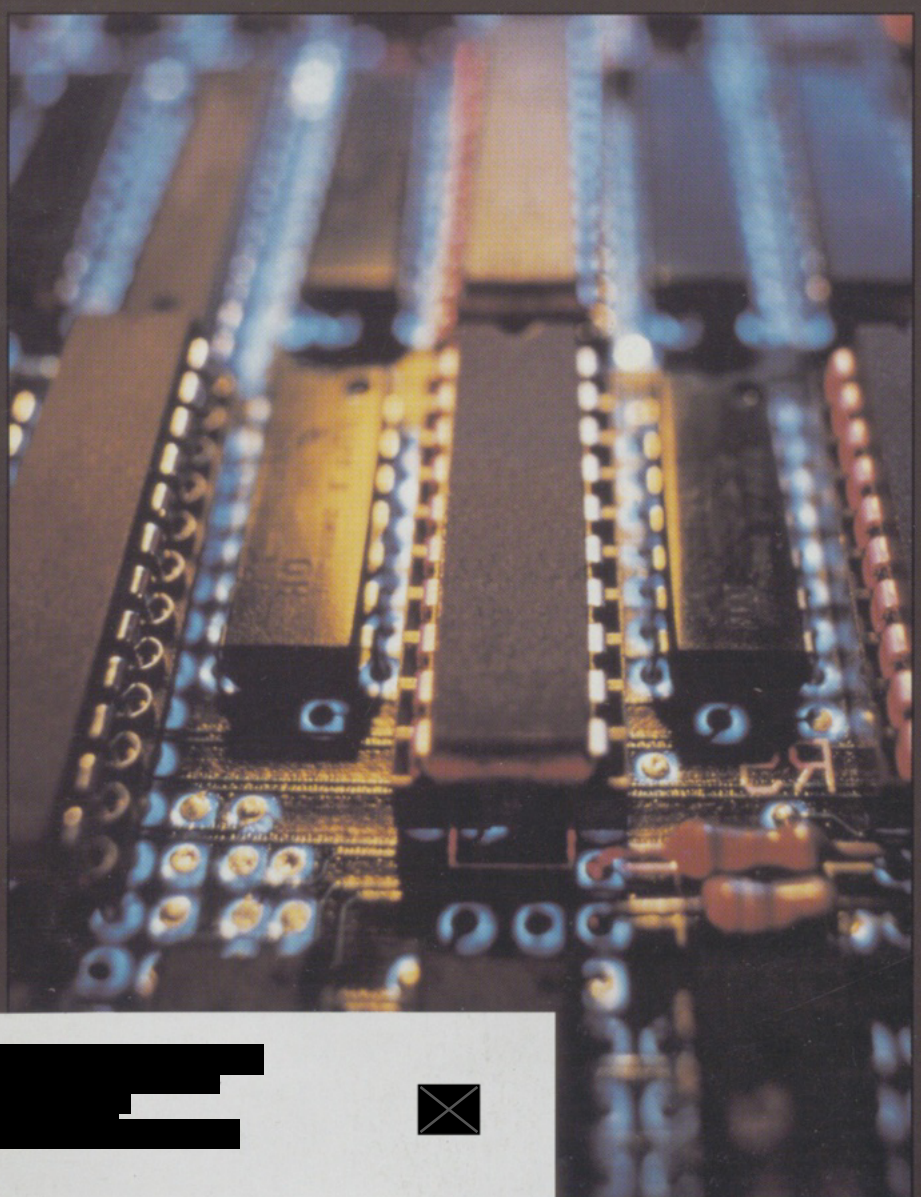
# Information Technology AND Libraries

ISSN 0730-9295

Volume 21 Number 2

JUNE 2002



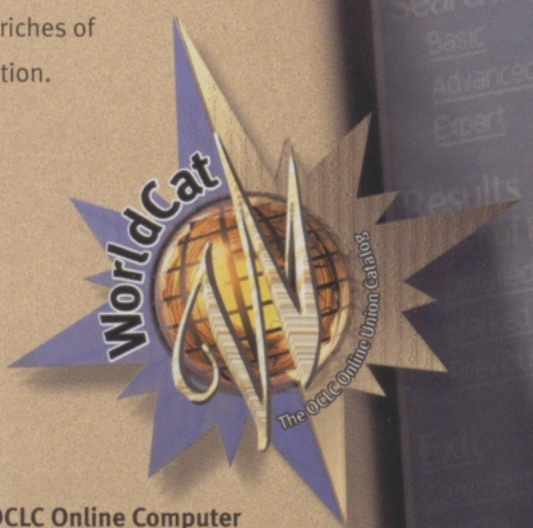


[redacted]

# Local control global reach

The OCLC FirstSearch service delivers a world of information to your users, with a focus on your collection. Through unparalleled resources like WorldCat that are enhanced by searching tools and system features you can customize, FirstSearch makes finding information convenient for your users and affordable for your budget. Your users have access to the holdings of thousands of libraries worldwide, full text from more than 9,000 serials (including 3,400 electronic journals), and best of all, the riches of your library's collection.

**www.oclc.org**





**OCLC Online Computer Library Center**

**OCLC—Furthering Access to the World's Information**

# LITA Regional Institutes

## New Regional Institutes now scheduling:

EBOOKS: AFTER HYPE AND DEATH, WHAT REALLY WORKS
Mark Beatty, presenter

A "state of the business" overview of eBooks. Emphasis will be library usability of eBooks. Covered topics include what librarians can actually do with eBooks now, what they can expect to happen, and not happen in the near future, practical applications and "buyer beware" warnings.

XML AND LIBRARIES
Ron Gilmour, presenter

This institute covers the structure and utility of Extensible Markup Language (XML) with an emphasis on the relevance of XML to libraries and the role XML plays in diverse applications such as web design, content management, and descriptive metadata.

If you are interested in hosting a Regional Institute, contact the LITA office at lita@ala.org.

For more information about hosting or attending a LITA Regional Institute, visit the LITA Regional Institutes web page

**www.lita.org/institut/index.html**

LITA Regional Institutes are one-day workshops on technology-related topics. Taught by experts and presented throughout the country, they reach beyond ALA Conference locations to bring you high quality continuing education.

*Information Technology and Libraries* (ISSN 0730-9295) is published quarterly in March, June, September, and December by the American Library Association, 50 E. Huron St., Chicago, IL 60611. It is the official publication of the Library and Information Technology Association, a division of the American Library Association. SUBSCRIPTIONS: Subscription price, $25, is included in membership dues. Nonmembers may subscribe for $50 per year in the U.S.; $55 in Canada, Mexico, Spain, and other PUAS countries; $60 in other foreign countries. Send subscription requests to *Information Technology and Libraries,* Subscription Department, American Library Association, 50 E. Huron St., Chicago, IL 60611; 800-545-2433; fax: (312) 944-2641; subscription@ala.org. Single copies, $15. Periodical-class postage paid at Chicago, Illinois, and at additional mailing offices. POSTMASTER: Send address changes to *Information Technology and Libraries,* 50 E. Huron St., Chicago, IL 60611.

*Information Technology and Libraries* publishes material related to all aspects of libraries and information technology, including digital libraries, metadata, authorization and authentication, electronic journals and electronic publishing, telecommunications, distributed systems and networks, computer security and intellectual property rights, technical standards, geographic information systems, desktop applications, online catalogs and bibliographic systems, optical information systems, software engineering, universal access to technology, futuristic forecasting, library consortia, vendor relations, and technology and the arts. *ITAL* welcomes unsolicited manuscripts. Submissions should follow the guidelines stated under "Instructions to Authors" on page 92 of this issue or on the LITA Web site at www.lita.org/ital/infoauth.htm. Manuscripts of feature articles, communications, and tutorials should be addressed to Dan Marmion, 221 Hesburgh Library, The University of Notre Dame, Notre Dame, IN 46556-5629. Books for review should be addressed to Tom Zillner, WILS, 464 Memorial Library, 728 State St., Madison, WI 53706. Software for review should be addressed to Andy Boze, Hesburgh Library, The University of Notre Dame, Notre Dame, IN 46556-5629.

ADVERTISING: The Goldman Group, 14497 N. Dale Mabry Hwy., Ste. 205N, Tampa, FL 33618; telephone (813) 264-2772; fax (813) 264-2343.

As a matter of policy, *ITAL,* as the scholarly organ of LITA, does not review LITA publications.

PRODUCTION: ALA Production Services (Troy D. Linker, Kevin Heubusch; Ellie Barta-Moran, Angela Hanshaw, Kristen McKulski, and Karen Sheets), American Library Association, 50 E. Huron St., Chicago, IL 60611.

Publication of material in *Information Technology and Libraries* does not constitute official endorsement by LITA or the ALA.

Abstracted in *Computer & Information Systems, Computing Reviews, Information Science Abstracts, Library & Information Science Abstracts, Referativnyi Zhurnal, Nauchnaya i Tekhnicheskaya Informatsiya, Otdyelnyi Vypusk,* and *Science Abstracts Publications.* Indexed in *CompuMath Citation Index, Computer Contents, Computer Literature Index, Current Contents/Health Services Administration, Current Contents/Social Behavioral Sciences, Current Index to Journals in Education, Education, Library Literature, Magazine Index, NewSearch,* and *Social Sciences Citation Index.* Microfilm copies available to subscribers from University Microfilms, Ann Arbor, Michigan.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1992.∞

# Information Technology AND Libraries

Volume 21, Number 2 June 2002 ISSN 0730-9295

Abstracts and contents pages of recent issues of *ITAL* can be found on the LITA Web site at www.lita.org/ital/index.htm.



**cover design** Kevin Heubusch • **interior design** Dianne M. Rooney

# Editorial: Listening to Our Users

Dan Marmion

What are your users telling you about your library? You are getting feedback from them, aren't you? If you are not, you should be. One of the traps we can fall into is thinking we are doing a good job of designing user interfaces when in fact we are not. One way to avoid that trap is to ask users to tell us how we are doing. There are several tools that can assist you in doing this, one of which is focus groups.

You can learn a lot by doing a series of focus group interviews. Get eight to ten library users together in a room for about an hour and a half, feed them or offer some other incentive, and ask five or six questions designed to get them talking about the issues of interest to you. Repeat several times. If you can invest a little money in a professional facilitator to moderate the sessions, plus record the discussions and have the tapes transcribed, you will increase the usefulness of the interviews. Keep in mind that the participants are not likely to be a statistically valid sample of your entire user population, so it's probably a good idea to employ other tools as well to gather input.

We recently did several focus group interviews in our library. We wanted to find out what problems our users encountered in accessing digital information, and we did indeed learn a lot. It turns out that we are not doing as good a job as we had hoped. A common theme among the focus group participants was that they often find it very difficult to find the information they need. Here are some of their comments:

> It might be lots of good stuff out there, but because it's so hard to find, you never bother to look for it.

> There's just so much there. I think every time you click on something, you're given twenty more choices and I don't feel like I know where to go.

> You've got to know how to do it in this library and you get the help page and you go through it and I'm not stupid, you know . . . but, I just can't figure it out.

The interface isn't what they would like it to be either:

> I want a single search to look at all the information as opposed to having to know that I need to go to a certain index and search for this type of information. I don't have to know whether it's in biological abstracts or whether it's an online journal that we hold, it would just go find it for me in whatever place it happened to be.

> It would be nice to be able to . . . have an interface that didn't require training and that was intuitive.

Over and over again, people kept asking for a simpler, less complex interface.

> My wish would be that the interface of the library's resources would look like Google, and it would operate that simply.

> It loads easily—you don't have to think—I can usually find what I need on the first page.

Yes, the G word surfaced many times—121 times, in fact, in fourteen hours of taped discussion. Our users want an interface that is as easy to use as Google. That's very clear. Faculty and graduate students mentioned Google most often as the first source of reference they consult. I think it is time that librarians face up to the fact that Internet search engines are having a profound effect upon those whom we regard as our primary clientele. They don't just wish that we made our systems as easy to use as Google. They are voting with their virtual feet and going to Google rather than to the library. We need to understand why and react accordingly.

Another thing to emerge from the interviews was that once they have managed to identify something, they would like it to be easier to obtain:

> It really frustrates me a lot if I go to search for an article and it's only an abstract and I have to go find the actual articles.

> I'm always crunched for time, as everybody is, and I'm designing my class and using the Web extensively to figure out what papers to put in the reading packet for these graduate students and if I had trouble finding it in the library and if I couldn't go get it immediately, it didn't end up on the reading list.

This puts me in mind of something that a colleague once pointed out to me. Savvy Internet users are accustomed to ending their transactions with having accomplished a goal. They go to an online merchant and place an order that results in what they purchased arriving on their doorstep. They can go to the Internet Movie Database and learn that Hugh Sothern played Andrew Jackson in the original 1938 version of *The Buccaneer* and again in 1939's *Old Hickory*. So many Internet-based transactions culminate with the user having possession of that which was sought. Not so in the library online catalog. Most OPAC-based transactions culminate with the user then having to go to the library and find that which was sought. We are seeing evidence that this is something users will actively avoid.

I've shared with you just a few of the comments captured in our focus group sessions, and I should hasten to add that there were many positive statements to go with the negative ones. Clearly, however, we have some work to do to make life easier for our users. We are still in the process of analyzing the data that we collected, but as you can see from the preceding, some trends have emerged. What we have learned so far adds emphasis to something that we already knew, which is that it is essential to do usability testing. That's where we are headed next.

**Dan Marmion** (dmarmion@nd.edu) is Associate Director for Information Systems and Access at the University of Notre Dame Libraries, Notre Dame, Indiana.

# Pinyin Romanization for OPAC Retrieval—Is Everyone Being Served?

Clément Arsenault

*As conversion of Wade-Giles entries to pinyin is underway, this paper reconsiders the usefulness of providing romanization-based retrieval in OPACs for Chinese-language resources. An in-house experiment designed to measure retrieval performance in OPAC title searches under different romanization methods revealed that while romanization is an efficient retrieval method that works relatively well for a large number of patrons, it remains problematic for a significant portion of end users who might be better served with character-based retrieval systems.*

Two romanization systems for Chinese data are currently in use in most libraries in the Western world: the Wade-Giles (WG) system, mainly used in North American libraries, and pinyin—called Hanyu pinyin but simply referred to as pinyin—mainly used in European and Australian libraries. In 1997 the Library of Congress (LC) announced its plan for converting to pinyin, which was unsurprisingly endorsed by the bibliographic community at large (except on the issue of word division) since "most librarians [. . .] have come to realize that conversion to pinyin will be necessary if North American libraries are to provide adequate service to their users."[1] The conversion process timeline, started in October 2000, extended over a one-year period, after which it was expected that the conversion of the records in individual libraries' OPACs would have been completed.[2]

On several points, inclusion of pinyin romanization is a remarkable added value to bibliographic records that facilitates retrieval for a great number of end users. However, relying solely on romanized entries for access may not be suitable for a portion of patrons who might be better served with character-based retrieval approaches. This paper intends to show that while romanization-based retrieval works well for most people, it remains problematic for a smaller but significant portion of catalog users.

## Background and Literature Review

### Comparing WG and Pinyin

The recent decision by the LC to convert from WG to the pinyin romanization system was long awaited by many library users in North America. Pinyin, promulgated in 1958, is now fully recognized as the official romanization scheme of the People's Republic of China (PRC); it was also recognized in 1977 as a United Nations Standard, and as an ISO standard (ISO 7098) in 1982.[3] Pinyin is now widely accepted in China and is used extensively by most government and press agencies around the world. It is used in the PRC to help first-graders and foreigners learn Chinese characters.[4] Pinyin is also used in publications such as dictionaries and maps, and sometimes for book and periodical titles. It is widely seen in public places such as building names; street, highway, and railway signboards; and on product labels.[5] WG still enjoys some popularity in Taiwan but on July 26, 1999, the Taiwanese government announced the use of pinyin (Hanyu pinyin from the mainland) for the romanization of street names, suggesting that Taiwan might soon officially adopt pinyin as its romanization scheme.[6]

When pinyin was developed, great care was taken to keep the notation as simple and as internationally acceptable as possible. For this reason, it was decided, "not to augment the Latin alphabet by adding new letters, such as [those used in] the International Phonetic Alphabet."[7] As a result, pinyin uses all the consonants of the roman alphabet—except *v*, which is only used for the transcription of foreign terms—in conjunction with four digraphs, namely *ch*, *sh*, *zh*, and *ng*.[8] Except for these four (or five) digraphs, all other consonant sounds are represented with a single roman letter. This, in some cases, proves to be bothersome for the native speaker of English. For instance, the phonetic values assigned to the pinyin letters *c*, *q*, *x*, and *z* do not correspond very closely to the values that are usually attributed to these letters in English. WG uses more digraphs (thirteen in total), producing a less compact notation, but one from which the sounds are usually, for a native English speaker, easier to infer.[9] For instance, the sound [X] (as in *sh*oe), represented in pinyin by *x*, is transcribed in WG with the digraph *hs*, which is easier to figure out for a native English speaker.[10] Another major difference between pinyin and WG notation is that pinyin uses different letters to mark the difference between aspirated and unaspirated consonants. Instead, WG (as shown in figure 1) uses an aspiration mark (often confused with the ayn) to distinguish between these aspirated and unaspirated pairs.[11]

In pinyin, the vowels *a*, *e*, and *o* are used fairly consistently, but *i* and *u* are used to represent various sounds, depending on the letter or digraph that precedes them. Hence, in pinyin, the vowel grapheme *u* is pronounced [u] after *b*, *c*, *ch*, *d*, *f*, *g*, *h*, *k*, *l*, *m*, *n*, *p*, *r*, *s*, *sh*, *t*, *w*, *z*, and *zh*, but [y] after *j*, *q*, *x*, and *y*. WG, on the other hand, distinguishes between [u] and [y] with the use of a dieresis: *u* and *ü* for each sound respectively. Pinyin makes use of

**Clément Arsenault** (clement.arsenault@umontreal.ca) is Assistant Professor at the School of Library and Information Sciences, University of Montreal, Canada.

| WG | Pinyin |
|---|---|
| k'/k | k/g |
| p'/p | p/b |
| t'/t | t/d |
| ch'/ch | q/j or ch/zh |

**Figure 1.** Graphemic Distinction of Aspirated and Unaspirated Consonants Pairs

the *u*-dieresis, but much more sparingly than WG. It is used in only four pinyin syllables, *lü*, *nü*, *lüe*, and *nüe*.[12] On the whole, the pinyin notation is somewhat simpler than the WG notation, as its use of diacritics, punctuation marks, and digraphs is noticeably lighter. On the other hand, for native speakers of English—and even for speakers of other western languages—WG produces romanized strings that are often easier to read and pronounce.

### Consequences on Retrieval

Previous analysis revealed that, in indexing, due to systematic removal of diacritics and punctuation, two hundred of the 410 WG base syllables (or toneless syllables) are collapsed into one hundred syllables. For example, *pa* and *p'a* are both searched with the string "pa." This represents a loss of one hundred syllables. Furthermore, twelve base syllables are collapsed into three syllables—*chu*, *ch'u*, *chü*, and *ch'ü* are all indexed as "chu"; the same for *chun* and *chuan*—which represents an additional loss of nine base syllables, for a total of 109 syllables lost. The remaining 198 syllables remain unaffected.[13] As for pinyin, only four of the syllables are collapsed into two syllables, a loss of two syllables, caused by the use of the dieresis on the *u* (*lu/lü* and *nu/nü*) as explained above. This leaves 301 "usable" WG syllables against 408 "usable"pinyin syllables. Table 1 gives a summary of the problem of loss of distinctly unique syllables due to transcription and indexing procedures.

## Romanizing Chinese Script

Transliteration is the operation that consists of representing the characters of one alphabet with those of another alphabet. By extension, romanization is the conversion of a nonroman script into roman script. When romanizing nonphonological writing systems such as languages that do not use alphabets or syllabaries—Chinese for example—the operation is slightly more complex since no direct graphemic correspondence can be established between the target and the host script. Technically speaking, transliteration (*script* to *script* conversion) between the Chinese and roman scripts is impossible; only transcription (*sound* to *script* conversion) is feasible.[14] Romanized Chinese is thus, in a sense, twice-removed from the original script. The pronunciation of characters in modern standard Chinese language is defined in the Mandarin standard, based on the Beijing dialect. Mandarin is the standard phonological and grammatical expression of the Chinese language, and is now accepted worldwide. Phonetic transcription is, therefore, usually based on the standardized sounds of the characters in Mandarin.

There are seven major Chinese dialects, along with hundreds of minor regional dialects.[15] It is recognized that, although "there is practically one universal Chinese grammar . . . apart from minor divergences . . . and differences . . . in some southern dialects," there is a tremendous phonological disparity between dialects, along with some variations in the vocabulary, to the point where the local speech of southerners is almost completely unintelligible to northerners.[16]

The Chinese syllable structure is relatively simple—Mandarin syllables can take the form (C)(Y)V(C or V), where each bracketed element is optional.[17] Because of this, Mandarin has

> an extremely high phonological load for all the variables involved; in other words, any little difference will make a great difference, and mispronouncing a word will very likely result in saying another word.[18]

The adverse consequence of this is readily observable when non-native Mandarin speakers romanize Chinese according to what they believe to be accurate pronunciation. For example, speakers from the Wu dialect area (Suzhou/Shanghai) tend to confuse dental and guttural nasal endings.[19] This can be explained by the fact that Wu has its own dental/guttural pattern (in Chinese *qianbiyin/houbiyin*), that is not necessarily the same as the one in Mandarin, hence the confusion. For instance, the character meaning "forest" is pronounced *lin* (dental) in Mandarin, but *ling* (guttural) in the Shanghai dialect.[20] King—referring to a personal communication with Professor Xu Baohua of Fudan University, Shanghai—also remarks that even people living in Mandarin-speaking regions, such as Sichuan and Manchuria, tend to make numerous errors in phonetic transcription since their native speech

> is a sub-dialect of Mandarin, with little difference from Putonghua [Mandarin], [and] when [they] learn

**Table 1.** Summary of Syllable Loss Due to Romanization and Indexing Procedures

| | WG | | Pinyin | |
|---|---|---|---|---|
| | No. | % | No. | % |
| Total number of syllables | ca. 1,300 | 100 | ca. 1,300 | 100 |
| Base syllables (toneless) | 410 | 31.5 | 410 | 31.5 |
| Indexing tokens created | 301 | 23.2 | 408 | 31.4 |

> Putonghua, they experience more interference from their own dialect than speakers of other languages do . . . [and they tend to] settle for non-standard speech [without being aware of it].[21]

As one can imagine, this can cause serious problems in information retrieval in terms of query formulation in romanized form.

## The Need for Transliterated Data in Bibliographic Records

The International Standard Bibliographic Description (ISBD) prescribes the use of the vernacular script in the production of bibliographic records, regardless of the script chosen by the cataloging agency.[22] It should be noted that this preference for using the vernacular script in cataloging records is also expressed in the revised version of the second edition of the Anglo-American Cataloguing Rules (AACR2r 1998, rule 1.0E), but under USMARC,

> true nonroman text, the source of the romanization, [was] regarded as the alternate graphic representation of its Latin [i.e. roman] script rendering, and is an optional addition to bibliographic records.[23]

One can understand the initial motives for ensuring that entries are at least indexed in roman script, because in an electronic environment, it is almost certain that roman is available on all platforms worldwide. It can only be hoped that the approval of the Universal Character Set (UCS)/Unicode in MARC 21 will promote and facilitate the inclusion, display, and exchange of Chinese vernacular script in electronic bibliographic records.[24] The recently released MARC 21 format also allows the inclusion of several scripts concurrently within one bibliographic record, either under the MARC-8 environment—where different sets of characters accessed via escape sequence are used concurrently within one record or field—or under the global UCS/Unicode environment.[25]

Nonetheless, in North America, roman is still regarded as the *de facto* required script, as the majority of computerized systems are not yet equipped with the necessary hardware architecture and appropriate software to handle nonroman script. Making romanized fields a required element in the records of nonroman items is, therefore, a safeguard to ensure the usability of all records on every kind of platform. In the case of Chinese, transliterated text may actually prove particularly useful for data entry and problems referring to filing. In addition, Anderson has shown that

> nearly 25 percent of all sources of bibliographic information about Chinese language materials used by all catalogue users [in North America] do not supply that information in Chinese characters, but only in romanized form.[26]

In these cases, unless romanized entries are included in bibliographic records, users will most likely be unable to trace back their references since it is often quite difficult to reconstruct with exactitude the original script from a romanized citation, especially from one rendered in monosyllables.

In the case of Chinese, romanization is an indispensable addition to electronic bibliographic records, one that provides an excellent means for retrieval, as it removes the burden of computing with Chinese characters. However, because of dialectal interference and the high phonological load of the Chinese language, a number of end users may experience problems in transcribing pinyin correctly, and may be better served by alternative retrieval methods.

## Method and Analysis

### Experimental Setting

During a controlled experiment primarily designed to measure variations in retrieval performance in OPAC title searches based on type of romanization, thirty library users were asked to use romanized data to search forty monograph titles given to them in Chinese characters.[27] Participants were divided into three treatment groups: WG, monosyllabic pinyin (mPY), and polysyllabic pinyin (pPY). The OPAC contained almost fifty thousand bibliographic records, each containing three separate title fields, one in each of the romanized forms mentioned. Half of the titles were searched by phrase (exact-title mode) and half were searched by keyword (keyword mode). Transaction logs, capturing the query strings of the participants, were generated by a concealed logging program. All participants were provided with identical lists, except for the sequential order of the titles, which was rotated to minimize the learning factor.

### Data Analysis

The success rate, defined as the proportion of records retrieved from the total number of records searched, ranged between 71 percent and 91 percent depending on romanization groups and search mode. Unsuccessful queries were extracted from the logs and stored in a separate file. Each query string was analyzed to determine the cause of failure. As reported in an earlier article by the author, unsuccessful queries were first classified into two categories: problems related to aggregation of syllables, and problems related to romanization errors.[28] The

findings reveal that romanization errors account for approximately one-third to one-half of all errors, depending on the romanization group (see table 2).

Romanization problems were further categorized into four main types:

- $R_1$: Character mispronounced or misread (*ji* read as *jie*)
- $R_2$: Pinyin was used instead of WG, or vice-versa (*song* instead of *sung*)
- $R_3$: Phonetic transcription error (*chen* instead of *zhen*, *cheng* instead of *chen*, *fu* instead of *hu*)
- $R_4$: Miscellaneous errors (*wuan* instead of *wan*)

As can be seen from the data in table 3, most of the romanization errors are transcription errors caused by sound confusion of phonetic nature ($R_3$), usually confusions between quasi-homophone syllable pairs, namely between dental nasals and guttural nasals (pinyin and WG: n/ng), between aspirated and unaspirated dental sibilants, retroflexes and palatals (pinyin: z/c, zh/ch and j/q; WG: ts/ts', ch/ch', ch/ch' respectively), and between dental sibilant fricatives and retroflex fricatives (pinyin: s/sh; WG: ss/sh). Figure 2 gives a breakdown of these errors by type. The proportions are highly indicative of the dialectal interference mentioned earlier.

## Summary and Conclusions

### WG versus Pinyin

It is interesting to note that use of wrong romanization ($R_2$ errors) only occurred in the WG group. During WG searches, participants would occasionally revert to using pinyin. Since study subjects were, by their own admission, more familiar with pinyin than with WG, one could speculate that, when this occurred, they were inadvertently slipping back into their dominant mode (pinyin). Not surprisingly, there was no manifestation of the opposite as the reverse phenomenon was not observed. Level of familiarity is also suspected to be the main factor for the lower success rate observed in the WG group as shown in table 4.[29]

Notice however that the number of phonetic transcription errors ($R_3$) is much smaller for WG searches. Although this may seem surprising at first glance, it can be explained by the fact that WG is, in a way, more "forgiving" than pinyin when using either aspirated/unaspirated initial sibilants or the retroflex or palatal consonants in search queries, since the aspiration mark is not indexed. Therefore, even if the end user confuses, for example, the sounds *chen* and *ch'en*, this has no consequence since both *chen* and *ch'en* are indexed as *chen*. This, however, impacts negatively on the precision rate. This phenomenon is not manifested in pinyin searches since the distinction between aspirated and unaspirated consonants is expressed by using a distinct roman letter or group of letters (z/c, zh/ch, j/q), and in the above exam-

**Table 2.** Number of Errors by Type (Average per Unsuccessful Query)

| | WG group | | mPY group | | pPY group | |
|---|---|---|---|---|---|---|
| | No. | % | No. | % | No. | % |
| Aggregation errors | 0.66 | 66 | 0.67 | 47 | 0.85 | 61 |
| Romanization errors | 0.35 | 34 | 0.76 | 53 | 0.55 | 39 |

**Table 3.** Average Number of Romanization Errors per Unsuccessful Query

| | WG group | | mPY group | | pPY group | |
|---|---|---|---|---|---|---|
| | No. | % | No. | % | No. | % |
| $R_1$ | 0.11 | 31.0 | 0.14 | 19.0 | 0.11 | 18.8 |
| $R_2$ | 0.05 | 15.6 | - | - | - | - |
| $R_3$ | 0.14 | 40.0 | 0.50 | 65.8 | 0.42 | 76.5 |
| $R_4$ | 0.05 | 13.4 | 0.11 | 14.3 | 0.03 | 4.7 |
| Total | 0.35 | 100 | 0.76 | 100 | 0.55 | 100 |

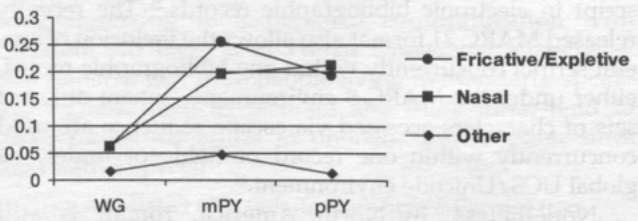



**Figure 2.** Average Number of Errors Related to Phonetic Transcription per Unsuccessful Query

**Table 4.** Success Rate (Mean)

| | WG group | | mPY group | | pPY group | |
|---|---|---|---|---|---|---|
| | % | n | % | n | % | n |
| Exact-title mode | 71.5 | 6 | 84.9 | 12 | 90.6 | 12 |
| Keyword mode | 70.6 | 5 | 80.0 | 12 | 80.9 | 11 |

ple the syllables are written *zhen* and *chen* respectively, which produces two different index tokens.

## Romanization As a Means for Retrieval

The high success rates measured for the retrieval task (table 4) indicate that romanization is, in most cases, a fairly effective means for retrieval of Chinese language titles. On the other hand, the number of errors directly attributable to romanization is fairly high. Figure 3 reveals that the distribution of the number of errors per participant follows a bimodal pattern, which indicates that although romanization was not too problematic for a large number of participants, it appeared to have been very problematic for some.

Some participants attested to this difficulty during post-search interviews:

> **Participant #29:** Mastering of pinyin is okay if you are from Beijing but for most people from the South the *z/zh, s/sh,* and also *n/ng* sounds are quite confusing. . . . [so] I had to modify the queries quite often.
>
> **Participant #12:** Spelling of pinyin is sometimes confusing.
>
> **Participant #30:** The task was not difficult but pronunciation is sometimes a problem.
>
> **Participant #22:** It was sometimes difficult for me to differentiate between sounds such as *s/sh, ch/zh,* so the keyword [mode] is better because you can select words that do not have these sounds.

## Further Research

As can be seen, a large number of search failures are directly attributed to romanization errors in the query strings input by the participants. The main conclusion to be drawn from this is that using romanization as the primary retrieval technique in OPACs still remains rather problematic for many end users. It would therefore be desirable to investigate the possibility of providing other means of retrieval, most probably character-based, as an alternative for retrieval of Chinese-language records in OPACs. This is certainly feasible, as most OPACs are now accessible and searchable through Web browsers. The adoption of the global UCS/Unicode character set in MARC 21 should facilitate and improve multiscript searching in Web-based bibliographic databases.

## References and Notes


1. Library of Congress, Library of Congress Will Convert to Pinyin for Romanization of Chinese, 1997. Accessed Aug. 29, 2001, http://lcweb.loc.gov/catdir/pinyin3.html.
2. Peter Zhou, "Summary Report on Pinyin Conversion Planning Meeting in Washington D.C.," *Chinese Librarianship: An International Electronic Journal*, no. 8 (1999). Accessed Aug. 29, 2001, www.whiteclouds.com/iclc/cliej/cl8.htm.

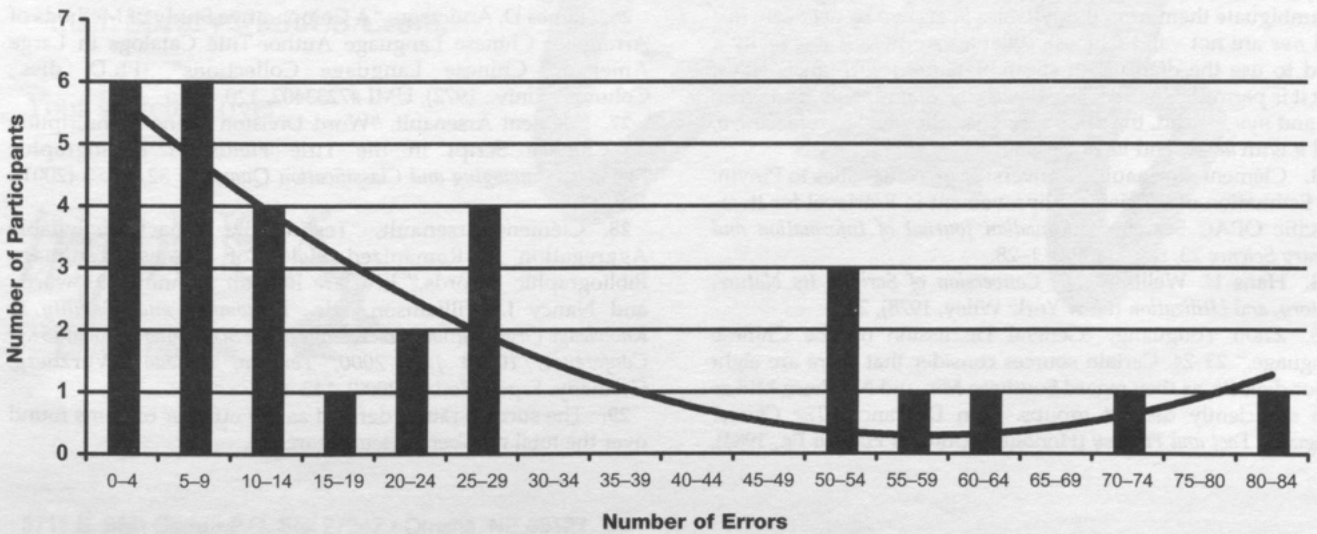




**Figure 3.** Total Number of Romanization Errors of Each Participant

3. Zhou Youguang, *Zhongguo yuwen zongheng tan* ("General Discussion on the Chinese Language" in Chinese) (Beijing: Renmin jiaoyu chubanshe, 1992).

4. Lee Shin-Ying, James W. Stigler, and Harold W. Stevenson, "Beginning Reading in Chinese and English," in *Acquisition of Reading Skills: Cultural Constraints and Cognitive Universals*, ed. B. R. Foorman and A. W. Siegel, (Hillsdale, N.J.: L. Erlbaum Associates, 1986), 123–50; Shu Hua and Richard C. Anderson, "Role of Radical Awareness in the Character and Word Acquisition of Chinese Children," *Reading Research Quarterly* 32, no. 1 (1997): 78–89.

5. John DeFrancis, "The Why of Pinyin Grapheme Selection," *Journal of the Chinese Language Teachers Association* 25, no. 3 (1990): 1–14.

6. Hwang Hsuan-fan et al., "Romanization Must Strike a Balance," *Taipei Times*, 9 Jan. 2000. Accessed Aug. 29, 2001, www.taipeitimes.com/chnews/2000/01/09/story/0000018872 [in Chinese; Big5 encoding].

7. DeFrancis, "The Why of Pinyin Grapheme Selection," 2.

8. Zhou Youguang, "General Discussion on the Chinese Language," 285. [Note that] *er* may be considered as a digraph, bringing the total to five (221).

9. Ibid, 212.

10. International Phonetic Association, *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet* (Cambridge: Cambridge Univ. Pr., 1999). Here and throughout this paper, pronunciations given in square brackets are given in the International Phonetic Alphabet notation.

11. The actual ayn is used, for example, in the transcription of Arabic to represent the pharyngeal voiced fricative. The character used in WG for the aspiration mark is the "reversed comma," character Unicode U+02BD, and has the appearance of an apostrophe, flipped left-to-right (from a personal e-mail communication with Ken Whistler, Technical Director, Unicode).

12. DeFrancis, "The Why of Pinyin Grapheme Selection," 4. Only in the case of *lü* and *nü* is the dieresis really necessary to disambiguate them from the syllables *lu* and *nu* respectively (*lue* and *nue* are not valid Chinese syllables so there is not really a need to use the dieresis for disambiguation). DeFrancis states that it is permissible to replace *ü* with the diagraph *yu*, and write *lyu* and *nyu* instead, the same way it is "allowed" to replace *ä, ö,* and *ü* with *ae, oe,* and *ue* in German.

13. Clément Arsenault, "Conversion of Wade-Giles to Pinyin: An Estimation of Efficiency Improvement in Retrieval for Item-Specific OPAC Searches," *Canadian Journal of Information and Library Science* 23, no. 3 (1998): 1–28.

14. Hans H. Wellisch, *The Conversion of Scripts: Its Nature, History, and Utilization* (New York: Wiley, 1978), 28.

15. Zhou Youguang, "General Discussion on the Chinese Language," 23–24. Certain sources consider that there are eight major dialects, as they regard Southern Min and Northern Min as two sufficiently distinct groups. John DeFrancis, *The Chinese Language: Fact and Fantasy* (Honolulu: Univ. of Hawaii Pr., 1984), 58; Bao Zhiming, *The Structure of Tone* (New York, Oxford: Oxford Univ. Pr., 1999), 8.

16. Chao Yuen Ren, *A Grammar of Spoken Chinese* (Berkeley: Univ. of California Pr., 1968), 13.

17. C stands for consonant, V for vowel, and Y for semi-vowel.

18. Chao, *A Grammar of Spoken Chinese*, 23.

19. Paul L. King, "Contextual Factors in Chinese Pinyin Writing" (Ph.D. diss., Cornell Univ., 1983), UMI #8321888, 70.

20. Chao Yuen Ren, *Mandarin Primer: An Intensive Course in Spoken Chinese* (Cambridge: Harvard Univ. Pr., 1961), 7.

21. King, "Contextual Factors in Chinese Pinyin Writing," 98–99. Quoting Xu Baohua in telephone conversation with author. See also Zhou Y., "General Discussion on the Chinese Language," 23, who estimates that over two-thirds of the population of mainland China speaks one of the four main forms of Mandarin.

22. ISBD Review Committee Working Group. *ISBD(G): General International Standard Bibliographic Description: Annotated Text*, rev. ed. (Munich: Saur, 1992), 0.6.

23. Joan M. Aliprand, "Linkage in USMARC Bibliographic Records," *Cataloging and Classification Quarterly* 16, no. 1 (1993): 5–37. Emphasis in original text. Library of Congress, MARC 21 Concise Bibliographic: Multiscript Records. Accessed Aug. 29, 2001, http://lcweb.loc.gov/marc/bibliographic/ecbdmulti.html. Note that in the newly adopted MARC 21 format any script can be considered as the primary script.

24. Library of Congress, MARC 21 Specifications for Record Structure, Character Sets, and Exchange Media. Character sets Pt. 2: USC/Unicode Environment. Accessed Aug. 29, 2001, http://lcweb.loc.gov/marc/specifications/speccharucs.html; Unicode Consortium, *The Unicode Standard*, Version 3.0. (Reading, Mass: Addison-Wesley, 2000).

25. Library of Congress, MARC 21 Specifications for Record Structure, Character Sets, and Exchange Media. Character Sets Pt. 1: MARC-8 Environment. Accessed Aug. 29, 2001, http://lcweb.loc.gov/marc/specifications/speccharmarc8.html.

26. James D. Anderson, "A Comparative Study of Methods of Arranging Chinese Language Author-Title Catalogs in Large American Chinese Language Collections" (Ph.D. diss., Columbia Univ., 1972), UMI #7233402, 120.

27. Clément Arsenault, "Word Division in the Transcription of Chinese Script in the Title Fields of Bibliographic Records,"*Cataloging and Classification Quarterly* 32, no. 3 (2001): 109–37.

28. Clément Arsenault, "Testing the Impact of Syllable Aggregation in Romanized Fields of Chinese Language Bibliographic Records," in Clare Beghtol, Lynne C. Howarth, and Nancy J. Williamson, eds., *Dynamism and Stability in Knowledge Organization: Proceedings of the Sixth International ISKO Conference, 10–13 July 2000, Toronto, Canada* (Wurzberg, Germany: Ergon Verlag, 2000), 143–49.

29. The success rate is defined as the number of items found over the total number of items searched.

# Find it. Faster.

You just got a request for a list of grocery stores in a three-state area with more than 20 employees.

You could search through dozens of reference sources to create that list. Or, you could call the Library Division of *info*USA.

We offer the country's most extensive databases of business and residential information. And, you choose the format you want: print, CD-ROM or via the Internet.

Want more information? Call us today at **1-800-808-1113** or e-mail: **library@infoUSA.com.**

Your search is over.

*info*USA®

*Library Division*

5711 S. 86th Circle • P.O. Box 27347 • Omaha, NE 68127
Phone: (402) 593-4523 • Fax: (402) 596-7688 • www.libraryUSA.com

23930

# Bringing Them In and Checking Them Out: Laptop Use in the Modern Academic Library

Jason Vaughan
and Brett Burnes

*Laptop computers provide unsurpassed flexibility and convenience for students in the modern academic environment. Many libraries, including the University of Nevada–Las Vegas's (UNLV) Lied Library, allow students to bring in their own laptops or check out library-owned laptops and connect to the institution's high-speed network. While offering unsurpassed convenience, laptop use must be tempered with appropriate authentication, security, and additional procedural policies to ensure that such privileges are not abused. In addition, libraries must be prepared to accommodate financial and staff-time costs associated with such programs.*

A hot new trend initiated by many academic libraries is allowing patrons to use laptops connected to the library's network. Modern libraries are often wired in many places, and older libraries have the option to add additional wiring or to investigate wireless solutions. This allows patrons to use a computer within quiet study rooms or while relaxing on lounge furniture. Since the opening of UNLV's Lied Library in January 2001, students have had the option to bring in their own laptops and take advantage of the library's fast network connection, using their own laptops with their own programs and personal setups with which they are familiar. Similarly, Lied Library began circulating library-owned laptops to students beginning in fall 2001.

From the start, the new 300,000-square-foot main Lied Library was designed and built to accommodate thousands of PCs. Data and power connections are integrated directly into both tables and carrels, and in the future, laptop connections will be activated within the floors near lounge chairs. Such flexibility allows users to bring in their laptops and plug in basically anywhere. Regardless of who owns the laptop—the student or the library—there are numerous issues and policy decisions that should be made prior to implementing such a program. For students bringing in their own units, a major issue is authentication to the library network. For those checking out laptops, issues include check-out policies, return polices, and damage assessment. This article looks at how the UNLV libraries are offering both of these services, as well as providing a summary analysis of how selected other libraries are offering similar services.

**Jason Vaughan** (jvaughan@ccmail.nevada.edu) is Head, Systems Section, and **Brett Burnes** (brett.burnes@ccmail.nevada.edu) is Network Specialist at the University Libraries, University of Nevada–Las Vegas.

## The LINAS System: Patrons Bringing in Their Own Laptops

A system allowing students to bring in their personal laptops and connect to the library's network was in place when Lied Library opened in January 2001. This was later expanded to also include the Architecture Studies branch library. Many issues affecting library-owned laptops are not relevant with personal laptops as there are no circulation parameters, damage assessments, or fine rates to determine. However, authentication and security are major concerns. By making available its network resources to patrons, the library inherits a certain amount of responsibility. As university policy and common sense dictate, all users must authenticate each time prior to gaining access to the library's network. A range of options exist and different institutions take various approaches with authentication. UNLV wanted a straightforward system, easy to maintain, that would not require the need to install special software on the patron's laptop.

When library staff controls the network hardware provided to the patron, appropriate system-enforced policies provide a level of security. For instance, the library can ameliorate the risk of a user executing malicious software against remote targets or abusing legitimate software (such as configuring a rogue file-transfer protocol (FTP) server with the intent to illegally distribute software without appropriate licensing). By allowing patrons to use their own laptops for network access, the library surrenders this ability to prevent the use of malicious software, or the abuse of legitimate software.

The range of possible malicious activities in which a patron could engage is limited only by the imagination. Two particularly salient concerns are the potential use of library resources to illegally distribute or collect electronic media or self-replicating software and to attempt to gain illegal access to systems. High bandwidth connectivity is crucial to one who wishes to make available, or to illegally obtain, copied software, music, videos, or books. Numerous ways exist to exchange this media over the Internet, including use of Internet Relay Chat (IRC), FTP, and Peer-to-Peer client software (such as KaZaA, Gnutella, or AudioGalaxy). Beyond the obvious legal implications, this activity can produce an unnecessary burden on the network infrastructure, usurping bandwidth that may otherwise be used for legitimate research activities.

A second concern, users attempting to crack and obtain access to other systems or networks, is becoming a serious threat. There was a time when a "cracker" (a programmer who engages in the act of attempting to gain access illegally to a computer system or network) had to be a very proficient programmer to accomplish their goals. Such crackers are often motivated by the challenge itself or by a desire to inflict damage. It was recently reported by

*Network World News* that the Kournikova virus which inflicted over $166,000 worth of damage to fifty-five victims was produced by an individual who was not a programmer.[1] The perpetrator used a worm-making tool kit to generate malicious software. Similar kits, known as root kits, are available that allow people with virtually no technical knowledge to obtain access to servers and inflict damage by removing or corrupting data, obtaining proprietary or confidential information, or using the breached server to launch attacks against another system.

Crackers can use breached servers to launch a Distributed Denial of Service (DDoS) attack against a particular server (usually a Web server). A DDoS attacker attempts to breach many servers, installing software that in essence makes many repeated requests against a particular Web server. If enough servers are running this software, the target server can be overwhelmed with traffic and stop responding to legitimate requests. Attempts to track the perpetrator of such an attack often dead-end at another breached server.

This very brief review of possible misuses of library resources illustrates that any attempt to anticipate and block all forms of malicious behavior is feeble. Instead, the staff has put their energies into ensuring that any malicious activity may be traced back to the perpetrator, who can then be held accountable. Should the library become aware of an attack which originates from one of the block of IP addresses reserved for use by the patron laptops, it must be able to associate the given address and the given time of attack with the patron to whom the address was leased. To accommodate these concerns, UNLV implemented the LIed Network Authentication System (LINAS).

## System Requirements

LINAS was designed at the outset to function with any laptop regardless of operating system or hardware manufacturer. The system has been used on Apple hardware running Mac OS 8, 9, and OS X and on Intel hardware running Microsoft Windows 95 through XP. In theory the system should work with any configuration that supports TCP/IP connectivity along with the Dynamic Host Configuration Protocol (DHCP). It was decided to prototype the system on Intel hardware running GNU/Linux 2.2.14 (Redhat 6.2). This system consists of ten HTML files and seven Common Gateway Interface (CGI) scripts that provide the user interface, three Perl modules that support the CGI scripts, three control daemons that automate various administrative functions, the Internet Software Consortium's DHCP daemon, and a MySQL™ 3.22.32 database. A Linux 2.2.17-21mdk kernel (Linux-Mandrake 7.2) running MySQL 3.23.23-beta was migrated to shortly after releasing the prototype into production. A summary of these components is provided in appendix A.

As the design and development of LINAS began, one of the utmost concerns was simplifying the registration process, so experienced and unexperienced patrons alike would enjoy a simple, efficient, one-time registration process. In particular, it was decided not to rely on the patron's knowledge of their network adapter's Media Access Control (MAC) address. To simplify the process, the patron's MAC address must be able to be detected when they access the Library's Web page-based registration form. This information is acquired using the Address Resolution Protocol (ARP) which is a link-layer protocol used with certain types of network interfaces including token-ring and Ethernet. ARP automatically provides a mapping from an IP address to its corresponding MAC address within a common broadcast domain.[2]

The authentication system is built around the Internet Software Consortium's (ISC) DHCP daemon; three control daemons; a collection of CGI scripts which provide the end-user registration interface to the library patron database; and the administrative interface used by library staff to approve, deny, and maintain patron records. The library utilizes Innovative Interfaces' Innopac integrated online library system in addition to an application programming interface supplied by Innovative. The ISC DHCP daemon configuration file allows one to class devices by MAC address. These classes may be allowed or denied from different pools of IP addresses. Two classes (approved and denied), and two IP address pools have been configured. The first pool contains one IP address only, and denies members of the approved and denied classes. This single IP address will only be granted to unknown or unregistered devices. The second pool, which contains the remaining portion of IP addresses, allows members of the approved class and denies members of denied class. The second pool only leases IP addresses to approved users. This mechanism provides the foundation upon which LINAS operates, allowing delineation between allowed and prohibited network connections.

## The Registration Process

As described above, a specific IP address has been allotted for use by unregistered laptops. To register a personal laptop, the patron first goes to a Web-page registration form. When first-time, unregistered patrons load the Web-page registration form on their browsers, the library server's ARP table should contain MAC address entries for the IP addresses allotted to unregistered patrons. After patrons complete the registration form, their names and library barcodes are stored in the MySQL database along with their MAC address as parsed from the server's ARP

table and assigned a pending status. At this point, library staff must approve the patron. The service-counter librarian, using the administrative CGI scripts, views the patron's account information as retrieved from the library's Innopac patron database (see figure).

After verifying the patron's information, the service counter staff approves the patron's record using the administrative CGI scripts. The CGI script updates the MySQL record from pending to approved, and generates a new copy of the DHCPd.conf file that includes that user's MAC address in the approved class. At this point, the second technical challenge arises. The ISC DHCP daemon does not automatically recognize changes made to its configuration file after its initial load.[3] It must be restarted when a newly approved or denied MAC address is added. Complicating this limitation is the fact that the DHCP daemon (DHCPd) should be run as the system's administrative user, in this case "root."[4] GNU/Linux security features prevent unprivileged users from arbitrarily restarting processes executed by privileged users, including the DHCPd process. To further automate the process, an intermediate program was desired to sense when a new laptop had been added in order to restart the DCHP daemon. In accomplishing this, to avoid unnecessary security risks, we implemented a control daemon—refreshd—that is capable of restarting DHCPd upon receiving notification of a change having been made to DHCPd.conf.

Refreshd is a daemonized process that runs as root. When a patron record is changed and a new copy of DHCPd.conf is generated, a trigger file is also created, indicating that DHCPd should be restarted. Refreshd awakens at a user-defined interval (in this case, five seconds) to test for the presence of the trigger file. If the file exists, DHCPd is restarted, and the trigger is deleted. At this point patrons may release and renew their leases using winipcfg, or restart their laptops to obtain a valid Internet routable address. A flowchart outlining the LINAS registration process is provided in appendix B.

In the design of the system, both the entire registration process and the day-to-day use of network access were made as intuitive as possible. During the initial implementation, unregistered laptops were automatically redirected to the registration form, asking the patrons to complete the form and approach a service counter to have their requests approved. It soon became obvious that the vast majority of patrons required a certain amount of assistance configuring their laptops correctly. In response, a registration station was created at a first floor service counter. Library staff that work at this counter have been trained in basic configuration and troubleshooting of laptops as well as the operation of the LINAS administration program. When the patron approaches the service counter, a staff member assists the patron in configuring the laptop, ensuring that the machine is connecting to the network and retrieving an appropriate IP address from

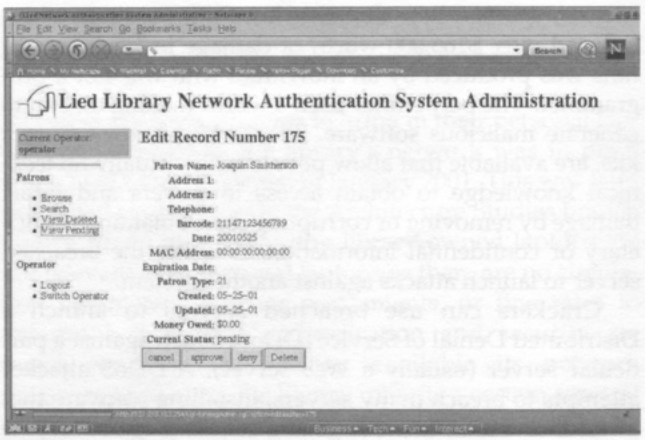



**Figure.** Patron Record in the Innopac Database

the DHCPd server. The staff member then provides the URL of the registration Web page to the patron.

When the form is completed, the patron must sign a laptop-use agreement, acknowledging receipt of a copy of the laptop-use policy. Both of these documents appear in appendix C. At this point, the staff member verifies the patron's picture ID, and approves or denies the request, using the LINAS administration program. While still at the registration station, patrons may verify functionality by rebooting their laptops or, for Windows-based laptops, using winipcfg to release and renew their IP Leases.

## Popularity

The system has been found to be useful beyond its initial scope. LINAS is being used to secure public access desktop PCs in two other departments that are housed in the library. As of March 2002, 374 individuals have registered their own laptops within LINAS. In Lied Library, there are 597 Ethernet drops currently available for students to plug in their laptops. Beginning in fall 2001, the capability of the system was expanded to include one of the UNLV branch libraries, the Architecture Studies Library. A dozen drops are present within that facility to accommodate users bringing in their own laptops.

## A Program for Circulating Library-Owned Laptops

In addition to permitting patrons to bring in their own laptops, UNLV also circulates library-owned laptops.

While the system described above involves a high degree of behind-the-scenes programming focused on security and authentication, the program of circulating library-owned laptops involved its own unique set of considerations. The idea of checking out laptops for patrons to use had been discussed at the library for two or three years prior to opening Lied Library. The discussion was somewhat moot in regards to the old library, as wiring and network equipment did not exist to accommodate such a program, nor were funds going to be invested in a facility that would soon be vacated. Early planning involved site visits to two institutions offering such programs. In addition, Lied Library plans called for network drops in myriad locations to accommodate future possible expansion of networked resources, including laptops. Later, after it was decided that the time had come to implement this service, a list of questions was formulated and talks began with other institutions offering such a service. This helped to avoid potential missteps in initiating the program at Lied. Twenty laptops were purchased in late spring 2001 and work was begun on configuration options, security implementations, and policy concerns.

In collecting information to start the program, other libraries were asked about their programs. A listing of several of these libraries, along with links to their respective laptop check-out programs, is provided in appendix D. Among the libraries surveyed, the number of laptops available for check out ranges from twenty to forty units. In addition to the base laptop, some libraries check out accessories such as an AC adapter or carrying case. At least one library did not check out AC adapters, letting the user rely only on battery power. This would help enforce a return of the laptop when the battery power was drained. One library provided spare batteries if a user requested them. Another library checks out AC adapters but does not include a battery whatsoever. This is the method followed at Lied, as it prevents battery theft, and power is available wherever laptop plug-in ports are located. In general, most libraries appear to check out and encourage the use of AC adapters. Libraries circulating battery-powered laptops will want extra batteries to avoid waiting for the battery to recharge before checking the laptop out to the next patron. Some libraries, such as the University of Virginia and the University of Maryland, have wireless laptops, and thus do not need to lend out patch cables. Otherwise, libraries check out patch cables with the laptops. Twenty-five-foot patch cables are circulated at Lied, allowing users to take advantage of practically any location they wish that is somewhat near a laptop drop. Some group study rooms have laptop drops in the floor, and by the time the patch cable is connected to the drop and winds its way to the top of the desk, there isn't much slack left. Each laptop circulated is a complete package, consisting of the laptop, AC adapter, patch cable, cheap headphones, and some basic general-user directions. It was decided not to circulate mice, as the laptops have two types of integrated pointing devices. Some libraries also circulate surge protectors and additional documentation. Following is additional information related to the laptop check-out program at Lied, as well as observations collected from other libraries with similar programs.

## Check-out and Check-in Procedures

Libraries vary in their check-out routines, though most keep track of what laptops are currently in use via their circulation system or other methods, such as an Access database. In general, most libraries don't allow holds, and, if they allow renewals at all, it is usually only if no one else is currently waiting for a laptop. Loan periods for other libraries ranged from two to four hours. Normally the laptops have to be returned an hour or so before the library closes each night. Steep overdue fines existed for practically every library. Lied charges twenty dollars per hour or part thereof. Some libraries charged in fifteen-minute increments, others by the hour or part thereof. One library charged five dollars an hour for an overdue laptop, another charged fifty dollars the first hour, and another blacklisted the user from checking out laptops in the future if it was returned late. While the majority of libraries check out laptops from the circulation desk, others circulate laptops from their media department, the reference desk, a combined services desk, or a PC lab desk. In Lied Library, the Media Resources department handles laptop circulations. They are well positioned to take on this task as the department has space to securely store the laptops, as well as service counter space to easily work with the user during the check out and return procedures. In addition, the staff was familiar with the circulation system, having circulated media materials. Libraries vary in the amount of attention given during the check out and return procedure. Some libraries have a staff member power-on the system in front of the user to ensure that the laptop is in good working order, instruct the user about where to plug in the patch cable and AC adapter, and show the user where the various drives are located. This also provides an opportunity to show the patron a laptop drop location map.

Most libraries require users to sign a use or liability statement or leave their IDs with the service desk. Some libraries require the user to sign the form only once, and make a note in their patron record that a form is on file. These use-forms require the user to agree with such statements as:

- I will comply with all university rules.
- I agree this laptop is in full working order.

- I am responsible for. . . .
- I understand the replacement and late return charges are as follows. . . .
- I will not tamper with the existing hardware and software.

The Lied laptop circulation form is found in appendix E. Patrons sign the form each time they check out the laptop. Some libraries perform a checklist of inspections when laptops are returned that is just as comprehensive as when they are checked out. These items include powering-on the laptop in front of the user to ensure everything is still in good working order, checking the floppy and CD/DVD drives, making sure the network cards are present and working, and looking for chips or cracks in the laptop casing. The check-in procedure can even be as thorough as taking out the network card and making sure it's the same one (serial number) that was there when the laptop went out. How easy it would be for a student to swap out a two-hundred-dollar library network card for a lesser one-hundred-dollar model. In developing a check-out program, a library may want to look for such items as laptops with integrated network cards and floppy disks. New laptop designs with multiple modular bays can be great, but they also make it easy for components to walk out the door. Users may sign the same form used for check out when they bring the laptops back, stating that the unit is in the same good working order, and agreeing to admit to and work with library staff to resolve any damage that may have occurred while in their possession. At least one library provides proof-of-return receipts after the laptop is returned. Some libraries, having such a popular check-out program, can check out a laptop in minutes, if not seconds, after it is returned, and thus rely more on the end user to report any problems. The staff at Lied follows the thorough model of powering-on the laptop upon both check out and return to help minimize any potential for abuse.

All libraries assess damages sustained to laptops, or have the groundwork to do so when the situation arises. Some libraries report that they have had to assess damages fines, while others report that they have yet to experience any major incidents. Should major damage be detected, some libraries have on file or on the user agreement form a damages rate chart. For example, a broken patch cable would result in no charge, but it would be a five-dollar replacement fee if the cable were missing. A missing drive might cost one hundred dollars, and a missing network card might cost two hundred dollars. All libraries had at the upper end a full replacement cost for stolen laptops or laptops that were totally irreparable. This price ranged anywhere from two thousand to four thousand dollars, depending on current market value. The full replacement cost at Lied is $2500. It was decided not to publish a damages rate chart, as damage can vary greatly depending on the situation. Instead, users are informed that they will be held responsible for any damage or replacement fees, up to the full replacement cost of the laptop.

## Maintenance and Security

In addition to staff time necessary to circulate laptops, libraries must be prepared to perform routine maintenance on the laptops, just as they do on other public PCs within their library. Libraries mention that it is typical to have four or five laptops unavailable for circulation because they are being repaired or otherwise undergoing maintenance. Some libraries use a General Hardware Oriented System Transfer (GHOST) to restore a fresh image to the laptop each time it is returned, known as ghosting, before it is circulated to another user. Some libraries moderate this approach and load a new image on all laptops on a regular basis, whether it be all laptops together once a week, or a certain number per day. The Lied laptops are ghosted on a regular schedule, at night, so that all receive a fresh image on a weekly basis. Some libraries install additional security software on the laptops to help keep them pristine from user abuse. Examples of such software include DeepFreeze, GoBack, or any of numerous other programs designed to lock down the units. In addition, simple batch files can perform such tasks as deleting files users may have downloaded or refreshing the wallpaper. Libraries which utilize such software may not ghost or reformat the hard drives on any regular schedule, relying on users to notify them of any major problems. At Lied, users are encouraged to save files to a floppy disk, and are not allowed to install their own software.

For the desktop PCs in Lied Library, an in-house authentication system was developed; future plans may call for implementing this system onto the circulating laptops. This system involves the use of the normal network client, several generic login accounts, and a Visual Basic script which queries the library's online database. The authentication process would be almost seamless to the user—when the laptop is switched on, a graphical screen appears in which information such as user name and library barcode number is entered. This information is sent to the Innopac patron database for verification, and, if determined to be an acceptable patron type, the user is permitted access to the computer. A log file would keep track of who is logged in, on what PC, and at what time. After authenticating, the normal Windows desktop and appropriate programs would be immediately available. For the present laptop circulation program at Lied, hard-copy user agreement forms are maintained, noting who had what laptop at what time. With DHCP, anyone who was responsible for any breach of network security or etiquette is pinpointed.

Software installed on the laptops includes Netscape and Internet Explorer Web browsers, FTP and telnet clients, and the full Microsoft Office 2000 Premium suite of productivity software. Physical security and theft prevention is a central issue in any laptop check-out program. Laptops are expensive, desirable, and easy to conceal in an academic environment where many other users have their own laptops and where book bags are prevalent. Without exception, all libraries surveyed did not allow laptops to leave the library. Nevertheless, several of the libraries investigated mentioned that laptops had been stolen from them. This ranged from a library reporting a single laptop stolen, to one library having a dozen stolen all at once from a locked storage cabinet.

Many libraries tattle tape their equipment, which activates the library security gates when a laptop is taken from the building. One library puts a garish-colored large sticker on the laptop to help identify those belonging to the library. A somewhat deep-end approach would be to apply sticker with superglue stating, "This laptop stolen from the A.B.C. Library." At UNLV, several strips of tattle tape are placed on the top cover and a large UNLV Library sticker is pasted down with heavy-duty wallpaper glue. In addition, tattle tape is inserted in the inside of the PC. As long as the tattle tape is not touching metal, it works reasonably well. Spray paint was used to stencil "Lied Library" onto the cover, in case the sticker came off. Another idea would be to use an engraving pen. One library mentioned that they would soon be installing Webcams to monitor the exits to the building and perhaps catch any laptop thefts. All libraries had stern warnings about never letting the laptop out of one's sight, and stated that the user is 100 percent responsible for stolen or damaged laptops. Only affiliated users (normally faculty, staff, and students) could check out the laptops, and usually only with a valid university identification card. Obviously, checking the laptops out via the library circulation system or another database helps to track not only the laptops but inventory numbers and other pertinent information. Holding a student's driver's license or student ID may also add a little incentive to return the laptop, though the possibility of fake IDs always exists.

## Financial and Staff Time Costs

An obvious question prior to starting such a check-out program relates to both initial startup costs, ongoing maintenance costs, and staff time costs. Initial startup costs include the laptop hardware, related peripherals one may want to check out (external mice, headphones, additional batteries, carrying case, patch cable), and software-licensing costs for the software loaded onto the laptops. At Lied, the initial cost per laptop for a complete unit (laptop, headphones, patch cable, carrying case, and software) totaled around $2500. At the time of purchase, while the laptops were cutting-edge, they were not based on the absolute highest iteration of processor, resulting in an overall savings of thousands of dollars for the twenty laptops. While traditionally laptops were both more expensive and less powerful than desktop units, this is changing. To put things in perspective, the Lied laptops are slightly more powerful than the desktop units that were ordered for the library a year earlier, and, everything considered, the same price. Since the laptops were purchased, they have dropped significantly in price, primarily due to both increased competition and various economic and supply factors.

Maintenance-related costs are another significant consideration. A laptop's compact and integrated design makes the unit, in general, a bit more challenging to maintain. Many components, such as network adapters, video controllers, and sound chips are typically integrated into the motherboard or associated chipsets. If a particular component of a laptop fails, (especially if it is a critical component), it may be necessary to send the unit back to the manufacturer, as it is often difficult, if not utterly inadvisable, for someone other than the manufacturer to attempt repair. This is obviously more costly than having local systems staff replace the components themselves, which can be easily done with desktop units. Related to this, libraries may opt to upgrade desktops in tight budget situations, rather than replace the units. Such typical upgrades may include memory, hard drives, video cards, or monitors. With few exceptions, laptop components can't be upgraded.

Laptops, in general, experience more wear and tear than desktops. Often, the only components of a desktop that a user typically touches are the keyboard and mouse. With laptops, they are touching not only these components, but handling the laptops more intimately. They are taking them out and putting them back into the carrying case. They are placing them in laps, on hard surfaces, or near drinks. They are plugging in patch cables, AC adapters, and perhaps swapping batteries. Considering the number of times such actions occur over a one-week period with a popular laptop check-out program, it is obvious that laptops receive significantly more general wear and tear. As an example, if a library has a standard four-year replacement cycle for fixed desktops, it might be prudent to knock this down to a three-year replacement cycle for laptops.

Other costs associated with the implementation of such a program are primarily staff-related, including routine maintenance and storage and circulation considerations, each of which requires some degree of staff involvement. Routine maintenance, such as ghosting machines with a fresh software image, doesn't take much more time than with fixed desktop PCs. Clearly, while desktops are

always connected to the network, prior to reimaging laptops, one must take the time to hook up x number of patch cables to x number of laptops to connect to the network and reimage. In a multicast scenario, using GHOST, in which multiple laptops can be reimaged at once, it can take a mere ten minutes or less to actually have the new image installed on all the laptops. This is followed by a few postregistration settings. Overall, reimaging takes a little time, but not much more than in a desktop environment. Twenty laptops could be easily reimaged in perhaps half an hour. If this is done once a week, it is not an excessively large amount of time, and student assistants could be trained for such work. Circulation considerations involve the fact that laptops have to be circulated (as opposed to fixed desktops), and have short circulation periods compared to books. If the program is popular, staff time involved in circulating PCs can be substantial, primarily depending on how much time a staff member spends with the patron upon check out and return of the laptops. If a thorough check of the laptop and components is performed, it can take up to ten minutes per circulation cycle of a laptop. Storage considerations involve making sure the laptop and all the associated components are in the storage case, and then stacked and locked up in a secure location at night. All of these items are necessary things to consider for a laptop check-out program.

## Popularity

Without exception, all libraries reported that their check-out program was very popular. At least one library has a two-part flip sign on the service desk that either reads "Laptops Available," or "No Laptops Available." This library stated that all of their laptops were typically in use by 11 A.M. The three most popular questions at this library were, "When will the next laptop be returned?" "Is there a waiting list?" and "Can I renew my laptop?" As mentioned above, laptops often recirculate just as soon as they are returned to the service desk. Most libraries don't allow reservations or holds, and allow renewals only if no one else is waiting. There is no mystery as to why the laptop check-out programs are so popular, and hence, successful. A laptop provides more privacy and quiet than using a regular, fixed-desktop PC. Students can take the laptop to a study room or a lounge chair on a floor far removed from the first floor hustle and bustle of the library. Libraries typically put an identical software suite on the laptops as found on their desktop student PCs, and users can print to the library network printers just as easily. Laptops are good for group gatherings in private study areas, and are compact, allowing users to spread out their schoolwork on a tabletop. To help with the potential questions of such a popular program, it's smart to have a well-developed Web page discussing such things as eligibility for checking out laptops, policies, a map of available plug-in locations, a hardware description of the laptop, software available, and perhaps even photos illustrating how to connect the patch cables and AC adapter.

## Conclusion

Mirroring the trend that laptop computers are increasing in popularity, students are ever more eager to use such technology in an academic setting. Many libraries provide students with the privilege of checking out library-owned laptops, or allow them to bring in their own laptops and connect to the campus backbone. Both of these services are offered at the UNLV Lied Library. With both services, it is important to have policies and procedures promoting network security. Circulating library-owned laptops requires careful considerations related to proper treatment of the laptops, circulation parameters, budgetary requirements of starting up and supporting such a program, and staff time considerations related to the maintenance and circulation of such laptops. Overall, despite such costs and considerations, laptop programs tend to be very successful and popular within the academic community.

## References


1. Joris Evers, Maker of Kournikova Virus Stands Trial, Accessed September 15, 2001, www.nwfusion.com/news/2001/0913annavirus.html.
2. W. Richard Stevens, *TCP/IP Illustrated, Volume 1: The Protocols* (Reading, Mass.: Addison-Wesley, 1994), 54.
3. Ted Lemon and Ralph Droms, *The DHCP Handbook: Understanding, Deploying, and Managing Automated Configuration Services* (Indianapolis: New Rider's Pubs., 1999), 195.
4. Lemon and Droms, *The DHCP Handbook*, 193.

## Appendix A. Components of LINAS

**Executable Scripts**

/LINAS/bin/leasesd — Tracks the /var/state/DHCP/DHCPd.leases file for updates, parses updates, and stores the info in the leases table of the database

/LINAS/bin/refreshd — Kills running DHCP daemon and restarts it. This forces a refresh of the DHCPd.conf file, thereby activating additions to the approve/deny lists.

**CGI Scripts**

/home/httpd/cgi-bin/regPatron.cgi
/home/httpd/cgi-bin/regAdmin.cgi
/home/httpd/cgi-bin/nav.cgi
/home/httpd/cgi-bin/logout.cgi
/home/httpd/cgi-bin/gatekeep.cgi
/home/httpd/cgi-bin/auth.cgi
/home/httpd/cgi-bin/admin.cgi

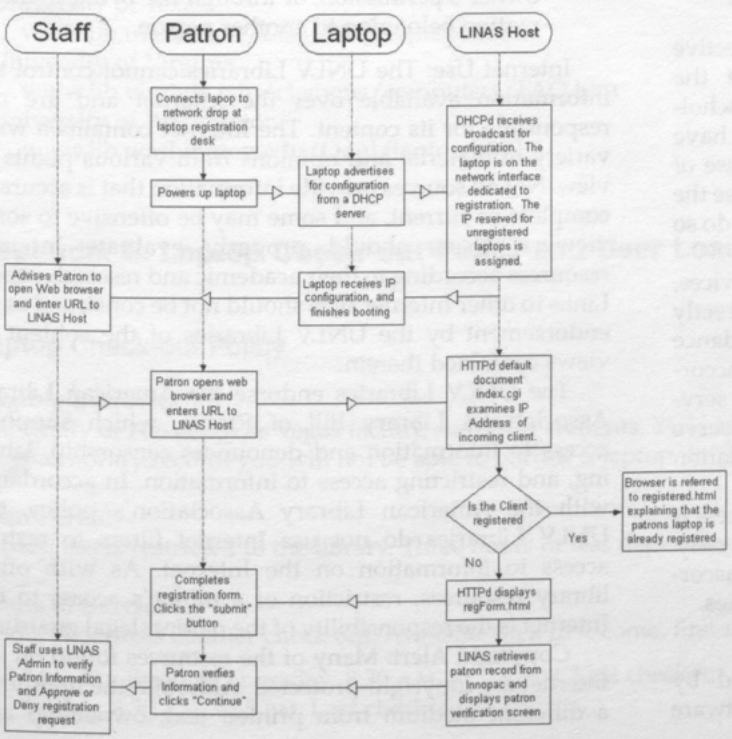
**Perl Modules** /usr/lib/perl5/site_perl/5.*/

db.pm
DHCP.pm
pViewBrowse.pm
ParseOPAC.pm

**Configuration Files** /etc

DHCPd.conf
leasesd.conf

## Appendix B. LINAS Registration Flow Overview



## Appendix C. Policies Governing Users Bringing in Their Own Laptops

### Laptop Use Agreement

1. I have received a copy of the UNLV Libraries' laptop use policy and I agree to abide by it.
2. I will report theft or loss of my laptop computer or its network card to the library.
3. I understand that I may be held responsible for any damage done to the network from my laptop.
4. I understand that my right to use the network may be revoked if I do not comply with the policies governing network use.

Name: (please print) _____
Signature _____ Date _____

### UNLV Libraries' Laptop Use Policy

Laptop Use: Students, faculty, and staff of the university are welcome to bring laptops with network cards and use them with our data drops to gain access to our network. The laptop must be registered in our laptop authentication system, and a valid library barcode is also required. Users are responsible for notifying the library promptly if their registered laptop (or its network card) is lost or stolen, since the user may be held responsible if their laptop is used to access and damage the network. Users taking advantage of this service are required to abide by all UCCSN and UNLV computer policies.

In pursuit of its goal to provide the most effective access to information resources in support of the University's programs of teaching, research, and scholarly/creative production, the UNLV Libraries have adopted policies governing electronic access and use of licensed software. It is expected that all those who use the libraries' network to access electronic resources will do so responsibly, respecting the rights of other users.

Authorized Users: Electronic information, services, software, and networks provided directly or indirectly by the UNLV Libraries shall be accessible, in accordance with licensing or contractual obligations and in accordance with existing UNLV and UCCSN computing services policies. The UNLV Libraries and their staff reserve the right to deny network access if they are in violation of any part of this policy.

Authorized and Unauthorized Use: Network access is to be used for academic research purposes only. Internet/World Wide Web searches must be in accordance with system and campus computer-use policies.

Users must not:

A. Copy any copyrighted software provided by UNLV. It is a criminal offense to copy any software that is protected by copyright, and UNLV will treat it as such.
B. Use licensed software in a manner inconsistent with the licensing arrangement.
C. Copy, rename, alter, examine, or delete the files or programs of another person or UNLV without permission.
D. Use a computer to annoy others, including, but not limited to, sending offensive messages, or knowingly causing a system crash.
E. Create, disseminate, or run a self-replicating program ("virus"), whether destructive in nature or not.
F. Use a computer for nonuniversity work, such as for private business, or clubs not sanctioned by UNLV.
G. Tamper with switch settings, move, reconfigure or do anything that could damage terminals, computers, printers, or other equipment.
H. Collect, read, or destroy output other than your own work without the permission of the owner.
I. Use the computer account of another person with or without their permission unless it is designated for group work.
J. Use software not provided by UNLV in the lab unless the student is legally authorized to do so.
K. Access or attempt to access a host computer, either at UNLV or through a network, without the owner's permission, or through use of log-in information belonging to another person

Internet Use: The UNLV Libraries cannot control the information available over the Internet and are not responsible for its content. The Internet contains a wide variety of material and opinions from various points of view. Not all sources provide information that is accurate, complete or current, and some may be offensive to some viewers. Users should properly evaluate Internet resources according to their academic and research needs. Links to other Internet sites should not be construed as an endorsement by the UNLV Libraries of the content or views contained therein.

The UNLV Libraries endorse the American Library Association's Library Bill of Rights, which supports access to information and denounces censorship, labeling, and restricting access to information. In accordance with the American Library Association's policy, the UNLV Libraries do not use Internet filters to restrict access to information on the Internet. As with other library resources, restriction of a minor's access to the Internet is the responsibility of the parent/legal guardian.

Copyright Alert: Many of the resources found on the Internet are copyright-protected. Although the Internet is a different medium from printed text, ownership and

intellectual property rights still exist. Check the documents for appropriate statements indicating ownership. Most of the electronic software and journal articles available on library servers and computers are also copyrighted. Users shall not violate the legal protection provided by copyrights and licenses held by the UNLV Libraries or others. Users shall not make copies of any licensed or copyrighted computer program found on a library computer.

## Appendix D. Additional Information from Libraries Offering Laptop Programs

Colorado State
manta.library.colostate.edu/access/policy.html
Concordia University
http://library.concordia.ca/ibooks
Oregon State University
http://osulibrary.orst.edu/computing/laptop1.htm
Penn State
www.libaries.psu.edu/crsweb/laptop/withid.htm
University of California, Los Angeles
www.clicc.ucla.edu/info/laptops.asp
University of Georgia
www.rx.uga.edu/main/home/cir/LaptopCheckOutPolicy.htm
University of Maryland
www.wireless.umd.edu/adchkout.html
University of Nevada, Las Vegas
www.library.unlv.edu/services/laptop
University of North Carolina
www.lib.unc.edu/circ/laptops
University of Oklahoma
http://faculty-staff.ou.edu/R/Sarah.E.Robbins-1/laptoppolicy.htm
University of Texas
www.lib.utexas.edu/Libs/PCL/laptops.html
University of Virginia
www.lib.virginia.edu/clemons/computers/FAQ.htm
University of Washington
www.lib.washington.edu/Ougl/laptop.html

## Appendix E. Laptop Check-out Policy and User Loan Agreement Form

### Laptop Check-out Policy

**Eligible Borrower**
University of Nevada, Las Vegas faculty, staff, and students. YOU MUST HAVE YOUR UNLV I.D. If you have any fines on your library record, you will not be able to borrow a laptop until the fines are cleared.

**Loan Period**
Laptop use is restricted to the library. Three hours or less depending on what time laptop is checked out. No renewals.

**Hours of Service**
Media Resources counter check out: Available on a first-come, first-served basis. No reservations or holds taken.

Monday through Thursday: 8:30 A.M. to 8:30 P.M. Last checkout at 7 P.M.
Friday: 8:30 A.M. to 5 P.M. Last checkout at 3:30 P.M.

Weekend hours of service:

Saturday: 10:30 A.M. to 6 P.M. Last checkout at 4:30 P.M.
Sunday: 11:30 A.M. to 8:30 P.M. Last checkout at 7 P.M.

**Responsibility**

Agreement form will be signed by the user that covers their responsibility and verifies condition of laptop at checkout and checkin.

The user who checks out the laptop is responsible for any loss or damage until the laptop is returned to the media resources desk. The user is responsible for up to the full replacement cost ($2500) if the computer is damaged or stolen.

All user files must be saved to a disk before returning laptop to media resources. Hard drives will be wiped clean of any files when returned to media resources. The library is not responsible for any lost files.

**Return of Laptops**

Laptops must be returned to a media resources staff member and not just left on the counter when returned. When returned, staff will verify that the laptop is in good working order before signing off on the user agreement form. This will take ten to fifteen minutes and user must be present.

In case of emergency evacuation of library please take the laptop with you and return to media resources when building is safe to reenter.

**Fines/Late Charges**

There will be a $20 late charge for each hour or portion thereof that the laptop is returned beyond the due time. The maximum overdue fine is $100. These fines will be attached to the patron's library card, which could affect students registering for classes or obtaining diploma/transcripts until paid in full.

## Laptop Computer Loan Agreement Lied Library UNLV

- I accept full responsibility for the laptop and accessories I am borrowing. I will reimburse the University of Nevada Las Vegas for the cost of repairing or replacing this laptop or accessories if they are damaged, lost, or stolen while checked out in my name.
- I understand that the replacement cost for this laptop computer will be no less than $2500.
- I understand the overdue fines are $20 per hour, $100 maximum overdue fine.
- I agree to comply with University rules regarding computer use.
- I understand that the circulation period is for three hours or less depending on the time the laptop is checked out.

Library use only. Please do not leave the laptop unattended.

Your signature below indicates that you recognize your responsibility in the care and custody of any laptop you borrow from Lied Library.

### Laptop Check Out

Physical inspection verification by staff
Fill in all blanks and initial corrections
__ Laptop computer # _____
__ PCMCIA slot/battery cover
__ Exterior appearance O.K.
__ Cat 5 cable
__ Successful boot-up
__ LCD screen
__ Floppy drive
__ Memory 130,528 KB RAM
__ DVD-ROM
__ Case
__ AC adaptor
__ Headphones

I have witnessed the physical inspection of the laptop computer and components. All checked parts are present and appear to be functioning.
Borrower signature_____
Print name _____
Date_____
Time checked out _____
E-mail address_____
Media staff name _____

### Laptop Check-in (Return)

Physical inspection verification by staff
Fill in all blanks and initial corrections
__ Exterior appearance O.K.
__ All checked-out components returned
__ Successful boot-up
__ Memory at 130,528 KB RAM
__ DVD-ROM/floppy drive ejected
__ Computer properly shut down
__ LCD screen
__ Documents off hard drive (ask patron)
Time checked in _____
Media staff name _____
Borrower signature _____

Comments, technical problems, or damage: please make comments or report problems on the back of this agreement and give it to Media Resources Staff.
Fines accessed: _____
Reason: _____

# Building Database-Backed Web Applications: Process and Issues

Stephen Westman

*In this article, the process of implementing database-backed Web applications from initial planning to final implementation and beyond will be explored. Both technical and "people" issues will be addressed, including choosing hardware and database platforms; selection of development tools and approaches; making decisions on the allocation of time, personnel, and equipment; and assigning data and system administration tasks.[1] It is necessary to consider these points, even if a project seems simple, to ensure that no issues have been overlooked. Carefully thinking through all aspects beforehand can minimize the number of problems one will encounter later.*

There is a growing body of literature on database-backed Web pages, both on general topics and specific platforms.[2] The library press is also beginning to publish in this area, including articles by Antleman, Mischo and Schlembach, Kiehl, Summers, and Westman.[3] There are several clear advantages of using a database rather than relying solely on static HTML pages: (1) it allows one to focus on inputting rather than formatting data; (2) it reduces the time required to maintain Web pages; and (3) it permits the reuse of information in a variety of situations and to serve a multiplicity purposes. However, when implementing this technology, one needs to undertake careful planning and weigh the costs versus the benefits. One will need to address a number of issues and make a number of decisions. Knowing what to expect will help to ensure success.

## Planning

### Deciding What You Want

When beginning, it is important to define the desired goal of the project and determine whether a database is the appropriate tool for that goal. Database-backed applications generally:

- use structured data where output via different parameters such as subject or material type is desired;
- enable searching of the data by various parameters;
- manage constantly changing data; and
- allow the reuse of information in different contexts and pages.

### Get Everybody on Board

Once it is determined that a database is the appropriate tool for the job, it is necessary to get buy-in from two areas, administration and staff, before proceeding. Database-backed Web projects can consume significant staffing resources, may require a substantial financial investment in the short-run, and involve ongoing support in the long run. It is therefore critical that the library administration understands and fully supports the project. To obtain this understanding and support, answers must be provided to a number of questions:

1. What will be the benefit of this project? Answers can include saving time, providing better access to information, or allowing for more efficient workflow.
2. What will the costs be in terms of money and staff time? What will need to be dropped to do this? What new hardware, software, and personnel will be needed? Can it be done with existing infrastructure?
3. How is this project to be accomplished? What is the timeframe for its implementation? What are the steps to be taken and when should each step be completed? What expertise is needed? Does such expertise exist in-house or can it be acquired?
4. How much time and effort will it take to maintain this application? How much data inputting and updating will be involved? Will it be worth the investment? What subsequent programming and technical resources will be needed?

It is also important to identify the persons who will fill various roles in the project. Note that one person may perform more than one role. Necessary functions include:

- Project coordinator—the person who oversees the project and makes sure things stay on-track
- Data users—the people who will help define the data and the functionality to be included in the application
- Designer—the person who finds out what data are needed and then creates the required data models
- Database administrator—the person who implements the data model and who administers the database management system (often the same person as the designer)
- Developer/programmer—the person who programs the application
- Testers—users who test the application[4]

The project coordinator, by having the answers to these questions when approaching the library administration, can provide the administration with the information needed to make an informed decision. Having these answers makes certain that the coordinator has thought the issues through, thus ensuring the project's success.

**Stephen Westman** (westman.2@osu.edu) is Digital Resources Systems Administrator for The Ohio State University Libraries, Columbus.

Staff members also need to support the project. Since the new system will involve changes in their work, they need to understand the reasons for the changes and to support them. In particular, they need to be shown how this new application will make their work easier. By involving staff in the application from the beginning, they will have an investment in its implementation.

### Think Big, Start Small

When starting the project, one should think big but start small. Thinking big means looking at the project from a larger perspective. What other uses might there be for the proposed database? Are there other types of information that could benefit from being made available in such a manner? If so, what are the relationships to the project under consideration? It is necessary to ask these sorts of questions to help assure that the project will work well with future products that the library might develop as well as permitting easy integration of new features into the proposed application at a later time.

Starting small means undertaking a project that realistically can be accomplished. This type of project may involve many people and will require planning, coordination, and communication. Staff will need to allocate time for evaluating, purchasing, and installing the hardware and software required to create the application. Starting small allows staff to define issues and to learn how to use and develop the tools and techniques needed for successful implementation. It also provides an idea of ongoing maintenance requirements and allows for gradual integration of the final product into the library environment.

Note that, while thinking big ensures the product will be flexible enough to support future development, it is important not to create false or unrealistic expectations. Getting people excited about possibilities that cannot be fulfilled can create frustration that could endanger the acceptance of the current application and possibly engender resistance to any new projects.

Once the project is defined, the project coordinator should establish a timeline to provide structure to the development process. This timeline should list each task to be done and indicate the order in which the tasks should be accomplished. It should also include realistic target dates with the names of those responsible for each task's completion. This schedule is valuable in making important time, money, and personnel allocations, as well as helping to keep the project on track.

## Defining the System

The next step is to define the application. As Greenspun points out, this requires one to:

> take [the] publishing idea . . . and apply three standards steps:
>
> 1. Develop a data model—Define what information is going to be stored and how it should be represented
> 2. Develop a collection of legal transactions on that model, e.g., inserts and updates
> 3. Come up with a set of Web forms that let the user build up to one of those transactions.[5]

## Develop a Data Model

Data modeling is perhaps the most critical part of the entire process and one that is too often overlooked. The designer will save a lot of time and avoid many problems and frustrations by focusing on proper data modeling before attempting implementation. While a full description of data modeling is beyond the scope of this article, there are several good books on the topic. One of them is Hernandez's *Database Design for Mere Mortals*.[6] In clear, nontechnical language, he gives a step-by-step process for doing interviews to obtain needed data elements from users; organizing those elements into groups and setting up relationships between the groups; and defining the constraints on the data (rules to which data elements input into the system must conform). To assist the reader, he provides example scenarios at each step. Because his focus is on design principles rather than specific implementation or programming techniques, the book can be useful, no matter what infrastructure may be used for the project.

Briefly put, data modeling is the process of defining all of the data to be included in the system and then structuring them to allow their easy and efficient maintenance and retrieval. To obtain this information, the designer talks with staff members that will work with the data (both inputting and using) to find out what data are needed and how they are to be used. This process includes asking a number of questions, such as:

1. What data are needed?
2. How do the data interrelate and how should they be organized?
3. What sorts of outputs (both Web pages and printed reports) are needed?
4. Who should be able to enter and access this data?
5. How can the design be made flexible enough to encompass future needs?

In the process, certain information about the data needs to be obtained. Questions that should be asked about each item include:

1. What type of datum (text, numeric, date) is it? How big?
2. What types of constraints should apply to it? For example, should users be able to enter anything for

this item or should they select from a predefined list of possible entries?
3. Should the user be able to search on it?
4. Does it need to be repeatable?
5. Is it mandatory? Optional?
6. Should access be restricted? If so, to whom?

The designer then groups the data elements groups and defines how each group relates to the others. Once this has been done, the results should be verified with the staff to determine if there are mistakes or whether additional data elements need to be included. This process goes on iteratively until the designer feels that all of the basic pieces have been discovered and entered into the model.

As part of the modeling process, the designer needs to know how users will want to retrieve information from the database and what types of output are desired. What fields will they need to be able to search by? What fields should be displayed? Is all desired content included? Is it organized in a logical manner? Will the proposed model make it easy to retrieve the data? What sorts of reports will be needed and how will they be done? The designer needs to ask these questions both to optimize the model to support all desired searches and to validate the data structures that have been developed.

## Define Interactions

Once the designer feels that a fairly complete model has been achieved, it is necessary to define the application processes needed to implement the system, and then to determine each step (interaction) within those processes required to get data into and out of the database. For each interaction, the designer defines the data elements involved and creates the structured query language (SQL) queries required to enter or retrieve information into or from wherever it resides in the proposed model.[7] This ensures that all data elements are present and validates (or show weaknesses in) the proposed model.

A process may involve multiple tasks. For example, adding a URL into a Web collection database may involve entering basic information for the site, assigning it subject headings, providing online licensing and support data on the resource, and entering public help-screen information. The designer may wish to break the process up and assign individual steps to different persons. In developing the system, one should ask:

1. Can this process be broken down into individual tasks?
2. Who will do each task?
3. How do these tasks fit into the existing workflow?
4. If a task doesn't fit in well, can it be changed? Can the workflow be changed?
5. What security levels should be assigned to each interaction?

By breaking processes down in this way, the designer is able to build a high degree of flexibility into the system and permits tasks to be reassigned as library structures and workflows change.

## Define the User Interface

Along with defining the interactions, it is necessary to design the HTML forms to be used in implementing each interaction within the application. These screens provide the primary window into the data; their design needs to be clear, allowing for easy navigation and use. Therefore, user input is crucial. The developer should create mockups of the application's inputting, editing, and searching/outputting screens and show them to users for feedback. User input can then be folded back into the screens' design. The process goes on iteratively until agreement is reached.

This iteration process may go back to any of the three steps listed above: data modeling, interaction definition, or screen design. The screen layout may change based on user input, allowing the user to view data and see a model of the final application elements and processes in ways that are more familiar. This then allows the user to see where additional data elements (or even processes) are required. Taking this approach ensures that the most complete data model is developed and allows for the building of an application that integrates well into the library workflow.

It is important that the designer and user(s) come to complete agreement on all three of these steps before proceeding with implementation. Adding further data fields or changing how the system works after this point could easily lead to significant delays due to the large amounts of time needed for data restructuring and additional programming or reprogramming. A generally accepted practice is that once agreement has been reached, a written document (containing documentation of the data models, transactions, interface screens) is created and all parties sign off on it. Any further changes are then subject to separate negotiations as to the possible effects of the proposed changes on the project's timeline.

## Implementation

### Select a Hardware Platform

For the most part, a full range of database and interface tools is available for both Windows and Unix platforms.

While a library may already have hardware in place that could do the job, there are a number of points the project coordinator should consider when deciding on a hardware platform and determining whether existing equipment could be used:

1. Servers must be adequate to support the new application. Issues such as disk storage space, required RAM, 24/7 availability, and supportability need to be considered.
2. How many simultaneous users need to be supported? Larger numbers of users mean more multiple simultaneous queries, requiring more robust infrastructure.
3. Is the proposed hardware scalable? Is it powerful enough to support future growth (both in terms of number of applications and number of users)?
4. What kinds of data are involved? Graphic images require more bandwidth and utilize more storage space than text.
5. Do the pages need to be created each time a user requests the page? Dynamic Web pages, while allowing for real-time availability of new information, can exact a serious performance price for often-accessed information on heavily used servers.
6. How complex will the data structures be? If the database has a large number of tables with multiple relations, it will take more computer processing time to get the answer to an information request than if only a few tables are involved.
7. What other applications are running on the same computer? What are their requirements?

In looking at these issues, one should note that the database management system does not need to reside on the same computer as the one that runs the Web server. Tools such as Database Interface (DBI), Open Database Connectivity (ODBC), Java Database Connectivity (JDBC), and others makes it possible to place the database on one computer and the Web server on another and then to establish communication between the two.[8] This allows the load to be distributed and permits the use of a different platform for the database than is used for the Web server.

When selecting a machine and operating system on which to run the database, there are some rules of thumb that the coordinator should consider when choosing between NT and Unix:

- Scalability—Either NT or Unix will be more than adequate for less complex or fewer numbers of applications. However, given that NT runs slower on a given Intel machine than Linux does, the latter is significantly more scalable (can support more activity) on a per-dollar basis.
- Affordability—Probably the most affordable solution would be a Pentium server running Linux. With the long tradition of sharing of Unix programs (particularly in the open source movement), one can easily assemble a fully functioning, fully featured database/Web server for around $1,000. NT on Intel is probably the next most affordable option. However, given scalability issues as well as licensing costs, NT will require a greater financial outlay.
- Staffing—NT is clearly the better-known operating system and, given the wealth of tools written for NT and the greater likelihood of finding NT expertise in-house, it probably has the edge in this area. However it is good to consider alternative approaches. In particular, the growing popularity of Linux means that it may not be as difficult as one might think to find Unix support at an affordable price.
- Ease of development—Traditionally this has been one of NT's strengths. However, with the increase in the number of development tools becoming available on both platforms, development on Unix/Linux servers is becoming easier.
- Security—Both operating systems need to be made secure, but there are tools to help the system administrator do so. While most flavors of Unix have a reputation for not being secure, the increasing number of exploits aimed at NT-based servers (particularly against Microsoft's IIS Web server) makes this point a draw.[9]

In deciding which platform to use, one should balance the pros and cons in each of these areas, giving each the proper weight based on the needs and resources of the particular institution.

## Select a Data Management System

Although there are numerous ways to store data, the use of a good relational database management system (RDBMS) is recommended. An RDBMS not only provides a means of storing the data, it also provides built-in mechanisms (APIs and SQL) for adding, editing, and deleting records, as well as searching and outputting data.[10] RDBMS products are available at three levels: free, workstation (inexpensive), and enterprise (expensive) database systems.[11] Table 1 provides a list of some of the more popular RDBMS products that can be used for Web database publishing.[12]

The project coordinator should consider several parameters when deciding which RDBMS product to use:

1. What products are already available in-house?
2. What can be afforded?
3. What features are desired?

4. What expertise is available?
5. Is the desired product scalable?[13]

When considering cost, one should look beyond just the price of the product in question. If certain features are needed, it may ultimately not be cost effective to purchase a less expensive product and then have staff program in the missing features. Not only could that end up costing a great deal more in terms of staff time, the features may not work correctly, creating problems that would not have existed had the more feature-filled product been chosen. In other words, the more expensive product may end up being less expensive in the long run.

When using an RDBMS, one needs to address the matter of who will administer the database. Database administration responsibilities include setting up and administering the RDBMS; creating databases and tables; and developing, implementing, documenting, and maintaining data models implemented within the RDBMS. Duties also include backing up (and restoring) databases and tables; importing and exporting data; ensuring database security and integrity; and working with programmers to get data into and out of the system. As the person who understands and maintains all the databases and how all of the data within them interrelate, the administrator is the one responsible for trouble-shooting any database problems that occur.

If it is decided that the scalability, flexibility, and features of an enterprise RDBMS are required, then an administrator who is able to support those features will be needed. This person would be responsible for properly configuring the system—defining adequate swap space, defining file mappings on various disk drives, setting block sizes—and undertaking performance tuning. In addition, given that these packages are very large and complex, this person would need to deal with bugs that will inevitably be encountered by developing workarounds or installing patches. Failure to have expertise available that can take on these kinds of tasks could result in significant down-time and major data problems. Thus, if an enterprise RDBMS is required for a project, a competent database administrator needs to be included in the financial calculations.[14]

## Select Development Tools

It is necessary to decide on which program development tools to use to create the application. Tools come in four types: programming languages, server-side scripting languages, application servers, and programs that create static pages. The first involves a programmer writing a CGI application using a traditional programming language (such as C/C++, perl, or VisualBasic) to get data in and out of the RDBMS.[15] While this approach provides the most flexibility and power, it has several drawbacks:

- It is extremely expensive in terms of time
- It requires a much greater level of developer expertise
- It involves more lines of code (with a proportionately greater possibility for bugs)
- It scales poorly for multiple or large projects in terms of development costs

Alternately, one can decide to use a server-side scripting language such as mod_php, VBScript, or Active Server Pages (ASP) (see table 2).[16] This approach involves writing code directly into HTML documents rather than as stand-alone programs. While server-side scripting does require some programming expertise, it is relatively minimal. In addition, developers have a number of resources to assist them, including books, function libraries, user's groups, and discussion forums. Also, since it utilizes HTML-based pages, it allows nonsystems staff to be involved in the program development process.

A closely related approach, using an application server such as ColdFusion, also requires programming expertise, but somewhat less than the previous two. Database interactions are carried out through the use of ColdFusion Markup Languages (CFML) tags (which look like HTML tags) embedded in the HTML document. In

**Table 1.** Database Products

| Product | Wintel | Unix |
|---|---|---|
| **Free** | | |
| FrontBase | Yes | Yes |
| MySQL | Yes | Yes |
| mSQL | No | Yes |
| PostgreSQL | Yes | Yes |
| **Desktop** | | |
| FilemakerPro | Yes | No |
| Microsoft Access | Yes | No |
| FourthDimension | Yes | No |
| MS SQL-Server | Yes | No |
| **Enterprise** | | |
| DB2 | No | Yes |
| Oracle | Yes | Yes |
| Informix | Yes | Yes |
| Sybase | Yes | Yes |

**Table 2.** Scripting Tools and Application Servers[17]

| Product | Language | Servers supported | Windows | Unix |
|---|---|---|---|---|
| ASP | VBScript | Microsoft IIS | Yes | No |
| Apache::ASP | perl | Apache | Yes | Yes |
| ColdFusion | CFML | Apache, Microsoft IIS, Netscape | Yes | Yes |
| EmbPerl | perl | Apache | Yes | Yes |
| mod_php | php | Apache, Microsoft IIS, Netscape | Yes | Yes |

addition, ColdFusion permits development of new tags if those that come with the program are not adequate.

Finally, one can use tools such as Access2000, GDIdb, and others to take a database and output HTML pages that can then be moved to the Web server.[18] However, with the exception of GDIdb, this approach requires that each page be manually created and published to the Web site. Since this approach requires a pre-extant database, the developer may still need to write data maintenance routines to add and edit data, particularly for multitable applications.

## Coding

Once the design process is completed, the necessary hardware, RDBMS, and development software have been installed and configured, and the data models have been implemented in the database, one can proceed to actual code development. This part of the process involves creating the HTML forms; creating code to process the data from those forms to build appropriate SQL queries; connecting to the RDBMS server and sending queries and retrieving results; and formatting query results returned by the database.

While the specifics on how to code vary according to the development tool selected, there are certain tool-independent rules of thumb to make the process easier. These include:

- Make sure that HTML form-variable names are the same as the name of the field into which the data will be placed. This makes programming and debugging the application much easier.
- Break the application down into sections and then program and test each section individually.
- Where the same task is done often, consider creating a function and putting that function into a separate file to be included as needed. This allows this same code to be used in multiple applications, thereby speeding up development of new projects and making it easier to maintain multiple applications.
- Investigate code libraries, script collections, and other Web programming sites as well as checking with colleagues. Often what one is trying to do has been done before and the developer has made it available. Significant amounts of time can be saved by not reinventing the wheel.
- Document the program as it is being developed by writing comments and explanations into the code itself. This is critical to the ongoing support of the application and one that is all-too-often ignored. There are two rules of undocumented code: The first is that no programmer will immediately understand someone else's code. The second is that no programmer will easily understand his or her own code six months later. Without proper code annotations, any programmer who needs to change or fix a section of code may need to relearn the entire program in order to properly fix it, resulting in significant down-time.[19]

An important part of the development process is to eliminate bugs (program glitches) before putting the product out for general use. As noted above, part of the planning process is defining the application program flow. The developer should use that information to develop formal testing procedures to verify that the program is working correctly. Then the developer separately tests each transaction with the database, verifying that each data element is being added and updated properly, and that all desired constraints are being observed by the application. By formally defining and performing these testing procedures, the developer will save much time and frustration down the road.

Once the application has gone through these testing procedures, it is ready for end users to try out. While problems may still occur at this point, they should be relatively minor. There are a number of things to remember about the process of end-user testing:

1. The developer should work directly with the testers so that, as problems are found, it is possible to see what is happening and, if feasible, fix the problem on the spot

2. There should be enough data in the database for users to test all desired functionality
3. Searchers should know what to search on to get actual results
4. As noted above, this is not the time to change things. Any changes to the system should be put off until a subsequent version

## Production

Once the application has been fully tested, it can be put it into production. Note that this is not an all-or-nothing decision. For example, one may create data maintenance routines first and leave other aspects, such as public access modules, for later implementation. In such cases, a copy should be made both of the application code and the database and put into a separate testing area, where they can then be used for further system development. Under no circumstances (aside from fixing bugs) should work be done on the production version. Otherwise, the result will be an unstable (and for the user a very frustrating) application.

Note that, once a complete application is in production, the work does not end. Bugs will turn up, features won't work as expected, fields may need to be added or changed, and new features or enhancements may be desired. One should therefore plan for ongoing application support.

As parts of the program are implemented, developers should provide documentation on how to use the application. While the staff, at least initially, will be fairly knowledgeable due to their experience in the development process, they may leave or forget what they have learned. Having documentation, no matter how easy the application may seem, can save great amounts of time and frustration for new users trying to learn the system.

## Issues

### Data Backups

Among the ongoing tasks that will need to be done, none are more important than doing regular backups. However, backing up a database, while seemingly a simple matter, can be a huge disaster waiting to happen if not done correctly. If a user submits an interaction that modifies data in the database while the database is being backed up, the result could be a backup containing pre-interaction (unchanged) content in some tables and post-interaction (changed) contents in the others. The result would be a backup that would be unusable in case of an emergency. As Greenspun puts it: "Be afraid. Be very afraid. Standard Unix or Windows NT file system backups will not leave a consistent and therefore restorable database on tape."[20]

There are three ways around this problem:

1. Take the database offline (or put a read-only lock on it) before the backup is started and then put it back online (or take the read-only lock off) when the backup is completed.
2. Put the database into read-only mode and do a complete dump of the data to disk and then take the lock off again. Due to the potential amount of disk space required, this is the least scalable of the three solutions.
3. Use specialized backup software that will maintain the integrity of the data while keeping the database up and running[21]

Given that most library applications do not require the ability to do data maintenance on a 24/7 basis, it should not be a problem to take the first course. However, it is important to be aware of the issue and to plan accordingly.

Since backups are usually done only once a day, applications should log all interactions that update the database to a journal file. This logging may be done either by the database itself or the developer may code that feature into the application. Not only does having a journal file make trouble-shooting significantly easier, it means that if there is a database crash, the administrator can restore the database from the backup system and then reapply any subsequent transactions from the log files to bring the system back up to date.

### Data Integrity

In writing data maintenance routines, the programmers need to ensure data integrity—making sure that, as information is added, modified, and deleted, the information in the database remains reliable. Failure to maintain this integrity can create a data disaster. There are a number of ways in which data integrity can be compromised:

1. Two people trying to make changes to the same record at the same time
2. Editing routines that do not properly maintain links between related records
3. Programs that end without completing the creation or elimination of links between records

One can address this problem in a number of ways:

1. Locking—by locking data-affected areas in the database when an interaction is taking place, no other interactions against that data area can be undertaken until the first one has completed. This locking can be done either at the table or the record

level, depending on RDBMS capabilities. Code usually needs to be included in the application to invoke these capabilities.
2. Transactions—transaction is a technique where the RDBMS waits until all queries have successfully completed, at which point they are *committed*. If there is a problem, a rollback can be performed to return the database to where it was before the transaction was attempted. Most of the databases mentioned above provide transaction support.
3. Foreign key integrity support—In a relational database, two tables are related if records in each table have fields that share a common value. If the value in the common field in one table changes, the records in the other table's common field containing that value must also have the value changed or the relationship will be broken. RDBMS products with foreign-key integrity are able to automatically maintain these relationships without requiring additional programming. If the RDBMS does not provide this feature, the programmer will need to code it into the application.[22]

One further data-integrity issue that the developer needs to address is that of validating user-data entry. Even though an application works correctly, users may create data problems by inputting anomalous data into fields, such as entering letters into a number field or mistyping a word. There are a number of approaches that can address this type of problem:

1. Rather than typing information in, one can create option lists, checkboxes, or radio buttons from which users select an entry.
2. Insert code that checks user input. If they have input incorrect information (or even no information), the program can then warn the user and not complete the interaction until the problem is corrected.
3. Automatically enter certain types of information, such as the user's name (from their login information) and the current (system) date, into the database.

## Security

As with any Internet-accessible application, security should be a primary concern in database-backed Web applications. Developers address security at three levels: system, database, and browser/server communication. While system security is beyond the scope of this article, the other two are not; both have issues associated with them. These issues center around two core concepts: authentication and authorization. Authentication lets the application know that the person attempting to access it is the person she or he claims to be. Authorization is of verifying that that person is allowed to perform the requested action.

While one probably won't authenticate end users going to the application's public pages, one should do so for those maintaining the database. This authentication can be done in one of three ways:

1. Securing the data maintenance directory using the Web server's security mechanism
2. Using the RDMBS's security tables
3. Using tables within the application database for authorization

The advantage of the first, and simplest, option is that it allows the developer to utilize an already-developed security system. Since the Web server authenticates the user, the developer doesn't need to build that type of functionality into the application. However, it has the disadvantage of not supporting authorization. Once users are authenticated, they all have the same rights within the application. If that is not desirable, another approach will be needed.

Most RDBMS products provide support for user authorization in that their security tables can associate different permissions (such as Select, Update, and Delete) within the RDBMS with different users at the database, table, and even field level. In other words, one may allow certain users to add records to one set of fields while allowing them only searching capability with another set. The drawback of using the RDBMS security tables is that maintaining these tables can be awkward and could lead to security problems if one is dealing with numerous users.

The third option is set up a table that will contain the user's username, password, and access levels(s)— admin, student, supervisor, and so on. A user who starts to work with the application is then presented with a login screen on which to enter this information. The application would then look the information up in the table for authentication. If a match is found, the user is provided with access to the application and the access levels are noted. Then, each time an interaction with the database is attempted, the application compares the user's access levels with those required to perform the interaction. If there is a match, the user is allowed. If there is not a match, the user is disallowed.

For this third option to work, the user's information needs to be available to every part of the application. Since the Web is stateless (each transaction with a Web server is a discrete event with the server usually keeping no information between transactions), some means of passing the user's information between the pages is needed. This can be done in the following ways:

- Put the information into hidden tags that can be passed from page to page and use the HTTP_REFERER environment variable to make sure that infor-

mation passed is coming from a trusted source. The problem is that it can be very tedious to make sure that each page receives and passes the information on correctly.

- Once the user is authenticated, set a cookie that contains the user information and pass it from page to page. A user attempting to access a page without the proper cookie is taken directly to the authentication form.
- Some development tools (such as php) support the concept of sessions and, within those sessions, the creation of session variables. Once these session variables are set, they can be accessed without having to explicitly pass them between pages (and, if they are not available, the user is forced to authenticate). The one caveat here is to make sure that only the application (and not malicious users) can set these session variables.

The third security area concerns the communication between the user's Web browser and the Web server. To protect the database from unauthorized access, one should explore implementing a secure sockets layer (SSL) on the Web server.[23] SSL causes the browser to send authentication information to the server in an encrypted form. If this is not done, hackers with freely available software could grab this information and use it to make unauthorized changes to the database.

## Communication

Finally, one must develop and maintain good lines of communication between both the library administration and staff members involved in the process. This communication is critical to ensuring (and maintaining) the success of the project. Given that there can be delays, it is important that the development team provide regular and frequent updates to both groups on how the project is going and inform them of any problems or delays that arise. That kind of feedback helps to maintain interest and assists in keeping support for the project from evaporating.

## Conclusions

Using a database can provide a multitude of benefits that make it well worth the investment of time, energy, and resources. It allows for easier Web-site maintenance, more efficient use of time, and a greater ability to keep data current. However, the output of the system will only be as good as the time, planning, and care that go into the system's creation. By taking the time to do things right rather than quickly, developers ensure that the investments made will be richly repaid many times over. Anticipating challenges and making the right decisions allows one to avoid many problems. Thus, by utilizing the expertise, vision, and energies of the library staff, the development team can create a product that will be useful for years to come and reflect the qualities of the staff who made it happen.

## References and Notes


1. Due to the predominance of various flavors of Unix and Windows (as well as the author's expertise), this article will concentrate primarily on those platforms. Note that, in addition to using the Unix tools contained in this article, Macintosh users can use native Mac OS applications to create database-backed Web pages. Accessed Apr. 12, 2002, www.nisto.com/mac/tool/database/html.

2. Philip Greenspun, *Philip and Alex's Guide to Web Publishing* (San Francisco: Morgan Kaufmann Pub., 1999.) (Also available on the Web at: www.arsdigita.com/books/panda); Luke Welling and Laura Thompson, *PHP and MySQL Web Development.* (Indianapolis: Sams, 2001); Jim Buyens, *Web Database Development Step by Step* (Redmond, Wash.: Microsoft Pr., 2001); Micah Brown and Mike Frederick, *Essential Cold Fusion 4.5 for Web Professionals* (Paramus, N.J.: Prentice Hall, 2001).

3. Kristin Antelman, "Getting Out of the HTML Business: The Database-Driven Web-Site Solution," *Information Technology and Libraries* 18, no. 4 (Dec., 1999): 176–81; William H. Mischo and Mary C. Schlembach, "Web-Based Access to Locally Developed Databases," *Library Computing* 18, no. 1 (1999): 51–58; and Carole A. Kiehl and Edward H. Summers, "Comprehensive Access to Periodicals: A Database Solution," *Library Collections, Acquisitions, and Technical Services* 24 (spring 2000): 33–44; and Stephen Westman, "Database-backed Library Web Pages,"*The Electronic Library* 19, no. 6 (Dec. 2001), 424–31 Antelman's article provides a good overview of database-backed Web pages.

4. While this group may include data users, it should also include others who have never seen the system before.

5. Greenspun, p. 313. Greenspun's use of the term "transaction" is potentially confusing, due to the fact that the word has a particular meaning in RDBMS systems. Therefore, this article will use the term "interactions" when referring to this particular concept.

6. Michael J. Hernandez, *Database Design for Mere Mortals: A Hands-on Guide to Relational Database Design* (Reading, Mass.: Addison-Wesley, 1997).

7. The term SQL is the acronym for "structured query language," which was first described in 1970 by E. F. Codd. The relational database model has quickly become industry standard for databases ranging from Microsoft Access for desktop applications to products such as Oracle which manage the data for huge multinational corporations. Codd's article is available in Michael Stonebreaker's *Readings in Database Systems* (San Mateo, Calif.: Morgan-Kaufmann Pubs., 1988). A good introduction to SQL by Philip Greenspun can be found at www.arsdigita.com/books/sql, accessed Apr. 12, 2002.

8. DBI (or Database Interfaces) are perl modules that provide a perl-based interface to a wide variety of RDBMS platforms. Accessed Apr. 12, 2002, http://dbi.perl.org/index.html contains a list of modules. Accessed Apr. 12, 2002, www.perl.com/pub/1999/10/DBI.html has more information available; open

database connectivity (ODBC) is a standard protocol for accessing different databases residing on different servers and platforms; JDBC (or Java Database Connectivity) is similar to ODBC and is an API developed by Sun Microsystems for use with their Java language. It provides Java programmers a means for integrating database queries to databases residing on the same or other servers into their programs. Accessed Apr. 12, 2002, http://java.sun.com/products/jdbc provides more information.

**9.** A notable exception being OpenBSD which claims "Three years without a remote hole in the default install." Accessed Apr. 12, 2002, www.openbsd.org.

**10.** Applications Programming Interface (API)—this is the set of hooks the RDBMS vendor provides to allow interactions to occur between the RDBMS and outside programs.

**11.** These categories are borrowed from Roy Tennant, "A Database for Every Need," *Library Journal* 124, no. 20 (Dec. 1999): 41,43.

**12.** More information on and downloads of these products are available: Accessed Apr. 12, 2002, FrontBase is available from: www.frontbase.com; accessed Apr. 12, 2002, mSQL at www.hughes.com.au; accessed Apr. 12, 2002, MySQL at: www.mysql.com; accessed Apr. 12, 2002, PostgreSQL at www.postgresql.org; accessed Apr. 12, 2002, Filemaker Pro at www.filemaker.com; accessed Apr. 12, 2002, Fourth Dimension at www.acius.com; accessed Apr. 12, 2002, Lasso from www.blueworld.com; accessed Apr. 12, 2002, MS-Access at www.Microsoft.com/office/access/default.htm; accessed Apr. 12, 2002, SQL Server at www.Microsoft.com/sql/default.htm); accessed Apr. 12, 2002, DB2 at www-4.ibm.com/software/data; accessed Apr. 12, 2002, Informix at www-4.ibm.com/software/data; accessed Apr. 12, 2002, Oracle at www.oracle.com; and accessed Apr. 12, 2002, Sybase at www.sybase.com; PostgreSQL, on Windows systems require the use of Cygwin. accessed Apr. 12, 2002, http://sources.redhat.com/cygwin.

**13.** One source for side-by-side feature comparisons, accessed Apr. 12, 2002, is www.mysql.com/information/crash-me.php. John Ashenfelter, *Choosing a Database for Your Web Site* (New York: Wiley, 1998) also discusses a variety of Web tools, focusing on Access and Filemaker Pro.

**14.** A good overview Oracle database administration is provided by Rick Greenwald, Robert Stackowiak, and Jonathan Stern, *Oracle Essentials: Oracle 8 and Oracle 8i* (Sebastopol, Calif.: O'Reilly and Assoc., 1999).

**15.** Perl is a very popular scripting language that is used extensively in Web programming, particularly CGI applications. Accessed Apr. 12, 2002, www.perl.org for more information; accessed Apr. 12, 2002, http://msdn.Microsoft.com/vbasic will provide information on Microsoft's VisualBasic.

**16.** Mod_php is an Apache server module that uses php, an open source scripting language that can be used on multiple platforms and can be integrated into an Apache server. It allows HTML page creators to embed commands that can search a database and then output the results directly into their pages. Accessed Apr. 12, 2002, www.php.net for more information on php. Accessed Apr. 12, 2002, http://hotwired.lycos.com/Webmonkey/databases/tutorials/tutorial4.html for a tutorial on how to include php searches in an HTML page. Note that php4 allows the same type of functionality to be included using Netscape and Microsoft servers using NSAPI and ISAPI respectively. Active Server Pages (ASP) is a Microsoft scripting language that works with their Internet Information Server. Accessed Apr. 12, 2002, http://msdn.microsoft.com/scripting/default.htm?/scripting/vbscript/default.htm for information. ASP, while originally a Microsoft technology using VisualBasic/VBScript, has been extended for use on other platforms, such as the Apache::ASP port, which is multiplatform. Information on the Microsoft product is available at http://msdn.microsoft.com/library/default.asp?url=/nhp/Default.asp?contented=28000522 and at www.nodeworks.com/asp/for the Apache product, accessed Apr. 12, 2002,

**17.** Apache requires the use of mod_perl for Apache::ASP. ColdFusion is a Web application server from Allaire that incorporates visual programming, as well as database and debugging tools, available in versions for Solaris, NT, and Linux. It uses its own proprietary language for integrating database support into HTML pages. Accessed Apr. 12, 2002, www.allaire.com/products/coldfusion for information. A free, limited-functionality version (for Linux and NT), ColdFusion Express, is available. CFML (ColdFusion Markup Language) available for Linux and Solaris only.

**18.** GDIdb is a Windows-based automated application that provides scripting that permits the creation of HTML files from any ODBC-compliant database. The company offers two versions, GDIdb and GDIdbPro, the latter stating that it can interactively receive and publish information in real time. Accessed Apr. 12, 2002, www.gdidb.com for information. See also Ernest Perez, "GDIdb: Getting Your Databases on the Web, with No Big Deal," *Library Computing* 18, no.1 (1999): 29–35. For other approaches, see www.tucows.com/htmlacc95.html, accessed Apr. 12, 2002.

**19.** While any good programming book should contain information on commenting code, a particularly good example can be found in Tobias Ratschiller and Till Gerken, *Web Application Development with PHP4.0* (Indianapolis: New Rider's Pubs., 2000).

**20.** Greenspun, 358.

**21.** There are a number of these types of products. In addition to those that come with the enterprise database servers, manufacturers of commercial backup software include packages that support major RDBMS vendors. These include Veritas Software's NetBackup, accessed Apr., 12, 2002, www.veritas.com; Legato, accessed Apr., 12, 2002, www.legato.com; and Sun's Solstice Backup accessed Apr., 12, 2002, www.sun.com/solstice/backup that provide this support.

**22.** Note that this is an important difference between lower- and higher-priced systems. If the system utilizes table-level locking, this means that nobody else can modify any record within an affected table until the process is completed, whereas row-level locking only affects the rows involved in the process. While this may not be a problem for many library programs, it could pose a problem for applications where there could be a great deal of data maintenance occurring concurrently. See Greenspun, 352–53 for caveats on this approach.

**23.** Secured Sockets Layer—a means by which transactions between users and Web servers can be made secure via certificates, and use encryption between the user and the server to ensure that transaction data are not visible to snoopers on the Web. Accessed Apr., 12, 2002, http://searchSecurity.techtarget.com/SDefinition/0,,sid14_gci343029,00.html for more information. Most commercial Web servers provide this capability. Users of Apache can have two alternatives, both accessed Apr., 12, 2002, www.modssl.org and www.apache-ssl.org.

# Multimedia Research Support for Visiting Scholars in Museums, Libraries, and Universities

*Marcia J. Bates, Catharine Hulsy, and Geoffrey Jost*

*Results are reported from a study conducted to determine the state of the art in the provision of multimedia support by cultural institutions to visiting scholars. In the study, conducted under the auspices of the Getty Research Institute (GRI), Los Angeles, California, experts were interviewed and cultural institutions were surveyed. The following was determined: In most museums and libraries, multimedia support and advanced computer support for scholars is still minimal to nonexistent, while support for scholars in universities is often more substantial. Responsibility for management of new media support is currently diffused among a number of different staff positions in these institutions.*

This article reports results of a study conducted to determine the state of the art in the provision of multimedia support by cultural institutions to visiting scholars. As multimedia and advanced computer system use becomes more common, cultural institutions of all types can expect visiting scholars to want to engage in any number of new-media-intensive activities, such as setting up their own multimedia Web site as a cultural or artistic contribution; doing advanced online research and downloading; entering research data into databases to be made Internet accessible; or any of numerous other uses of recently developed information technologies.

The results are drawn from a study funded by the GRI. It was initiated after staff began receiving requests from prospective visiting scholars for various forms of multimedia support during their sabbatical year at the GRI. These requests raised questions about how the GRI could anticipate likely future requests, and what a fair level of support would be. Though the study was originally done for the GRI's own purposes, the results reported here are those that are likely to be of interest to other cultural institutions or to researchers studying multimedia and information technology usage by scholars.

In the following, the methods used are reviewed, then the results presented. The paper concludes with a summary and recommendations section.

## Methods

Four different methods were used to identify the information needed to answer the GRI's questions: literature and Web-site review, interviews, survey, and Web-site analysis. The literature review, developed by Bates, is available at www.gseis.ucla.edu/faculty/bates/scholars.html and the Web-site review, prepared by Hulsy, is available at www.gseis.ucla.edu/faculty/bates/multimedia.html. The reviews, prepared in 1999, are very extensive, containing a discussion of scholarly use of information technology, a review of technology/arts-related Web sites, organizations, and other relevant material.

A large number of resources were discovered to be relevant to scholarly use of technology, especially in the arts, but no literature was found that directly addressed the specific questions studied here, namely the needed support for visiting scholars at cultural institutions.

The interviews were targeted at creative and representative leaders associated with various aspects of the survey questions who would perhaps have the insight and foresight about these questions that a survey would not reveal. The survey was intended to provide a rough indication of the range and frequency of types of multimedia support. Finally, the Web-site analyses were done to determine what kind of support had been needed to create various novel and leading edge cultural Web sites. All data were collected during spring and summer of 1998. The data thus constitute a snapshot of the state of the art at that time.

## Interviews

Eleven people were interviewed in sessions lasting up to one-and-a-half hours. Interviewees were selected for their expertise in such areas as academic support for faculty in universities, library management, museum management, digital archives, multimedia design, art history, and the practice of art. See appendix A for a list of the interviewees.

## Survey

While getting a general sense of the popularity of various types of support for scholars was desirable, it was decided at the outset that the more important objective was to uncover the full variety of institutional responses to the new media, at least among major institutions. The sampling, therefore, was broad rather than deep, including museums, universities, and libraries, and drew strongly on cohorts of institutions thought likely to be engaging in new

**Marcia J. Bates** (mjbates@ucla.edu) is a Professor in the Department of Information Studies, University of California, Los Angeles; **Catharine Hulsy** is Systems Analyst, Internet Technologies Department, Pacific Health Systems, Cypress, California; and **Geoffrey Jost** is Lead Programmer, education.com/Vivendi Universal, Los Angeles, California.

approaches. The GRI also wanted extra attention given to southern California institutions in order to learn of the activities of local institutions. As a result, the emphasis in the sampling design was on selecting relatively small subsamples of a variety of institutional types—museums, libraries, and universities, both cutting-edge and the general-run, both national and local.

For the purposes of this survey, cutting-edge consisted of institutions thought likely to emphasize new media, electronic information, and the Internet. The national cutting-edge samples were drawn from the directory of the Museum Computer Network (www.mcn.edu) for museums, from a listing of New Media Centers (www.csulb.edu/gc/nmc) for universities, and from the Los Angeles Culture Net (www.lacn.org) for all three institution types in southern California. Where institutional lists were available, systematic sampling was used, wherein every "nth" listing was selected.

For the general run of museums, it was determined that an appropriate cohort would be major, premier institutions, comparable to the Getty Museum, rather than a random selection from the thousands listed overall, many of which are small and do not have the resources or interests of the larger institutions. No rankings of the top one hundred art museums or anything of the like were found; in the end, the knowledge of people in the museum world was drawn on for the list of premier museums to sample.

Premier universities were selected both for their general preeminence as well as prominence in the arts and humanities, and libraries were selected from the Association of Research Libraries' (www.arl.org) list of top institutions. In all cases, institutions were selected for the survey, but the questionnaire was directed to particular individuals in the institutions.

In the end, a sample of 101 contacted institutions was achieved, from which were gathered a total of sixty-three complete surveys. Table 1 provides the sample sizes, planned and actually achieved, for each category of institution in the study. The questionnaire is reproduced in appendix B.

The original plan was to survey as many of the institutions as possible (especially the cutting-edge ones) by e-mail, and supplement this means of contact with telephone contact in case of nonresponse or unavailability of e-mail addresses.

The actual conduct and data collection for the survey proved to be difficult. Though all e-mail contacts were made to individuals, not institutions, responsiveness to e-mail was nonetheless disappointing, and many responses had to be taken by telephone. The reasons for this difficulty possibly have much to do with the early stage of development of this new media area. Labels and roles for professional positions associated with multimedia in the various institutions have not even begun to stabilize. It was often necessary for project staff to make numerous telephone or e-mail contacts just to determine whom they should be talking with. Since two or more different individuals were responsible for various functions covered in the survey instrument, with different jobs in each institution, one person often turned out to be several.

These difficulties overlay the usual problems with survey sampling, namely nonresponse or partial response. In the case of partial responses, telephone calls or e-mail contacts helped to complete some questionnaires. In the case of nonresponse, particular efforts were made to contact at least some of the institutions in each category of the above sample so that the full range of types of respondents were included. This approach helped achieve the desired result of surveying widely to capture the variety of multimedia support activities across institutions and environment types, rather than surveying more narrowly and deeply to verify statistical frequencies.

## Web-Site Analysis

Eleven Web sites displaying innovative humanities and social science scholarly material were analyzed to determine the hardware and software that was likely behind them. In this case, no formal sampling was attempted. Sites were discovered through exploration and following up arts-related links. Though complete detection of software and systems used through Web-site analysis alone was not possible, a great deal was learned and the method has promise for certain purposes. The full set of sites analyzed is listed in appendix C.

## Results

### Interview Results

Key points drawn from the interviews:

- High-end needs of scholars and artists doing original multimedia work are so diverse that anticipation of their needs through advance purchase of sophisticated hardware and software is not cost-effective.
- All scholars should be supported by a baseline hardware and software configuration consisting of powerful but not extremely high-end Macintosh and PC computers (both platforms are popular), coupled with a standard office suite, communications software, and a popular general-purpose image processing program such as Adobe Photoshop.
- Human technical support and assistance is always needed, whatever the level of hardware and software support. Problems always arise with respect to configuring hardware and software

**Table 1.** Survey Samples

| | Respondents | Nonrespondents | Total Unique Surveyed |
|---|---|---|---|
| **Cutting Edge** | | | |
| Museum Computer Network | 14 | 7 | 21 |
| New Media Centers | 10 | 8 | 18 |
| Totals | 24 | 15 | 39 |
| **Premier comparison** | | | |
| Museums | 11 | 3 | 14 |
| Universities | 3 | 12 | 15 |
| Libraries/Archives | 5 | 3 | 8 |
| Totals | 19 | 18 | 37 |
| **Southern California** | | | |
| **Cutting Edge—LA Culture Net** | | | |
| Museums | 3 | 2 | 5 |
| Universities | 3 | 2 | 5 |
| Libraries | 4 | 1 | 5 |
| Totals | 10 | 5 | 15 |
| **Premier comparison** | | | |
| Museums | 5 | 0 | 5 |
| Universities | 1 | 0 | 1 |
| Libraries | 4 | 0 | 4 |
| Totals | 10 | 0 | 10 |
| Grand Totals | 63 | 38 | 101 |
| **Crosstabulations** | | | |
| Cutting Edge Total | 34 | 20 | 54 |
| Premier Total | 29 | 18 | 47 |
| Grand Totals | 63 | 38 | 101 |
| Museums Total | 33 | 12 | 45 |
| Universities Total | 17 | 22 | 39 |
| Libraries/Archives Total | 13 | 4 | 17 |
| Grand Totals | 63 | 38 | 101 |
| General Geog. Location Total | 43 | 33 | 76 |
| Southern California Total | 20 | 5 | 25 |
| Grand Totals | 63 | 38 | 101 |

NOTE: Percentages are not provided, because with a total sample of 101, percentages and raw numbers would be almost identical.

to the local network of the visited institution, as well as in helping scholars with problems that are novel to them, no matter what level of sophistication they have upon arrival.

- Ability for scholars to do sophisticated research with their hardware/software configuration is as important as ability to do multimedia authoring. Consideration should therefore be given to provision of access to electronic databases—bibliographic, original materials (such as the Perseus database of classical literature), and image bases. Broad bandwidth permitting fast downloading of high-resolution images is important to that research.
- Museum designs that are dynamic, integrated systems will be increasingly common. Collections management systems will be expanded to be end-user–accessible search systems, and will be used both for research and collection management. Staff and scholar users will jointly contribute to the description of museum objects.
- Museum collections of real and virtual objects will supplement each other. Users will not need to handle objects as much when they can see them in virtual form, yet availability of the virtual form will acquaint many more users with the fuller museum collection for which there is no space in museum galleries.
- Many creative arts and humanities multimedia projects are also put on the Web. Management of Web-based projects of visiting scholars, including decisions about what will be mounted on the visited institution's local system, what can and cannot be migrated to the scholar's home institution, and issues of copyright and image integrity in such Web sites are important to the success of multimedia elements of any visiting scholars program.

- Collaboration and interdisciplinary work, particularly associated with e-mail and use of the Web, is increasing in humanities research. Support for communication is vital.
- Institutions should be prepared to provide for handicapped scholars, including special computer equipment for the partially sighted.
- Due to shifts in orientation of humanities research in recent years, access to secondary sources is growing in importance. Since much research involves "repositioning and recontextualizing" art objects, it is important to see what other scholars have written, in addition to studying the object itself.

## Survey Results

In general, the most advanced institutions in the development of multimedia support for scholars are the new media centers in universities. These have been developed to provide active support and encouragement to faculty to use new media and the Web for research and teaching, especially teaching. Universities without new media centers tend to offer more conventional computer support such as computer labs and troubleshooting for faculty office computers. However, despite the current low demand from scholars, some leaders in the museum and library world are also developing new models of service that incorporate the scholar in an integrated research and artifact image base environment. (See later discussion.)

- In general, differences in conventional and multimedia support were greater across the three types of institution—museums, universities, and libraries—than they were between cutting-edge and general premier institutions.
- The most commonly mentioned forms of new media and multimedia support in the three cultural institutions are digital scanning (regular and slides, either by the institution or permitted of users), CD-ROM burning and archiving, and helping users with Web sites, including mounting Web sites for users. All three types of institutions offer various forms of scanning, but universities offer the bulk of the other common forms of assistance, such as for Web sites and CD-ROM archiving. Nearly a third of the respondents—all of them museums or libraries—said they offered no multimedia support of any kind. Two premier museums mentioned that they prohibited scanners.
- For end-user service, both Macintosh and PC platforms are popular, though by no means universally available, and were available mostly at museums and universities. Standard PC's and Macs were about equally common at universities, but PC's dominated at museums. High-end Mac and PC machines are fairly common, but are mostly to be found at universities. Likewise, universities were the home of the handful of high-end machines of other brands that were mentioned. A few museums and libraries provide electrical outlets for computers or permit laptops.
- Web access is sometimes made available at all three types of institution; Netscape is the usual Web communication software employed. Telnet is occasionally made available for remote e-mail at all three types of institution as well.
- The total availability of software for scholar use is not high, and most of what is provided is to be found in universities. One form or another of word processing software or office suites was the most common type made available. Among other types of software mentioned, such as multimedia and desktop publishing, the most frequently mentioned packages were Photoshop, Premier, PageMaker, Illustrator, Director, and QuarkXpress. Of these, only the first was available at more than ten institutions.
- Technical assistance and training took on a variety of configurations, from nonexistent, to passive support through the availability of help-lines or consultants on call, to proactive encouragement and assistance in creation of multimedia and Web products. The last configuration was found only in university new media centers.
- Though not specifically asked, seven museums mentioned plans to develop and make possible public access to museum collections management systems. This raises many issues of usability, as most such systems are not currently designed for unskilled use.
- Museums and libraries (rather than universities) were most likely to provide conventional support, such as photocopying, microform readers, and study space. A minority of the museums and a majority of the libraries provide online catalogs as well. A few of both institutions also provide online databases and CD-ROM databases.
- Conventional media technical support, such as the provision of photographic services and color copying, are widely available, especially at museums and libraries.
- Institutions are generally aware of the thorniness of questions regarding copyright of digitized materials, and many are thinking about it but do not yet have a policy. In the survey findings, the University of Texas at Austin has by far the most developed response on copyright of this sort, including specific attention to multimedia copyright issues

(available on the Web at www.utsystem.edu/ogc/intellectualproperty/distance.htm).

## Multimedia Job Responsibilities of Institutional Staff

Responsibility for multimedia/new media resources and activities is spread across a wide range of positions in the cultural institutions surveyed, while standardized job titles have not yet emerged. To illustrate the current flux, figure 1 presents selected job titles for the respondents surveyed in this project. The job titles are grouped by cultural institution type. Duplicative or very similar job titles are not listed.[1]

**Museums**

Librarian

Assistant Director of Research

Director of Research and Collections

Manager, Imaging and Technology Development

Manager of Information Services

Executive Director of Imaging

Registrar

Curator

Manager, Photographic Services

Director of Systems

Associate Director of Publications

Research Associate

Director, Print and Electronic Media

Director of Information Resources

MIS Manager

Assistant Curator for Research

Head of Collections

**Libraries**

Head Librarian

Director, Library and Information Resources

Archivist

Reference Archivist

**Universities**

Computer Manager

Lead Consultant, Academic Technologies Group

Manager, Academic Computing and Software Development

Humanities Information Technology Specialist

Assistant Director for Consulting Services

Acting Manager, Web, Instructional and Media Services

Faculty Room Consultant, Instructional Technology Services

Academic Technology Specialist for Art

Manager, School of Humanities Computing Facility

Director, New Media Center

Director, ad interim, Center for Instructional Technologies

**Figure 1.** Selected Job Titles

### Institutional Configurations

One of the aims of the survey was to see if new organizational and service configurations around multimedia are developing in the surveyed institutions. Eight distinctive institutional arrangements found in the survey are described below. Other configurations that are more conventional, or are mixes of those below, are not reviewed.

**Museum: Moderate to High Services**—California Museum of Photography, University of California at Riverside. Conventional support: Copystand for still photography and video can be arranged with advance notice. Media technical support: wide range of photographic services. New media technical support: scanning station is available. Resolution of scanned images monitored: Web designers asked to use low-resolution images; for users with long-term or specialized relationship with the museum, higher resolution images may be available. Computer technical support: Mac Quadra with scanner and zip drive attached. People to help users: museum staff available to support and troubleshoot. Software available for use by scholars and artist users: Adobe Photoshop, scanning software.

**Museum: Balance of Real and Web-based Resources**—Fine Arts Museums of San Francisco. Of all the museums surveyed, this one had by far the largest online image collection, available via the Web. Users may browse and download images remotely from a seventy-thousand-item collection (www.thinker.org/fam/thinker.html), many of which are not on public display. The museum's offerings to users constitute a truly mixed collection of the physical and the virtual.

"The Fine Arts Museums of San Francisco is a public museum with an evolving mission to behave more like a resource and less like a repository" (from Web site). According to the survey respondent, overall use of the collection is up, while direct handling of objects is down. The image base is also serving the internal needs of the curatorial staff who now can make selections online, print out copies from the image base, and assemble an exhibition list before having to consult the actual object for the final decision. The image base is also making it possible to link (and integrate) various museum functions together. Most staff have direct access to the records and can make changes or add information to the descriptions. In this way, it is functioning as a true knowledge center as opposed

to having pockets of knowledge scattered around the institution.

**Museum: Two Institutions Cooperating**—Minneapolis Institute of the Arts and Walker Art Center. ArtsConnectEd (www.artsconnected.org) and the Integrated Arts Information Access Project are complementary, interdependent collaborations between the Minneapolis Institute of Arts and the Walker Art Center that provide online access to the rich collections and reference, archive, media, and curriculum resources of both institutions through one shared point of entry or gateway (Web site). The two museums are collaborating in depth to create a single point of access to their multiple resources.

Teachers may use the site to search for and access art-related curriculum materials and student activities and may participate in online discussions between school and museum educators. Students can use it to conduct project-related research, and they and the general public can explore the collections and resources of both museums through online tours and browsing. The Minneapolis Institute of the Arts is currently in the process of integrating the library catalog with the collection management system so that bibliographic and collection or object records can be accessed through one integrated system.

**Museum: Special Center for Electronic Information Access**—Metropolitan Museum of Art, New York. The Met is very large and organizationally decentralized, with more than eighteen different curatorial departments. Fellows may have access to an office with a computer and a fax machine, but availability of such resources is entirely department-dependent. There is no museumwide policy regarding computing arrangements for visiting fellows. There are several different libraries as well.

In this decentralized environment, the new Hazen Center for Electronic Information Resources has been developed to provide access in a single place to an extensive range of electronic information resources. The Center provides training and support in the use of an expanding collection of CD-ROMs, online journals, and Internet resources. The Web-based resources collected by the center are accessible through the main library's Web page. Future plans include selective access to the museum's text and image-based collections management system, currently being developed to provide information on the museum's more than two million works of art.

**Library: Large, Classic, Full-Service**—Huntington Library, San Marino, California. Conventional support: staff operate photocopier for users. Media technical support: wide range of photographic services, including microfilm, infrared, Beta radiograph, and ultra violet; all require approval of curator. New media technical support: scanning, image capture, CD-ROM burning available. Users may not digitize copyrighted material. Computer support: only online catalog and CD-ROM terminals available. People to help: no technical help for personal computers brought to library by users. Library staff will provide appropriate reference services, including mediated electronic searches of bibliographic databases. Software support: library will print Word Perfect or Word documents from disks supplied by readers at their request for a nominal price per page.

**Library: Mix of New and Old Services**—Architecture and Fine Arts Library/Arts Center, University of Southern California, Los Angeles. This library offers general library resources, a slide library, and special collections. The arts and architecture departments of the university offer Web authoring assistance to faculty, but the library serves a variety of other newer types of needs. The library has a slide photographer and offers scanning (with staff assistance) for both slides and other materials. There are graduate students to assist with higher-level computing needs; undergraduate students assist with Web authoring and scanning. The library has Adobe Photoshop and PageMill, as well as scanning software.

**Note on university support for scholars**—Several patterns in university information technology support were detected. Universities varied in their placement of specialized personnel and services. Some had computing centers for the whole of the humanities, while other centers were focused on specific departments, such as art or music. There is also frequently a division of services offered to faculty and to students, though this division is less likely to be present in those cases where the technical support is centered within a single department. Some universities employed subject-specialist consultants within a more generalized academic or faculty-computing center.

In those cases where a formally organized new media center existed to support faculty (usually to provide campuswide support), there was generally a more concentrated, extensive, and consistent array of services offered to faculty. Where no formal new media center existed, services were more scattered and uneven.

**University: Proactive New Media Center**—Claremont Graduate University Humanities Center Technology Lab, Claremont, California (http://newmedia.cgu.edu/lab.htm). "[T]his facility offers humanities faculty, staff, and students a serene and comfortable computing environment" (from the Web site). Further, "the staff is here to work with the humanities faculty across the Claremont Colleges." On a Web page titled Faculty Support, What We Can Do For You, the following forms of support are listed: "multimedia project planning; layout and design techniques; digitizing photographs, graphics, sound and video; HTML programming; testing of new-media course materials; classroom facilitation of your project."

There is also the Humanities Electronic Media Project Professional Development Program, which is

> designed to equip faculty, staff, and students with the knowledge and skills necessary to incorporate new technologies into their own learning and professional environment. Through an array of innovative seminars and workshops, join us and learn to develop materials that integrate multimedia on-line resources. (Web site)

**University: Advanced Media Center for Arts Training**—City University of New York Brooklyn College Digital Media Center for Arts. This center, part of the Art Department at the University, is intended to train artists to produce the most technically advanced digital materials. Starting with desktop publishing, the center moved on to multimedia production, then to nonlinear animation and video production, and is now focusing on digital video disk.

## Web-Site Analysis Results

More and more scholars are engaged in creating Web sites containing a variety of different types of information and underlying technology. Researchers might want to bring such sites with them, or create them at the institution they visit. An analysis of current innovative humanities Web sites was undertaken in order to determine the types of demands that might be placed on cultural institutions by visiting researchers. The results are as follows.

Of the eleven sites that were examined, all:

- employed some type of HTML editor;
- utilized a scanner and accompanying software; and
- contained images generated or modified by some sort of image-editing software

CGI scripting was used in several sites, indicating that the producers either had direct access to the Web server or were able to have the server administrator perform set-up and maintenance. One of the sites used Microsoft FrontPage CGI modules, which require special server plug-ins, but is otherwise unexceptional. One site used simple audio clips. And finally, FileMaker Pro database software was used in the production of at least one site, though not for hosting.

Most of the sites require hardware available on current mid-range computer systems, or feature flashy and complicated commercial graphic design. The sites are built with standard tools on standard computer systems. Unless it is stated explicitly on a site, it is difficult to tell exactly what software programs were used. This is because most text and images can be copied and pasted between programs, and many designers use a combination of programs to achieve their effects.

## Summary and Recommendations

In this study of multimedia support afforded to scholars in museums, libraries, and universities, it was determined that such support is still minimal to nonexistent in most museums and libraries. Demands placed by scholars on these institutions for new media support remain at a low level. A number of cultural institutions are experimenting with new organizational and technological arrangements; however, responsibility for management of new media support is currently diffused among a number of different positions in these institutions. Copyright issues are yet to be worked out, and practices for handling fragile materials in the multimedia environment (with scanning machines, for example) have not been formalized.

The most active support for scholars is provided in university new media centers that are dedicated to assisting scholars and students in use of new media and the Internet, especially, but not exclusively, for teaching purposes.

Museums provide extensive but more conventional support, such as microfilm and fiche readers, photographic services, and dedicated space. Libraries and museums often provide access to online catalogs; less frequently, museums provide access to collections management systems and other online databases.

Drawing on the results of all elements of this study—interviews, survey, and Web-site analysis—it was concluded that most technically advanced needs of potential visiting scholars can be met with a generic and relatively inexpensive package of resources:

- Provide both powerful PC and Macintosh computers, so scholars may choose. Top-of-the-line computers are not needed, however.
- Provide each machine with a standard office suite, such as Microsoft Office; communications software for the Internet and the Web, such as Telnet and Netscape; and a popular multimedia image processing program such as Adobe Photoshop.

Given the current state of usage of sophisticated computer capabilities, higher-level support should not be routinely provided, but rather negotiated on an individual basis on those occasions when needed.

It is also suggested that institutions consider providing each scholar his or her own laptop computer to use while visiting. The prospect of having one's own laptop may forestall requests to bring computers from the home institution, thereby reducing complications of setting up "alien" computers on internal museum or library networks. Such an arrangement would also enable scholars to work both at the cultural institution and in their apartments during their visit.

In addition to facilitating multimedia authoring, support for scholars should promote the ability to do

online research and to communicate electronically with other scholars and online groups.

Finally, the enormously varied and nonstandardized management of new media in the studied institutions makes survey sampling, such as was done for this study, extremely labor-intensive and not a cost-effective means for the annual assessment of the status of multimedia support at other institutions. Instead, it is recommended that institutions use a comparison sample of a handful of other similar institutions as the basis for any occasional assessment of the state of the art in the provision of multimedia support for visiting scholars.

## Note

1. In museums, the person responsible for scholars' support was sometimes the librarian. In those cases the librarian was surveyed, even though the sample member drawn was a museum.

## Appendix A. Persons Interviewed

Following is the final list of eleven people interviewed in depth. Annotations explain the reasons for their choice as respondents:

Howard Besser, Associate Professor, Dept. of Information Studies, UCLA; image databases and museum networking specialist

Robert Flick, artist, former Visiting Artist at GRI

Bernard Frischer, Professor, Classics, UCLA; creator of an innovative new media resource

Anne Gilliland-Swetland, Assistant Professor, Dept. of Information Studies, UCLA; digital archive and multimedia production specialist

Ted Kahn, DesignWorlds for Learning; multimedia and educational technology specialist

Clifford Lynch, Director, Coalition for Networked Information of the Association of Research Libraries

Paul Marty, Librarian, Spurlock Museum, University of Illinois, site of an innovative approach to museum management online

Wayne Miller, Programmer/Analyst, Humanities Computing Facility, and Assistant Adjunct Professor, German, UCLA

Donald Preziosi, Professor, Art History, and Head, Program in Museum Studies, UCLA

Steven Rossen, Supervisor, Faculty Media Center, Office of Instructional Development, UCLA

Victoria Steele, Head of Special Collections, University of Southern California Library, and art historian

## Appendix B. Survey Text

You may be able to help us with a survey commissioned by the Getty Research Institute for the History of Art and the Humanities in Los Angeles, California. The purpose of the survey is to discover the sorts of resources and services that are provided for scholars and artists who are working with new media and electronic information in arts-related institutions. The survey project is headed by Dr. Marcia Bates. My name is _____ and I am an interviewer for the project.

Please fill out the following questionnaire. If you have questions or prefer to talk with me by phone rather than filling out the questionnaire, please e-mail me at _____ .

Please note that you will be asked for types of information that are generally publicly known in and among organizations, so we will not attempt to maintain your anonymity or the confidentiality of the information you provide in the project report.

Please provide answers to the following questions:

List the facilities and conventional tools available for qualified scholars or artist users of your institution:
(For example: copier, micro reader)

Media technical support:
(For example: make slides, color photocopying, desktop publishing)

New media technical support:
(For example: scan and digitize, mount on Web site . . .)

Computer technical support:
Types of computers and workstations available for users to use:
(For example: two Mac Centris 610 computers)

People to help users—in what ways, how much?
(For example: technical person available to troubleshoot, but no formal classes in system use)

Software available for use by scholars and artist users:

Do you have a policy regarding copyright rights in association with the use of new media?
(For example: can users digitize copyrighted material?)

If YES, please send or fax to us a copy of the policy (Fax #:_____)

If NO, but you have standard practices, what are those practices?

Do you have any policies regarding the handling of library/museum resources by users to prevent damage or wear and tear to those resources?

If YES, please send or fax to us a copy of the policy.

If NO, but you have standard practices, what are those practices?

Thank you very much for taking the time to complete the survey. Your responses are most appreciated.

## Appendix C. Web Sites Sampled for Web-Site Analysis

The Interactive Ç/Zifar (this site no longer active)

www.acad.cua.edu/as/ml/barletta/cifar

Proto-Cuneiform Texts From Archaic Babylonia
http://early-cuneiform.humnet.ucla.edu/index.html

El Camino De Santiago
www.humnet.ucla.edu/santiago/iagohome.html

Duke Papyrus Archive
http://scriptorium.lib.duke.edu/papyrus

Isfahan—Islamic Architecture
http://isfahan.anglia.ac.uk:8200

The Complete Writings and Pictures of Dante Gabriel Rossetti: A Hypermedia Research Archive
http://jefferson.village.virginia.edu/rossetti/rossetti.html

The Labyrinth: Resources for Medieval Studies
www.georgetown.edu/labyrinth

Art Crimes: The Writing on the Wall
www.graffiti.org

Gargoyles Then and Now
http://ils.unc.edu/garg/garghp4.html

Amiens Cathedral
www.arch.columbia.edu/DDL/projects/amiens/index.html

Ars Electronica Center
http://web.aec.at

## Bel Jour: A Discipline-Specific Portal to Periodicals

*Ryan Womack*

*Convenient, subject-based access to current periodicals can be difficult to achieve in today's mixed and changing journals environment. This paper describes the creation of a Web-accessible database of journals in business and economics, using Microsoft Access and ColdFusion. The Business and Economics Locator for Journals at Rutgers (Bel Jour), provides value-added descriptive information about a specific subject-based journal collection along with convenient access to content from this collection. Although the actual journal collection is dispersed among many physical and online locations, Bel Jour provides a single, virtual point of access for researchers in the discipline.*

Printed lists of journals in a particular subject area or academic discipline were once a common sight in libraries, filed along with pathfinders and other handouts that have now migrated to the Web. These lists were quite popular with patrons who wanted a convenient way to survey resources in their field. Although libraries now provide lists of e-journals alongside subject-specific indexes and databases, few of these services give a complete picture of the library's holdings in the current environment of mixed online and print resources. Furthermore, it is difficult to extract a complete list of journals in a particular discipline from the online catalog, given the complexities of the subject headings assigned to journals. Providing institution-specific information about the full spectrum of journals available in print, online from the publisher, or available through an electronic aggregator is difficult. The Business and Economics Locator for Journals at Rutgers (Bel Jour) is a project designed to overcome these difficulties and provide a useful tool for researchers in the fields of business and economics.[1] It attempts to provide a complete listing of the journal holdings available to Rutgers patrons within its subject area, along with other descriptive data that add value beyond the online catalog record. Bel Jour is a virtual space that defines a coherent collection of sources that are physically and electronically dispersed among many locations.

Many libraries have integrated their e-journal holdings into their online catalogs, providing a necessary point of entry to these resources. However, catalogs have some disadvantages as quick-finding aids for journals. The catalog serves as a detailed record of all of the library's holdings. A search for an individual journal title frequently does not point to a single record, but to multiple records referring to variant titles, alternative formats, and differing holdings among branch libraries. While this level of detail is necessary for definitively describing the library's collection, it is often superfluous for someone who is seeking the most convenient way to retrieve a known item from a 1999 issue of the *Review of Economic Studies*. Bel Jour abbreviates the catalog record so that only the most frequently used information is reported.

Bound by the MARC format, flexible as it is, the catalog lacks detail in some areas that are useful to researchers. When links to e-journals exist, dates of online holdings are often lacking. Without clicking through several links, it is difficult to determine whether a particular issue of a journal is available online or not. Also, descriptive information about the journal is based on historical cataloging. Most notably, the publisher of a journal may have changed several times since the catalog record for the journal was created. Information about where the journal is indexed and descriptions of its editorial practices are available elsewhere on the Web, but not within the catalog record. Bel Jour adds this information to the journal descriptions in order to increase its utility to researchers.

Libraries have often produced lengthy lists of e-journals to create a more convenient way to access a known title. However, these lists are usually separate from other finding aids, and they quickly become difficult to browse as libraries add more and more electronic holdings. Some

**Ryan Womack** (rwomack@rci.rutgers.edu) is Business and Economics Librarian, Rutgers University Libraries, New Brunswick, New Jersey.

libraries have created subject-specific finding aids that include e-journal listings, but these lists are static.[2] Bel Jour is an interactive, searchable database of journals that offers a flexible presentation of its contents.

While some libraries have used database-driven Web sites to provide access to their journal collections, these efforts do not have the subject focus or enhanced information for the end user that motivates the Bel Jour project.[3]

Indexes and full-text databases are another place that researchers may turn to in search of journals. Indexes such as EconLit are comprehensive within a discipline, but lack convenience since they are not tied to a particular institution's holdings. Full-text databases such as ABI/INFORM may provide many full-text journals, but they fail to indicate journals available online from other sources, or available in print at the library. Bel Jour is designed to answer the question, "Do we have access to this title?" as quickly as possible, and to provide the answer regardless of the source's provenance.

The future promises a fully integrated online environment where journal articles will seamlessly link to one another through DOIs, OpenURLs, and the like.[4] At present, however, a collection-specific journal portal such as Bel Jour can add value by making searches more convenient and bridging some of the gaps between print and online journal collections. Since Bel Jour is built on a database and a flexible Web interface, it can be expanded and customized to meet changing user needs.

## Description

Bel Jour is built from a Microsoft Access database that contains descriptive information on nearly one thousand journals in business and economics. Fields include title, publisher, publisher URL, multiple URLs for online access to the journal, a URL for descriptive information about the journal, locations of print holdings, dates of print and online coverage, flags to indicate where the journal is indexed, the number of citations in Web of Science over the previous five years, the Web of Science abbreviation for the journal, and journal pricing information. There is also a Notes field to indicate any special access instructions or content restrictions.

Web access to the database is provided by ColdFusion software, which allows for the dynamic presentation of search results. An initial search screen allows searching by title, publisher, and keyword, along with the ability to restrict results according to whether or not the journal is indexed in EconLit or ABI/INFORM (figure 1). The search results are displayed in a brief format, which includes journal title, dates of print and online holdings, and URLs that link to the online locations of the journal (figure 2). Clicking on the title of a journal on the brief-results screen brings up the detailed-results screen, where most of the fields listed above can be viewed (figure 3). The format is deliberately simple in order to allow quick access to online journals with a minimum of mouse clicks.

## Goals

Cameron describes the creation of a database of computer science journals and outlines some useful criteria for an Internet electronic library, which apply equally to the Bel Jour project. These are (1) authoritativeness and currency of information, (2) comprehensiveness within a well-defined focus, (3) integrated access to print and electronic resources, and (4) efficiency and reliability.[5]

Bel Jour is designed to function as its full title implies: it is a way to locate journals in business and economics held by the Rutgers University Libraries (RUL). Journals are selected for inclusion if they fall within the subject areas of business and economics.

Bel Jour lists only currently subscribed journals under their current titles. It is meant as a guide to the active holdings of the library in business and economics, and strives to be current, accurate, and comprehensive within these guidelines. Bel Jour provides a single point of access for both print and electronic holdings.

Bel Jour is designed to function as a quick-finding aid. The Web interface comprises only three screens—the search entry screen, a brief-results screen, and a detailed-results screen. The minimal interface speeds page loading and navigation. The brief-results screen includes abbreviated holdings information and quick links to online versions of the journal. This allows the researcher to go to the online journal from this screen with a minimum of clicks. Information about print holdings allows the researcher to quickly ascertain whether a trip to the library is necessary. Multiple links to online journals are also provided, including a special category for JSTOR journals. The full spectrum of journal coverage is thus present on one line, avoiding the need to go back and forth among multiple records to determine where the appropriate version of the journal is located. Holdings information is abbreviated to the year of coverage only. Users are referred to the online catalog for more detailed information when it is necessary to resolve ambiguities.

The detailed-results screen is designed to provide value added information about the journal for researchers and librarians. From this screen, the searcher can connect to the journal home page describing editorial and submission policies, which is often separately maintained from the publisher's homepage where the journal text is available. The researcher can check to see how many times the journal was cited according to Web of Science, to gauge the journal's relative standing in the field. The researcher can determine

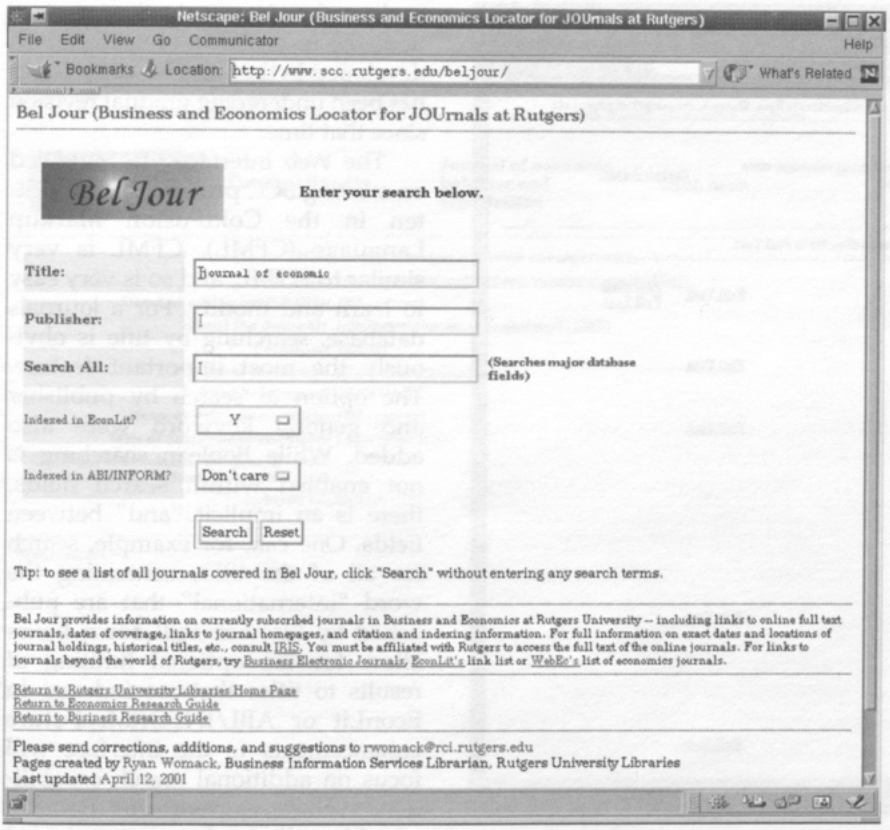



**Figure 1.** Search Entry Screen

where and if the journal is indexed. Bel Jour's indexing information is currently limited to EconLit and ABI/INFORM, but other indexing services may be added in the future.

Bel Jour is intended to facilitate browsing of journal titles as well. Clicking the "Search" button without entering any title retrieves the entire list of journals in the database. One can also restrict the search to journals indexed in EconLit or ABI/INFORM or to journals from a particular publisher. Future enhancements will include further subject classification of journals so that the researcher can easily access lists of journals within a subfield, such as accounting. A field to indicate whether or not the journal is peer-reviewed is planned.

Graduate students and faculty are Bel Jour's primary intended audience. The interface assumes familiarity with the scholarly journal environment and with the basic tools for searching and accessing these journals. The descriptive information about the journals was selected with this audience in mind.

Bel Jour is also designed to function as a tool for subject bibliographers. It allows a librarian to monitor a controlled list of titles and selectively include material from large, heterogeneous packages, from subscriptions that are outside of the selector's direct control, or from freely available government and Web publications. At Rutgers, for example, there are three business libraries on different campuses, and there are many full-text databases that are not cataloged in the libraries' online catalog. Bel Jour is one way to integrate these holdings so that a bibliographer can easily determine if a journal is available somewhere else in another format, and make informed collection decisions. The database also includes information on subscription prices for many journals. Although this information is not currently available via the Web interface, it can easily be added should this become desirable.

Maintaining the journal information in database format facilitates the addition of notes and fields that inform subject selectors about the collections, such as the fund code associated with a journal. Rutgers, like many institutions, uses a central fund to purchase electronic resources. Many journal titles that were formerly part of individual subject collections have been transferred to the central electronic fund as a result of the purchase of large packages like ScienceDirect. However, the subject bibliographer still needs to be aware that the relevant journals from these packages are available as part of the subject collection. Bel Jour facilitates tracking these titles.

The overarching goal of Bel Jour is to provide added value to the existing online catalog, journal collection, and related Web resources. It does this through the integration of these sources into one interface, and through the addition of extra search capabilities and descriptive information.

## Development of the Database and Interface

The Scholarly Communication Center (SCC) of RUL has developed several Web-accessible databases using ColdFusion software to connect to Microsoft Access databases. Ron Jantz, data librarian at the SCC, made available template ColdFusion files that provided the basic structure for the search interface.[6] These files were easily modified to search the specific fields of the Bel Jour database

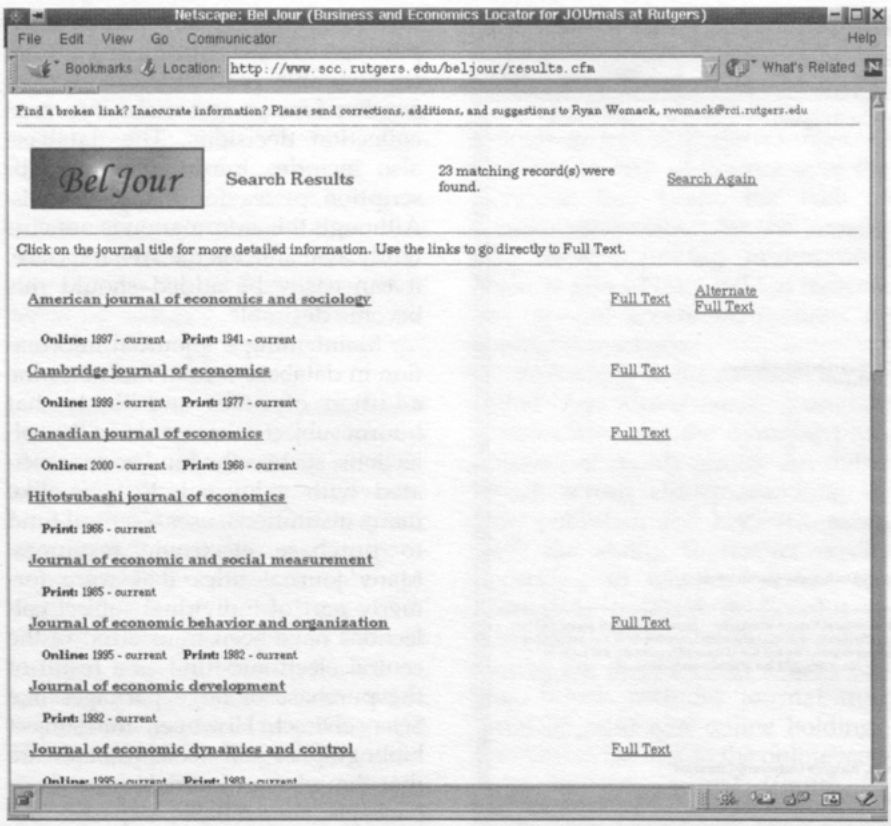



**Figure 2.** Brief Results Screen

and create an appropriate interface. The SCC provides technical support to librarians involved in creating digital projects and maintains the ColdFusion server that hosts many of these projects. This in-house wealth of expertise made the creation of the database and interface quite simple.

The journals data originated from Excel spreadsheets that were created by the author as part of a serials cancellation project in 1999. These spreadsheets contained pricing and holdings information for a large percentage of Rutgers' subscriptions in business and economics. The author added fields for indexing, citation counts, publisher information, URLs for online access to these files, and the other fields described above. Full-text titles from ABI/INFORM were imported as a separate process. A searchable interface to ABI/INFORM titles was a high priority since these important business journals were not included in the library's online catalog. Converting the data from Excel to Access was a simple process from within Access. The resulting Access database is mounted on the SCC's ColdFusion server so that the interactive Web pages can be delivered.

The initial phase of the project focused on entering complete information for economics titles, since these were the most complete in the original spreadsheets. The database is continually updated to expand the level of detail and inclusiveness for business-related journals across the RUL system. Approximately one thousand titles are currently included. Over time, the database will grow less rapidly and require only maintenance updates for changes in the information. The first version of Bel Jour was developed in the summer of 2000 and has been undergoing gradual revision since that time.

The Web interface was modeled on existing SCC projects, and is written in the ColdFusion Markup Language (CFML). CFML is very similar to HTML, and so is very easy to learn and modify. For a journals database, searching by title is obviously the most important feature. The option to search by publisher and general keyword were also added. While Boolean searching is not enabled within search fields, there is an implicit "and" between fields. One can, for example, search for all of the titles containing the word "international" that are published by Elsevier. Pull-down boxes allow for the optional limiting of results to titles that are indexed in EconLit or ABI/INFORM. Future enhancements to the interface will focus on additional search capabilities, such as support for Boolean searching, limits for peer-reviewed journals, and the like. Changing the interface takes relatively little time compared to the creation and maintenance of the database itself.

## Discussion

While anecdotal evidence suggests that Bel Jour is a useful service that meets the goals outlined above, no formal study of its impact has yet been done. The author plans a survey of Bel Jour users, combined with a study of database usage as recorded by the Web server logs. Information about the actual user population of the database would allow services to be designed that better meet the audience's needs. A survey would reveal which services would be the most useful additions to the database, and which aspects of the current service are not meeting user needs.

A database that focuses on current resources requires regular main-

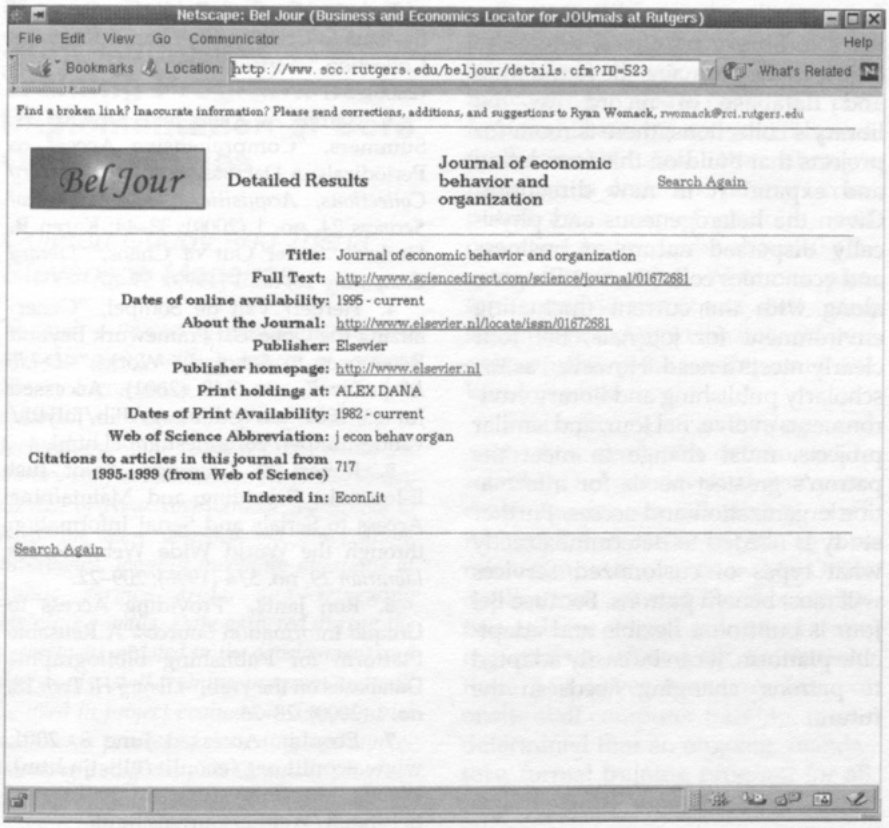



**Figure 3.** Detailed Results Screen

tenance to maintain its usefulness. It remains to be seen how much maintenance Bel Jour will require. If publishers' URLs remain relatively stable, and the library does not frequently switch among different journal providers, the maintenance burden should be manageable. However, if the Web environment changes more rapidly, the database may require frequent and intense reworking.

Bel Jour's utility is also closely related to developments in the library's online catalog and the Web environment. One inspiration for the development of Bel Jour was the increasing popularity of link lists of online journals in economics, such as those available from the EconLit Web site or through WebEc.[7] Due to their simple format and convenience, such lists are often the first stop for researchers who may be daunted by the complexities of the library's online catalog. Such global journal lists are not linked to the library's holdings, and may cause frustration when the researcher discovers that many of the journals referred to are inaccessible. Projects such as jake or databases like Ulrich's online may serve as the foundation for future finding aids, but at present they are more comprehensible to librarians than to the general scholarly audience.[8] However, if linking and authentication mechanisms become more seamless, so that patrons can check library holdings or authenticate themselves for access directly from an index or other intermediary, then the need for a locally developed service like Bel Jour may diminish.

In addition, as the information reported by the library's catalog changes, the information provided by the intermediary service must change. For example, the online catalog at Rutgers is beginning to integrate previously missing records for some journals in full-text databases like ABI/INFORM. If e-journal holdings information becomes available through the online catalog, Bel Jour will lose some of its value-added qualities. The information provided by Bel Jour must evolve to add new features of use to patrons; otherwise it will unnecessarily duplicate other efforts.

The question of which features to add becomes critical. Having the journal information in a flexible database format combined with an interactive Web interface opens many possibilities. One possibility is to develop a MyLibrary-style service, where user profiles are stored and researchers can create "hot lists" of their favorite journals, along with customized views of the related journal information. Bel Jour should also take advantage of linking technologies as they develop in order to deliver more information to the end user. The service should function as a central node that links together the many information retrieval services available to library patrons. Bel Jour can thus evolve to retain its place as the virtual center of the business and economics journal collection.

## Conclusion

Bel Jour demonstrates the benefits of using new database and Web technologies to create value-added library services that meet the needs of a specific group of patrons. Although the initial metaphor for Bel Jour was the printed list of journals, using an interactive database allows the service to be much more than a mere list. With only a few clicks of the mouse, Bel Jour allows patrons to search for and access the contents and descriptive information of journals in a carefully selected subject-specific collection

that crosses the boundaries of publishers, formats, and physical locations. Just as a pathfinder or Web-based research guide can direct patrons quickly to valuable resources, Bel Jour gives researchers greater access to the diverse journal collections of a large, heterogeneous institution. The additional descriptive information on citation rates, indexing, and links to pages about the journal provide a more complete picture of the journal than the static catalog record. Since the contents of Bel Jour are under the control of the subject bibliographer, the collection has a well-defined scope. Free resources can be added if they are of scholarly quality, and irrelevant journals that may be part of a subscribed package can be removed from the database. As the scholarly journal environment experiments with e-prints and other forms of online collaboration that compete with more traditional—and expensive—commercial publications, this feature of the database takes on added value.

The benefits of such a subject-specific enhanced journal finder must be weighed against the costs in time and effort required to develop and maintain it, and against the potential for fragmentation among different gateways to library resources. While the online catalog remains the foundation and database of record for the library's collections, there is room for projects that build on this foundation and expand it in new directions. Given the heterogeneous and physically dispersed nature of business and economics collections at Rutgers, along with the current fluctuating environment for journals, Bel Jour clearly meets a need. However, as the scholarly publishing and library environments evolve, Bel Jour, and similar projects, must change to meet the patron's greatest needs for information organization and access. Further study is needed to determine exactly what types of customized services will most benefit patrons. Because Bel Jour is built on a flexible and adaptable platform, it can be easily adapted to patrons' changing needs in the future.

## References and Notes


1. Bel Jour. Accessed June 11, 2001, www.scc.rutgers.edu/beljour.

2. Steve Shadle and Alex Wade, "Putting It All Together: The Involvement of Technical Services, Public Services, and Systems to Create a Web-based Resource Collection," *Serials Librarian* 38, no. 3/4 (2000): 341–7.

3. Carole A. Kiehl and Edward H. Summers, "Comprehensive Access to Periodicals: A Database Solution," *Library Collections, Acquisitions, and Technical Services* 24, no. 1 (2000): 33–44; Karen R. Harker, "Order Out of Chaos," *Library Computing* 18, no. 1 (1999): 59–67.

4. Herbert Van de Sompel, "Generalizing the Open Url Framework Beyond References to Scholarly Works," *D-Lib Magazine* 7, no. 7/8 (2001). Accessed June 5, 2001, www.dlib.org/dlib/july01/vandestompel/v07andestompel.html.

5. Robert D. Cameron, "Not Just E-Journals: Providing and Maintaining Access to Serials and Serial Information through the World Wide Web," *Serials Librarian* 29, no. 3/4 (1996): 209–22.

6. Ron Jantz, "Providing Access to Unique Information Sources: A Reusable Platform for Publishing Bibliographic Databases on the Web," *Library Hi Tech* 18, no. 1 (2000): 28–36.

7. EconLit. Accessed June 8, 2001, www.econlit.org/econlit/ellistjn.html. WebEc. Accessed June 8, 2001,www.helsinki.fi/WebEc/journals.html.

8. Daniel Chudnov, Cynthia Crooker, and Kimberly Parker, "Jake: Overview and Status Report," *Serials Review* 26, no. 4 (2000): 12–17.

# Tutorial

## Designing a Library Staff Computer Training Program: Implementation of Core Competencies

*Colleen Cuddy and Trisha Stevenson Medeiros*

*Through a process of data gathering and analysis along with a review of current library literature, a need was identified to establish a continuum of staff training in technology. The educational needs of library staff at the Ehrman Medical Library (EML) of New York University School of Medicine were assessed via observation, interviews, logging incoming questions at public service areas, and reviewing Helpline e-mails. Data gathered during the assessment assisted in the development and design of a staff training program and will be used in project evaluation. This tutorial outlines a process of designing and implementing a staff computer-training program while providing general tips and strategies.*

The Ehrman Medical Library (EML) has sixty-one FTE staff, a mix of professional and paraprofessionals working both part- and full-time. The majority of staff are onsite; however, about 5 percent work at affiliated libraries offsite at branch facilities. The level of comfort with technology and computer proficiency varies greatly among staff. Each staff member has access to a personal computer and most staff members serve the public at one of three areas: the circulation desk, the information desk, or the computer media center. All professional librarians and several paraprofessionals staff the library's information desk. The library's computers are a mix of platforms and operating systems with about 60 percent of staff using Windows platforms and 40 percent of staff using Macintoshes. Approximately 90 percent of the public computers are Windows-based.

Through data gathering and analysis, a need was determined to establish a base level of computer competencies for all library staff. Although they were eligible to take microcomputer classes in MS Office applications offered by the institution at large as well as Internet training courses offered by the library, there had never been any form of onsite staff computer training. It was determined that an ongoing, mandatory, formal training program for all library staff would prove most advantageous.

The project as conceived by the systems librarian and the coordinator for user services had two main objectives—first, to improve the staff's ability to troubleshoot their own desktop machines, and second, to improve the level of assistance provided to patrons in public service areas. An upcoming library renovation would combine all points of public service into one location. This too was considered in the need for a core level of skills.

Top-level library administrators were queried during one of their regular weekly meetings to ascertain their interest in such a project and to determine if such a program would be supported. Once verbal approval was received from the library director, project development got underway. It is vital to get the support of the library director and department heads before such an undertaking. There must be a buy-in at all levels, but rallying department heads at the onset proved to be an important step in securing attendance and overall support of all staff members for the training program.

## Data Gathering and Analysis

A literature review is a good start in developing an ongoing training program. A review of articles on staff computer training reinforced the value of such an undertaking. By further focusing the literature review on academic libraries, good starting points for actual program design began to emerge. In the academic institutions, in-house training programs were more likely to be developed and implemented, while training programs in corporate settings tended to be outsourced, something that library budgets do not often provide for. Academic libraries, in particular, are able to use valuable staff resources to cull instructors and budget constraints have reinforced the do-it-yourself attitude in many individual libraries.

A training model designed for the University of Wisconsin–Madison was ultimately adopted. This model provided three areas for training: the operating system, hardware basics and troubleshooting, and software basics.[1] Searching skills and techniques were considered an important fourth area of training for the staff that worked in public service areas. Using these four broad areas as a recommended basis, it is important to next get library-specific information to cover in each class. Starting with a basic model is handy, but in order for the courses to be truly effective at individual libraries, course designers must understand which areas are deemed the most important by their staff and also which areas are causing the most problems or computer down-time.

Four methods were used to gather information pertinent to the

**Colleen Cuddy** (cuddy@library.med.nyu.edu) is the Systems Librarian at the Ehrman Medical Library, New York University School of Medicine. **Trisha Stevenson Medeiros** (tmedeiros@purnell.org) is currently Library Director at the Purnell School in Pottersville, New Jersey. Previously, she served as Coordinator of User Services at the Ehrman Medical Library of the New York University School of Medicine.

EML: reviewing Helpline e-mail logs, logging questions at public service points, observing staff, and interviewing staff. Staff and patrons are encouraged to report all computer problems to Helpline, an e-mail alias that is directed to all members of the library's systems staff. Systems staff is then encouraged to respond to problems and sum up the solution for the Helpline alias. These systems-staff replies, junk mail, and all duplicate messages were removed from the logs; the remaining e-mail messages were analyzed for two three-month time periods in 1999. The problems reported were assigned to one of thirteen categories that became apparent as the Helpline logs were reviewed. Of these areas, printing problems were the most frequently reported, followed by hardware and network problems. Other problem areas uncovered included: general software issues; problems using Microsoft Office; nonbooting or freezing of computers; problems accessing servers; student academic courseware; e-mail; Web-browser problems; computer backup; terminal problems; and problems with locally administered database accounts. These thirteen categories were then assigned to one of the four training areas. Categories will differ between institutions given their individual needs. If this analysis was completed today on more recent logs, different categories might emerge.

Realizing that all computer problems are not reported to systems staff via e-mail, it is important to have other means of data gathering. Therefore, staff participated in the study by logging questions and problems at public service areas. Additionally, the systems staff logged all problems reported via phone or in person. Logs were kept for one-week periods at the circulation desk, the information desk, the computer media center, educational services, and the library systems office. These logs can be reviewed to assist with adding pertinent material to each of the four broad course areas: the operating system, hardware troubleshooting, software troubleshooting, and searching skills and techniques.

Observation and interviews with library staff should be used to flesh out the process to find out if there are problem areas that simply are not being reported. Observation and interviews proved to be the most difficult of data-gathering methods to carry out, but ultimately the information gathered thereby was useful and is recommended to others completing this process. It was difficult to simply observe and let the problem run its course without jumping in to assist. Some staff seemed nervous during the interview process, and some felt that they were being evaluated. It was important for the observer to stress the goals of the project clearly during these processes.

## Course Design

After analyzing the data it was decided to split the training into two sections: basic computer classes and information services classes. Course design would focus on basic computer classes first, and information services classes would be designed and taught upon completion of the first round of computer training. The basic computer classes consisted of six courses, including: The Operating System, Computer Hardware and Troubleshooting, Net Applications, Protecting Your Computer, The Ontime Calendar (scheduling software), and Introduction to Microsoft Office. Because staff use two different hardware platforms it was necessary to design sections for both Windows and Macintosh users.

The topics covered in each Windows class are outlined in the figure; similar objectives were met in the Macintosh classes. Key elements of the Windows and Macintosh operating systems were covered in the first course. Class two was a basic hardware and troubleshooting class. The idea was to establish a list of procedures to try before reporting a problem to Helpline, and to help staff feel more comfortable with all the devices attached their computers. Reports of printing problems were widespread; therefore, separate sessions for each library printer were offered to staff. These sessions, about fifteen minutes in length, were held in each department or unit and focused on such skills as clearing paper jams, dealing with common error messages, and replacing toner.

Four classes were chosen to represent the area of software reviews. The idea was to make staff more productive users of frequently used software applications and help them troubleshoot software errors. The first class, Net Applications, focused on using Netscape and Eudora, the library's supported browser and e-mail client. The second class, Protecting Your Computer, discussed using virus protection software, file backup and recovery, and basic security dos and don'ts. The third class focused exclusively on the library's Ontime calendar system, and the fourth class was an introduction to MS Office Suite providing productivity tips and tricks.

Outside of content, there are several additional areas that need to be considered with regard to class length, location, and evaluation when designing the courses. It was decided that each class would be held for ninety minutes in the library's computer classroom. Online courses which may have been convenient for offsite staff were rejected for the initial stage of the project because development would be too time consuming for course developers and not accommodating to beginners. However, online courses are being considered for review courses. The systems librarian would teach classes with the help of the coordinator of user services. Other library staff are recruited as necessary.

One of the most difficult challenges in designing the courses can

| Title | Topics |
|---|---|
| The Operating System | The Windows desktop; creating folders, icons, files; the control panel; Explorer; error messages; how to start, shutdown, and log on to network resources |
| Basic Hardware and Troubleshooting | What is a computer; input and output devices; step-by-step guide to troubleshooting; printer basics |
| Net Applications | Eudora (creating mailboxes; address books; tools, options, and settings, including directories like PH; memory usage tips); Netscape (preferences and settings, plugins, common error messages) |
| Protecting Your Computer | Security do's and don'ts; installing, updating, and maintaining virus protection; file backup and recovery |
| Using the Ontime Calendar | Calendar navigation; calendar and group types; scheduling appointments and meetings; general calendar maintenance |
| Introduction to Microsoft Office | Key concepts of Word, Excel, PowerPoint, and Access: What program to use? When? Review of staff questions |

**Figure.** Topics in Windows Classes

be deciding on a level of instruction, due to the varied competency levels of staff. It was decided that the best approach in this case would be to make each course a beginners-level (so everyone could understand the concepts) while keeping the pace fast. Intermediate courses are added as appropriate; for example, an intermediate-level course for the library's calendar system was requested and created. All staff are encouraged to bring specific problems to the classes and to have questions ready. Competencies were reinforced by requiring class registration via the library's calendar system. Staff who are proficient in one of the areas being taught can opt out of a class with permission of their department head and the course instructor.

Five or six sections of each course were held from April through August 2000 with make-up classes held during the fall of 2000. Additional classes are held as new staff join the library or new applications or software are adopted. Each participant receives a certificate signed by the library director for the successful completion of a course and a copy is placed in the personnel record file. An evaluation form was developed and is given to each participant for every class. Finally, attendance is taken and reports are sent to department heads who follow up with their staff regarding attendance, relieving the instructors of this duty. Course handouts, syllabi, and schedules are available on the library's Web site.

## Evaluation

At this point the quantitative research on the EML's staff training project is just beginning, but anecdotal research has proven the classes to have been a success. It was challenging to coordinate a training program in addition to one's day-to-day duties, but the time was well spent. The staff seem more confident in using their computers and helping others. The use of the library's calendar system has increased, and staff appear to be using e-mail more effectively both in reporting computer related problems to Helpline and in general communication.

Initially, the instructors considered implementing a pre- and post-test of individuals' computer skills. This idea was rejected because the library was reorganizing its department structures and preparing for a renovation, and it was felt that testing the staff on their computer skills would cause additional stress. However, without pre- and post-tests it is difficult to show statistically significant changes in computer use. Institutions desiring to show statistical significance should consider a standard pre- and post-test for each course.

The course evaluations give an indication of the students' perceptions of their computer knowledge pre- and post-training. An evaluation developed for the library's Internet training courses was adapted for the computer training courses. Using a rating scale of one (lowest) to five (highest), students were asked to rate their knowledge of the subject matter covered before and after each class. Overall, an increase was seen in the attendees' perception of their computer skills.

Participants were asked to suggest areas for improvement and future training on the evaluation forms. Future training was requested in file management, installing and uninstalling software, Web publishing, and cleaning viruses. Most participants requested handouts for such areas as key commands, troubleshooting checklists, and lists of common error messages. Incidentally, these documents were available on the course Web site. It is not apparent if students found the materials lacking or if they were not fully aware of the Web-site resource. There were

many requests for longer and more in-depth sessions on MS Office, Netscape, and Eudora.

A comparison of pre- and post-training Helpline logs and logs of questions at key information points has begun. It is hoped that this difficult process will provide further information on the benefits of the training program. It is also hoped that new areas for training will emerge from this process.

## Conclusion

This tutorial is provided as a basic model. Individual libraries will need to alter the specifics, but by focusing on the five steps outlined above—performing a literature review, data gathering, data analysis, course design, and course evaluation (and continued re-evaluation), course instructors should be well on the way to implementing an ongoing staff training program. A new position of coordinator of computer technology training was developed to address future training needs of both the library staff and the staff of the school at large. This person will be responsible for furthering the goals of this project by developing online tutorials and new areas for training. The staff of the EML is fortunate to now have a staff member dedicated to staff training. However, much can be gained by a little staff cooperation and basic computer skills training.

## Reference


1. Krissoff, Alan and Lee Konrad. "Computer Training for Staff and Patrons." *Computers in Libraries* 18, no. 1 (1998) : 28–32.


## Bibliography


Balas, Janet. "Online Training Resources." *Computers in Libraries* 18, no. 1 (1998): 36–38.

Brandt, D Scott. "Compartmentalizing Computer Training." *Computers in Libraries* 18, no. 1 (1998): 41–44.

Kirkpatrick, Teresa. "The Training of Academic Library Staff on Information Technology within the Libraries of the Minnesota State Colleges and Universities System." *College and Research Libraries* 59, no. 1 (1998): 55–59.

Krager, Dwayne. "Technology Training Dilemma." *Legal Assistant Today* 18, no. 5 (2001): 38–40.

Krissoff, Alan and Lee Konrad. "Computer Training for Staff and Patrons." *Computers in Libraries* 18, no. 1 (1998): 28–32.

Lippert, Margret. "Continuing Computer Competence: A Training Program for the '90s." *Bulletin of the American Society for Information Science* 20, no. 3 (1994): 18–19.

Marmion, Dan. "Facing the Challenge: Technology Training in Libraries." *Information Technology and Libraries* 17, no. 4 (1998): 216–18.

Massis, Bruce E. "How to Create and Implement a Technology Training Program." *American Libraries* 32, no. 9 (2001): 49–51.

Messmer, Max. "Building Employee Knowledge: A Look at Training Programs." *Strategic Finance* 83, no. 2 (2001): 8.

Morgan, Eric Lease. "Computer Literacy for Librarians." *Computers in Libraries* 18, no. 1 (1998): 39–40.

Tennant, Roy. "The Virtual Library Foundation: Staff Training and Support." *Information Technology and Libraries* 14, no. 1 (1995): 46–50.

# Book Review

Tom Zillner, Editor

## Information Architecture for the World Wide Web

by Louis Rosenfeld and Peter Morville. Sebastopol, Calif.: O'Reilly, 1998. 202p. $29.95 (ISBN 1-56592-282-4).

*Information Architecture,* published in 1998, is a relatively old book. In Internet time, that's at least a couple of generations ago. I think it's worth attention because it's a classic volume that will help anyone interested in designing a Web site from the ground up or attempting to fix a Web site that is not as usable as it might be.

This is a book by librarians (although they've migrated into the field of information architecture) but it is not a book that is primarily for librarians. The language of the librarian creeps in a bit, most notably when the authors discuss controlled vocabulary, but the real meat is not masked by librarian-speak. Rosenfeld and Morville are subtler than that. They present important concepts that are familiar to librarians but couched in plain language that uses plenty of examples.

The book hits the ground running. Rosenfeld and Morville recommend that the people who will be building a Web site get together in what they call "Consumer Sensitivity Boot Camp," in which they air both the things they hate and like about the Web. Ultimately the goal is to produce lists of things that work and don't work and use this as a background for the actual design and production of the site.

The book then pulls back from an active start to describe the role of the information architect and some suggestions for who might take on this role. Because many people are unfamiliar with information architecture and architects, it is a good idea to provide some coverage of these terms as well as contrasting the role of the information architect with the roles of others on the design team. It's important to note that this book targets both the large commercial project where information architects might be part of an outside consulting firm and the smaller project consisting of just a few people and a miniscule budget.

Creating a Web site starts with determining an organization scheme. An example of a organization scheme is the alphabetical arrangement of a telephone book or the layout of a supermarket. These are two very different schemes, suggestive of the wide range of possibilities. It should ordinarily take some time to determine the best organization scheme or schemes for a Web site. A different but related task is looking at possible organization structures for the site. An example of an organization structure is the hierarchy, the most commonly used structure in site design. Other examples of Web site organization structure include hypertext, where there are "content chunks" connected via links. There are pros and cons to each structure and these are well covered by Rosenfeld and Morville. An interesting aspect of the discussion of organization structures is a seeming bias toward a relational database model. Although one can glean some of the benefits of such a model from the discussion of database implementation, it's not clear that the author's explication of this model provides a solid rationale for its use.

If this all sounds a bit complicated that's because it is, but it's important to have a grounding in organization schemes and structures before taking any further steps in the design process. The next step is to consider possible navigation systems to employ. Hierarchical navigation systems are the norm, starting out at a home screen and filing down to subsidiary pages through the use of top-level links. Rosenfeld and Morville point out that this sort of navigation is often of limited use and other navigation methods should be considered, such as global navigation, where it is possible on any given page to navigate via tools similar to navigation bars. Local navigation is the use of additional tools for a subsite of the larger site and it is a complement, not a replacement, for global navigation.

The authors also address labeling systems. It really takes a librarian (or an information architect who has a librarian's training) to appreciate the importance of labeling systems. There are many species of labels within Web sites, but they should all be consistent and clear with each other and fit into a cohesive whole. While such a directive may sound vague and unclear, the authors do a good job of explaining the importance and application of labeling systems within a Web site. In this chapter, they discuss controlled vocabularies and thesauri and the value of their use in selecting appropriate labels. The authors even mention Library of Congress Subject Headings, although they point out that given its goal of describing "the universe of knowledge," the subject headings are wildly inappropriate for use as a thesaurus. The key is to find a narrowly focused thesaurus specific to the information the Web site will carry. Other types of tools to construct labeling systems are also covered. All in all, this chapter well deserves the twenty-five pages devoted to it.

Similarly, there is a comprehensive chapter on selecting indexing and searching systems. While most of us tend to throw in a search engine that's linked from any page and performs a simple search on the entire site, this isn't necessarily desirable. In fact, the chapter starts with a discussion of instances when you should not make your site searchable. Following are sections discussing searching behavior and search interface design. The important point of this chapter is that indexing and searching are not one-size-fits-all propositions; real work is required to determine how people will use your Web site and to construct a scheme that matches their behavior.

In these beginning chapters, which account for a good portion of the book, emphasis is placed on the foundation pieces of information architecture. The authors then move from these foundation pieces to actual practice. The first step in creating a Web site is research. It's important to stay with this process rather than leap into construction. This is a common mistake, particularly where budgets are small or where there is a rush to get "something" up, no matter how ill thought out that something might be.

Research must include face-to-face meetings. Other types of communication just don't work when discussing and working on an inherently visual medium. Rosenfeld and Morville recommend that the first item on the agenda for the first meeting be a critique of existing Web sites. This critique cannot effectively take place via phone or e-mail. What the participants say is partly a matter of consensus reached through exploration rather than just a yea or nay vote. The critique also is an ice-breaker, getting the members of the team better acquainted. If there is just one person responsible for the site, that one person should put together a team of people who have a vested interest in the site's content, whether they are colleagues in other departments of the same institution or users from outside the organization.

The meetings should include the definition of goals, intended audience, and content. All of this is easy enough to understand from most people's experience and the clear explication of the authors. Content inventory completes the first stage of the design and implementation process. This inventory is developed from wish lists and a setting of priorities regarding information to be included on the site.

The next steps in the process fall under the heading of conceptual design. As with the rest of the book, the reader may find these activities are covered in extensive detail, but as usual the devil is in the details. Rosenfeld and Morville discuss the respective merits of high-tech white boards that capture what's written for subsequent viewing and printout versus flip charts with a bit of levity. The authors' state, "We're guessing many of these gadgets are more trouble than they're worth. Sorry for the skepticism, but what do you expect from librarians?"

Metaphor exploration is also examined. I am skeptical of the heavy-handed use of metaphor in our computerized and netted life. The desktop metaphor is useful, but it's unclear if we will ever escape it to something that better suits our increasingly Internet-centric lives. Categories of metaphor are presented and one example provided is of the Internet Public Library's (IPL) reference center (http://www.ipl.org/ref). The center is graphically represented as a room in a library, with the friendly librarian behind a desk and books labeled with hotlinks to various subject areas. This mainly works well, but the addition of a link to a multi-user object oriented environment (MOO) results in a possibly jarring addition of a sign pointing to another room. Unlike the authors, I think this metaphor works just fine. But we've all seen plenty of cases where metaphor clutters a site with graphics. Judicious use of graphics in the service of a metaphor that assists the user in navigating the site is wonderful, but metaphor is easily abused.

Scenarios are presented as another part of the conceptual design process. They are simply descriptions of how various sorts of users would use the site. I say "simply," but the construction of realistic scenarios can be a daunting task. For one thing, if you don't have enough outsiders working with you to put together the site plan, it is very easy to make ill-conceived assumptions about users. Perhaps the most common tendency is to assume that users are very much like oneself, but other equally suspect assumptions may be made. This is the stage where relative outsiders may ask hard questions that result in a better set of working assumptions underlying scenarios.

One thing that Rosenfeld and Morville underemphasize is the involvement of users at some early point in the process of design and construction. Scenario building is a great place to begin user involvement, but even earlier is better. Users often traverse and utilize a site in ways never envisioned by architects and designers.

Next in the process are architectural blueprints. This is the point at which the information architect puts together a high-level diagram of the site. The high-level blueprint "shows pages, components within pages, groups of pages, and relationships between pages." Sounds more like a detailed blueprint, but that will be even more in depth and occurs at the last design stage. Prior to this the graphic designer comes forward with page design ideas leading to page prototypes and then templates. Both the high-level blueprints and the page design need to be approved by the committee formed at the project's beginning.

At the prototype stage the graphic designer comes to the fore and mounts pages for evaluation and critiquing. This is one of the first of the tangible "deliverables," and it can dazzle. As Rosenfeld and Morville suggest, this is a time when esthetics take center stage. But its important to look at the prototype with something of a jaundiced eye. The graphics can play an important role in complementing the content or they can overpower or confuse the user. It's best to use a graphic designer with Web experience instead of breaking in someone from the print world. It's also important to not spend an overly long time on this phase. I have been involved in a corporate identity project where weeks were spent getting interested parties to agree on a logo. I am sure this same phenomenon can

rear its ugly head in Web prototyping. Try to get participants to focus on how the graphics work for the site and worry less about the esthetics of the graphics themselves.

As mentioned previously the production stage is the point for detailed blueprints, as well as content mapping and a Web page inventory. For a small site this is fairly easy to put together, although small sites tend to have fewer staff committed to the project. For large projects there needs to be a small cadre of people dedicating much of their time to the site's planning and production. It's a lot of work. And then there comes the actual production of pages. Templates can help, but it's still hard work.

*Information Architecture for the World Wide Web* is certainly worth a read. It's a few years old, but has aged more gracefully than many others of its ilk by concentrating on the essentials. Anyone who creates a site is an information architect whether they know it or not. It's better to make choices consciously rather than unconsciously. I think that a weakness of the book is a failure to explore user testing more extensively, but this is something that is missed by most books in this category. All in all, a book to read in its entirety and use as a guide along the way.

## Index to Advertisers

# Instructions to Authors

## Subject Content

*Information Technology and Libraries (ITAL)* publishes material related to all aspects of libraries and information technology, including digital libraries, metadata, authorization and authentication, electronic journals and electronic publishing, telecommunications, distributed systems and networks, computer security and intellectual property rights, technical standards, geographic information systems, desktop applications, online catalogs and bibliographic systems, optical information systems, software engineering, universal access to technology, futuristic forecasting, library consortia, vendor relations, and technology and the arts.

## Types of Contributions

*ITAL* includes feature articles, communications, tutorials, reviews, and letters to the editor:

**Feature articles** consist of original research or comprehensive and in-depth analyses. Although longer works may be considered, manuscripts of 3,000 to 5,000 words are most common.

**Communications** consist of brief research reports, technical findings, and application notes. Generally communications are 1,000 to 3,000 words in length.

**Tutorials** are brief, instructional pieces on topics or issues of relevance.

**Book reviews** and **software reviews** are assigned by the respective editors.

**Letters to the editor** may be submitted to the editor for possible inclusion in the journal.

## Instruction for Submissions

1. Submit original, unpublished manuscripts only. Do not submit a manuscript being considered for publication elsewhere. Authors are responsible for the accuracy of the information in the manuscript, including references and statistics.
2. Manuscripts should be machine-printed and double-spaced. Two copies should be provided. It is not necessary to provide the manuscript in electronic format until it has been accepted for publication. Author name, title, and affiliation should appear on a separate cover page only. This information should not be repeated in the text of the manuscript, including the abstract. Pages should be numbered. An abstract of 100 words or less should be provided on a separate sheet.
3. *ITAL* follows *The Chicago Manual of Style, 14th edition* for capitalization, punctuation, quotations, tables, captions, and elements of bibliographic style, including references. Spelling adheres to Webster's Collegiate Dictionary, 10th edition.
4. Information on submission of electronic copy, tables and illustrations, copyright forms, and other aspects of final copy will be provided when a manuscript has been accepted for publication.

## Review of Manuscripts

*ITAL is* a refereed journal using double-blind reviewing. The editor assigns manuscripts to reviewers who receive the manuscript with no direct information on the author or author's affiliation. Reviewers examine the manuscript considering the following questions:

- Is the topic within the scope of *ITAL?*
- Is it meaningful and relevant to *ITAL* readers?
- Does it offer a valuable addition to the literature?
- Is it timely?
- Is the presentation that of an article or merely the text of an oral presentation?
- Is it organized well, and does it have a point?
- Are the citations complete and accurate?

Upon completing the review, a recommendation is presented to the editor as to the suitability of the manuscript for publication in *ITAL.* Recommendations fall into one of the following categories:

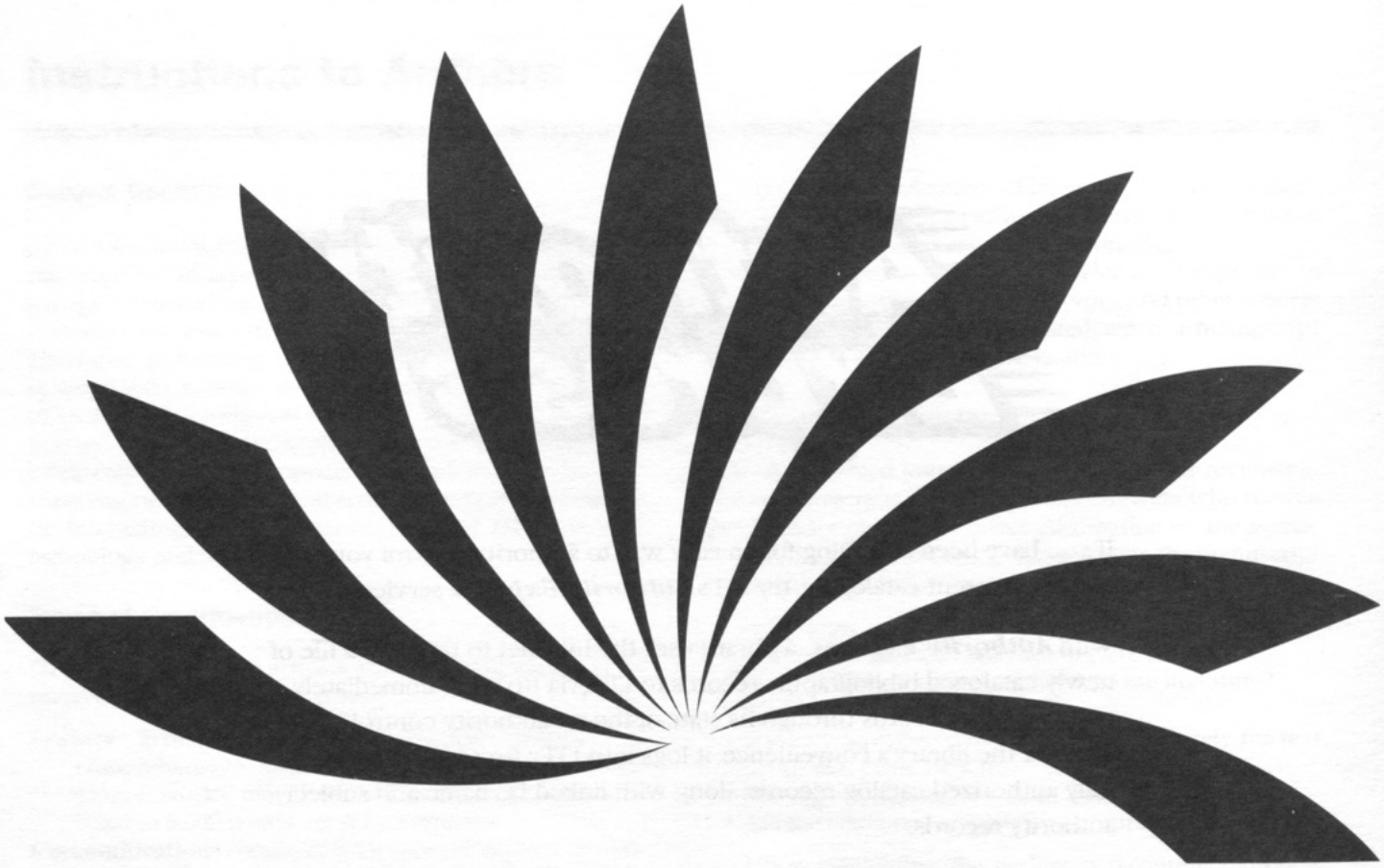
- Publishable in its current condition
- Significant and sound and should be published with only minor editorial revisions
- Basically significant and sound but requires some rewriting to make it a solid, publishable contribution
- Requires major rewriting, and it should be reviewed again after a revised draft has been received
- Does not warrant publication as a full article but might be published as a Communication
- Does not warrant further consideration by *ITAL*

The author is informed of the recommendation and any comments made by the reviewers. The review process takes four to six weeks.

## Submissions

Manuscripts for consideration should be sent to:

Dan Marmion
*ITAL* Editor
University of Notre Dame Libraries
221 Hesburgh Library
Notre Dame, IN 46556-5629



# Content Transformation Solutions

- SGML, HTML, XML
- DTD Development
- Imaging
- Web-site Development
- eBook & PDF
- *and so much more....*


Corporate Publishing Group, 11150 Main Street, Suite 402, Fairfax, VA 22030
703.352.0001 ext. 123 • iris.hanney@techbooks.com • www.techbooks.com

# NETWORK CDs... WITHOUT CDs

## SuperCD

***AFFORDABLE AND FAST***

Network your CD-ROMs today, but don't settle for yesterday's technology. Forget those big, slow and expensive CD towers. Using one high-speed CD drive and TurboCaching, SuperCD mirrors CD disks to a large hard drive. Super CD eliminates the need for multiple CD drives. It's super fast and super affordable, easy to use and easy to install.

*Thousands of satisfied users appreciate:*

- **Lowest cost per CD.**
- **Networks up to 512 CD titles.**
- **Serves 1,000 users.**
- **Fastest access times plus 66X transfer rates.**
- **Compact, fits on a shelf.**
- **Supports NT, Netware & Mac, simultaneously.**
- **Web Browser accessible.**
- **DVD compatible.**
- **Unburdens your LAN.**

BEST OF LAN TIMES 1996 SOFTWARE

Buy with confidence. Axonix backs its products with a FULL ONE YEAR WARRANTY. Plus, if you're not completely satisfied, you can RETURN THE PRODUCT WITHIN 30 DAYS for a complete refund. Axonix has a 18-year history and we understand the value of customer satisfaction.

Check our website for more SuperCD information.

For a free catalog, contact:
**Axonix Corporation**
844 South 200 East, Salt Lake City, UT 84111
800-866-9797 or sales@axonix.com

EDUCATION DISCOUNTS ARE AVAILABLE.

***Tomorrow's Library TODAY***

**1.800 866.9797**
**www.axonix.com**

AXONIX®
***Share and Save***