Communications

The Intelligent Reference Information System Project: A Merger of CD-ROM LAN and Expert System Technologies

Charles W. Bailey, Jr.

The University Libraries of the University of Houston created an experimental Intelligent Reference Information System (IRIS) over a two-year period. A ten-workstation CD-ROM LAN was implemented that provided access to nineteen citation, full-text, graphic, and numeric databases. An expert system, Reference Expert, was developed to assist users in selecting appropriate printed and electronic reference sources. This expert system was made available on both network and stand-alone workstations. Three research studies were conducted.

INTRODUCTION

From October 1989 to September 1991, the University Libraries of the University of Houston developed a prototype Intelligent Reference Information System (IRIS) that integrated CD-ROM LAN and expert system technologies. The IRIS Project was partially funded by a $99,852 Research and Demonstration Grant from the U.S. Department of Education's College Library Technology and Cooperation Grants Program.

The principal goals of the IRIS Project were to: (1) implement a CD-ROM LAN that would provide access to citation, full-text, graphic, and numeric databases; and (2) develop an expert system that would recommend appropriate CD-ROM and print reference sources.

Three research studies were conducted: (1) a CD-ROM LAN performance bench-mark; (2) a survey of user perceptions of the CD-ROM LAN; and (3) a survey of user perceptions of the expert system.

The IRIS project was intended to assist all members of the university community, plus the numerous Houston citizens who use the services of the University of Houston Libraries. This user population is ethnically diverse and multilingual.

The IRIS Project evolved from two earlier projects: (1) the Intelligent Reference Systems project, which developed an expert system for indexes and abstracts (Index Expert)[1,2]; and (2) the Electronic Publications Center project, which established a CD-ROM service that employed stand-alone workstations.[3]

EDUCOM's Educational Uses of Information Technology (EUIT) Program acknowledged the accomplishments of the IRIS Project when it named the project one of its Joe Wyatt Challenge Success Stories. The Joe Wyatt Challenge was intended to identify 100 successful applications of information technology in U.S. and Canadian colleges and universities; 101 projects were actually chosen.

PROJECT STAFFING

The IRIS Project involved staff from many parts of the library. The staff identified in the original grant proposal were mainly involved in expert system development and CD-ROM LAN technical support. As the project evolved, it became clear that additional project staff were required to plan and implement major new electronic information services, provide end-user support services, and conduct project research.

The IRIS Project Director was Robin N. Downes, director of the University Libraries. Reporting to Downes, the Project Management Group supervised the efforts of the Electronic Publications Instruction Group (bibliographic instruction and user documentation), the Knowledge Engineering Group (expert system development), and the Research and Evaluation Group (CD-ROM LAN performance benchmark and user

studies). The final project structure was as follows:

Project Director
Robin N. Downes, Director of the University Libraries

Project Management Group
Charles W. Bailey, Jr., Project Manager and Assistant Director for Systems (Chair)
Cherie Colbert, Coordinator of Library Instruction and Information Literacy
Kathleen Gunning, Assistant Director for Public Services and Collection Development
Donna Hitchings, Head of Information Services
Judy Myers, Assistant to the Director
Thomas Wilson, Head of Systems

Electronic Publications Instruction Group
Cherie Colbert, Chair
Ivan Calimano, Information Services Librarian
Carolyn Meanley, Coordinator of Government Documents
Derral Parkin, Head of Branch Libraries

Knowledge Engineering Group
Judy Myers, Chair
Charles Bailey
Jeff Fadell, Information Services Librarian
Jill Hackenberg, Coordinator of Electronic Services
Thomas Wilson

Research and Evaluation Group
Kathleen Gunning, Chair
Donna Hitchings
Kimberly Spyers-Duran, Information Services Librarian

MAJOR ACTIVITIES OF THE IRIS PROJECT

There were four major activities of the IRIS Project: (1) selection of CD-ROM databases and negotiation of network licenses for these databases; (2) selection of the hardware and software components of the CD-ROM network, installation of these components, and network implementation; (3) development of the expert system; and (4) evaluation of the performance of the CD-ROM LAN and assessment of user reactions to the CD-ROM LAN and expert system.

CD-ROM Database Selection

In order to explore the full potential of electronic information resources, the IRIS Project wanted to provide users with access to a mix of citation, full-text, graphic, and numeric CD-ROM databases. The project also wanted to select databases that supported the major disciplines taught at the University of Houston. When the project began, some CD-ROM vendors were hesitant to consider network licenses, and this limited the databases that the project could consider.

Since CD-ROM vendors were uncertain about how to price network licenses to CD-ROM products, negotiations with vendors were lengthy and there was little similarity in the obtained license agreements.[4] Initially, vendors restricted access to their products in various ways, such as the number of overall network workstations, the number of simultaneous users per database, and the workstation location. When the licenses were renewed, many vendors focused on the number of simultaneous users as the primary way of contractually limiting database access.

The nineteen networked CD-ROM databases used in the IRIS Project were:
- ABI/INFORM Ondisc
- Art Index
- Biological and Agricultural Index
- Business Dateline Ondisc
- Compact Disclosure
- Compendex Plus
- Computer Select
- General Science Index
- Humanities Index
- Microsoft Bookshelf
- The New Grolier Electronic Encyclopedia
- Ondisc ERIC
- Periodical Abstracts Ondisc
- PsycLIT
- Social Sciences Index
- sociofile
- Software Toolworks World Atlas
- Statistical Masterfile
- SUPERMAP

CD-ROM LAN

In order to provide optimal performance and maximum system design flexibility, the University Libraries decided to purchase the various components of the CD-ROM network and integrate them itself, rather than buy a "turnkey" system. This required technical expertise and a significant investment of time, especially since the University Libraries had to purchase many "plug-compatible" hardware components on state contract.

The IRIS Project designed and installed a ten-workstation CD-ROM LAN. The workstations were 16 MHz 80386SX microcomputers with 1 MB of RAM, 40 MB hard discs, and EGA monitors. The project used an IBM Token-Ring LAN that ran under Novell Advanced NetWare 2.15 Revision B. Two Meridian Data CD Net 314 CD-ROM servers were employed. A 20 MHz 80386 server was used to support NetWare and centrally provided software. The Meridian servers and the network workstations used the Meridian CD Net, NetBIOS, and NetWare DOS Client programs. The Microsoft CD-ROM Extensions software was used on workstations. The Saber LAN Administration Pack was used to provide menu-driven access to CD-ROM resources, log CD-ROM sessions in a dBASE-compatible file, control access to CD-ROM resources to comply with license agreements, and provide enhanced network security. From eight public workstations, library patrons were easily able to access any desired CD-ROM database. Library staff used two additional workstations: (1) the Information Services desk workstation was employed for ready reference searching; and (2) the computer room workstation was used to manage the CD-ROM LAN. During most of the project period, there were also four stand-alone CD-ROM workstations in the Electronic Publications Center that provided access to five checkout databases (there were three workstations in use after Reference Expert was implemented).

The CD-ROM LAN implementation process was challenging.[5] The purchase of "plug-compatible" Token-Ring boards resulted in compatibility problems with the Meridian CD-ROM server software that needed to be resolved. Some network cables were defective and had to be replaced. It was apparent that few CD-ROM vendors had designed their searching software to operate on a network, and the process of getting a CD-ROM database up and running on the network was often far more complex than it should have been. Some CD-ROM searching software would run from the Novell NetWare server; some would not. Workstation memory was another problem area; the AboveLAN software was required to free up some conventional memory in order to run some CD-ROM searching software. (At a later date, migration to MS-DOS 5.0 made it unnecessary to continue using the AboveLAN program.)

The CD-ROM LAN was made public in August 1990. It has been heavily used by the patrons of the University of Houston Libraries, and it has significantly increased user access to electronic information. In the first year of the grant (October 1989 to September 1990), CD-ROM databases were used 25,264 times (7,628 network uses in the two months that the network was available). In the second year of the grant (October 1990 to September 1991), CD-ROM use skyrocketed to 77,031 database uses (68,784 network uses), which reflects the fact that the network was available during the entire year. The highest number of CD-ROM database uses in a single month of the grant period occurred in April 1991, with 9,678 uses (8,581 network uses).

Effective management of the IRIS CD-ROM LAN required several policy decisions.[6] In order to provide equitable access to the popular CD-ROM LAN, workstations were scheduled in one-half-hour blocks during peak use periods. To provide the widest breadth of resources, very few CD-ROM back files were mounted on the CD-ROM servers; the majority of back files were made available on a checkout basis. Staffing was significantly increased in the Electronic Publications Center to provide more user assistance.

To support the effective use of the IRIS CD-ROM network, the University Libraries developed new user documentation and expanded its instructional efforts. It created a Quick Reference Card for each CD-ROM database that provided brief instructions for using the database.[7] The Quick Reference Card was available as a handout. It was also included in a three-ring notebook that provided users with more in-depth information about the CD-ROM database, including an Advanced Search Tips guide and, for citation databases, a list of indexed journals. Staff also modified existing publications, course-related workshops, and ongoing library tours to include coverage of IRIS resources. Special hour-long classes about generic CD-ROM searching techniques were initiated.

The CD-ROM LAN has been highly reliable. There is a scheduled weekly

maintenance period for the installation of new CD-ROM releases, hardware and software upgrades, minor workstation repairs, and other purposes.

Reference Expert

The development of the expert system was a complex process.[8,9] Three prototypes were developed using KnowledgePro (expert system shell with built-in programming language), VP-Expert (expert system shell), and PDC Prolog (logic programming language) prior to the development of the production system, which was written in PDC Prolog. Library staff informally evaluated the Reference Expert prototypes during the system development process. By using expert system shells, the developers could quickly create working mock-ups of the desired system. Programming in PDC Prolog required significantly more effort; however, it gave the developers a higher level of performance and much greater control over the workings of the system than an expert system shell would. The expert system was designed so that the knowledge base was contained in ASCII files, which nonprogrammer library staff could modify using word processing software. This strategy also enhanced the transportability of the system, because other libraries could customize the knowledge base for local use. To improve transportability further, a window on the first screen of the system was designed to display introductory text that was contained in an ASCII file. The system was menu driven, simplifying its use for the diverse user population of the University of Houston Libraries.

Much of the expert system design process focused on questions about what user, information need, and reference resource characteristics were important and how these characteristics were related to each other. The Knowledge Engineering Group (KEG) identified numerous potentially useful characteristics, but determining how they related to each other was very difficult. Those relationships that were deemed to be important had to be embodied in a menu-driven user interface.

After lengthy analysis of the reference process, KEG determined that the central variables in Reference Expert would be three reference resource characteristics: content type (e.g., addresses, citations, and definitions), format (i.e., CD-ROM or print), and

subject coverage.[10] It was a significant breakthrough when KEG decided that the traditional category that a reference work was placed in (e.g., dictionary, directory, or index) was not important. Reference works in the same category could be quite different (e.g., an engineering handbook and a psychology handbook). The key to describing reference works effectively was to identify the kinds of information that they contained—their content types. This simple idea was a powerful tool for classifying specific reference works.

The Knowledge Engineering Group developed a frame-based knowledge representation scheme to describe reference materials. Frames are a compact and flexible way of organizing knowledge in a hierarchical structure.[11]

To build the knowledge base, KEG selected 340 printed and CD-ROM reference sources, and it coded these reference sources using the knowledge representation scheme. This effort focused on the most heavily used reference sources in the collection. KEG recorded title and location information for each reference source, and the committee assigned appropriate subject headings and content types to the work. Comments about the proper use of the reference source were added as needed, and information about the coverage (i.e., selective or comprehensive) of indexes was optionally included for those sources. KEG recorded the relationship between each subject heading and its content types, which required that the committee identify what content types had been assigned to the sources classified under the subject heading. KEG also added descriptions for major subject headings to the knowledge base. The resultant knowledge base was fairly large, containing over 230,000 bytes of information.

The subject heading scheme used in Reference Expert was adopted from an earlier one used in the Index Expert system. The content type scheme was created from scratch. Both were refined in an iterative fashion, with changes being made as new sources were examined.

PDC Prolog, a logic programming language, proved be a good tool for developing Reference Expert's inference engine. An inference engine is the expert system component that emulates the human reasoning

process. The program was designed so that the entire knowledge base was loaded into memory from a disc file. The inference engine then used the information in the knowledge base to recommend reference sources. The knowledge base was near its maximum size for network workstations, which had less free memory than stand-alone workstations.

Unfortunately, PDC Prolog was a difficult tool to use for interface design tasks that would be relatively simple with a procedural language like C. The backtracking mechanism used in PDC Prolog, which searches the knowledge base for additional information when the inference engine reaches a dead end in its reasoning process, gives the language great power and flexibility; however, it also makes it difficult to program the parts of the system that require a predictable sequence of activity. Backtracking can be selectively disabled, and this was done, but this is a rather arcane art. A considerable amount of energy was devoted to making the user interface visually attractive and easy to use.

Given the significant differences between logic and procedural programming languages, a novice logic language programmer faces unknown territory, where his or her prior procedural language programming experience may be more of a hinderance than a help. The fact that the University Libraries had previously developed the Index Expert system using both Turbo and PDC Prolog made the programming effort go more quickly than it would have otherwise.

Although PDC Prolog made more effective use of conventional memory (first 640 KB of memory) and ran faster than the expert system shells that were employed in the project, the size of the knowledge base severely degraded system performance until a method of reducing system load was devised. Since this solution required that portions of the knowledge base be deleted from memory during program execution, the knowledge base had to be reloaded each time the program ran. System performance was now speedy, but there was a startup delay each time the system began a new session. It was decided that this was an acceptable tradeoff. It would have been possible to store the knowledge base in a B+ tree disc file; however, this would have made the program considerably more complex. Subsequent testing on an 33 MHz 80386 workstation revealed

that this platform provided very good system performance.

Many expert systems developed by libraries use low-cost expert system shells, and they often have a fairly limited scope. For example, a recent survey of Association of Research Libraries members found that four out of the six identified expert system development projects were using expert system shells.[12] Expert system shells can be effective for small-scale systems, but may not be adequate for larger, more complex systems.[13] The experience of the IRIS Project shows that expert system development with a logic programming language also can push the limits of affordable microcomputer technology.

Originally, we had intended to connect users to recommended CD-ROM LAN databases from within the expert system. It was decided that, given the ease of CD-ROM access from the LAN menu system, this linkage was unnecessary. Since many libraries who might be interested in using Reference Expert may not have CD-ROM LANs, omitting this feature also improved the transportability of the expert system.

Reference Expert was made public in June 1991. The system was available on the ten CD-ROM LAN workstations, on three stand-alone CD-ROM workstations, and on a dedicated workstation at the entrance of the library. Preliminary data indicate that it will be a popular service. From the time the system became public to the close of the grant period (June 1991 to Sept. 1991), Reference Expert was used 3,571 times. There was a steady increase in use of the system: 229 uses in June, 656 uses in July, 937 uses in August, and 1,749 uses in September. (These figures exclude use on a staff workstation used to manage the network.)

The system works as follows. After viewing an introductory screen, the user is presented with a subject menu. Each subject is described in more detail in a window at the bottom of the screen. If the user selects a subject that has lower-level headings in the subject hierarchy, the user can either choose the current subject heading or pick a lower-level heading. The user is shown the lower-level subject headings in a window at the bottom of the screen. After a subject is selected, the system determines whether there are content types associated with the subject (e.g., addresses, brief biographies, or

definitions), and if so, it displays a content type menu. Content types may be qualified by language, such as "Translations of words (English, French)," location, such as "Addresses (Houston)," and other criteria. Once the user selects a content type, the system determines if both print and CD-ROM resources exist for the chosen subject and content type, and if so, it displays a format selection menu. Finally, the system retrieves reference sources that match the selected subject, content type, and format criteria. After reading the list of sources, the user can print it. The current date, selected subject, and selected content type are automatically recorded in an ASCII log file, which is designed so that it can be easily imported into a dBASE-compatible system.

Two versions of the system are used in the University Libraries: (1) the version used on LAN workstations that executes once, returning the user to the LAN menu; and (2) the version used on the stand-alone workstation that automatically reloads after executing.

Reference Expert is available at no charge.[14] Recipients are licensed to use the program for noncommercial, educational purposes. As of July 1992, over 400 copies of the program have been distributed to libraries, library schools, computer centers, and other users. A 16 mHz 80386SX computer with an EGA or VGA monitor, 1 MB of RAM, and a hard disc is the minimum recommended hardware configuration needed to run the system.

Research Studies

Three formal system evaluations were conducted during the IRIS Project: (1) a CD-ROM LAN performance benchmark; (2) an assessment of user reactions to the CD-ROM LAN; and (3) an assessment of user reactions to the expert system. Selected highlights of these studies are presented here. The detailed results of these studies will be presented in future papers by members of the IRIS Research and Evaluation Group.

The performance benchmark showed that response time increased substantially as the number of simultaneous users of a CD-ROM LAN database increased. For example, the average increase in response time between one user and nine users for three of the CD-ROM databases (ERIC, Humanities Index, and PsycLIT) was 59.09 seconds. The bench-

mark also revealed that the degree of performance degradation under load varied considerably by CD-ROM product. With nine simultaneous users, there was a 65.08 second difference between the fastest and slowest response time for the three previously mentioned databases. The results of the performance benchmark reflect the specific hardware, software, and CD-ROM databases used by the IRIS Project and the particular testing methodology employed. Given the benchmark results, we currently plan to limit access to a maximum of ten simultaneous users per CD-ROM database.

The CD-ROM LAN survey indicated that the majority of users reacted very favorably to CD-ROM databases, saying that they found information more quickly (89.4%) and more easily (84.2%) than in printed sources. Most users (85.1%) believed that they found more helpful information in CD-ROM databases than in print sources. An overwhelming majority of users (97.5%) agreed that it was valuable having CD-ROM databases in the library.

The Reference Expert survey revealed that the majority of users preferred to use Reference Expert than to use printed guides (62.9%) or to ask library staff for assistance (56.3%). Most users said that they would consult the sources recommended by Reference Expert (72.5%) and use the system in the future to find reference sources (74%).

CONCLUSION

Given the state-of-the-art of the underlying technologies, CD-ROM LANs and expert systems still suffer from performance and other technical limitations; however, the Intelligent Reference Information System Project demonstrated that they can be used very effectively in libraries. The IRIS Project required a fairly high-level of technical support and the participation of a wide variety of library staff. Other libraries that want to develop systems of similar scope and complexity may have comparable staffing needs.

Given the success of the IRIS Project, the University Libraries are planning to expand the CD-ROM network significantly, increasing both the number of workstations and the number of networked CD-ROM databases.

The University Libraries have established the Reference Expert Task Force to continue the development of the expert system. This

group will design and program a prototype of the next version of the system using Visual Basic in a Microsoft Windows environment. The University Libraries are interested in adding significant depth to the decision making process in Reference Expert by taking into account more reference resource characteristics and adding both user and information need characteristics. The University Libraries are also interested in exploring the use of an object-oriented programming language like C++ for producing the production version of the new Reference Expert system. This should reduce the complexity of interface design and increase the overall capacity and performance of the system.

REFERENCES AND NOTES

1. Charles W. Bailey, Jr., "Building Knowledge-Based Systems for Public Use: The Intelligent Reference Systems Project at the University of Houston Libraries," in *Convergence: Proceedings of the Second National Conference of the Library and Information Technology Association, October 2–6, 1988*, ed. Michael Gorman (Chicago: American Library Assn., 1990), p.190–94.
2. Charles W. Bailey, Jr., and others, "The Index Expert System: A Knowledge-Based System to Assist Users in Index Selection," *Reference Services Review* 17, no.4:19–28 (1989).
3. Charles W. Bailey, Jr., and Kathleen Gunning, "The Intelligent Reference Information System," *CD-ROM Librarian* 5:14,16 (Sept. 1990).
4. Thomas C. Wilson, "Zen and the Art of CD-ROM Network License Negotiation," *The Public-Access Computer Systems Review* 1, no.2:4–14 (1990). To retrieve this article from the University of Houston's list server, send the e-mail message GET WILSON PRV1N2 F=MAIL to LISTSERV@UHUPVM1 (BITNET) or LISTSERV@UHUPVM1.UH. EDU (Internet).
5. Thomas C. Wilson and Charles W. Bailey, Jr., "The Intelligent Reference Information System CD-ROM Network," in *Library LANs: Case Studies in Practice and Application*, ed. Marshall Breeding (Westport, Conn.: Meckler, 1992), p.157–71.
6. Kathleen Gunning, "The Intelligent Reference Information System: The Effect on Public Services of Implementing a CD-ROM LAN and Expert System," *Library Administration & Management* 6:146–53 (Summer 1992).
7. The Quick Reference Cards developed by the IRIS Project are available from: LOEX Clear-

inghouse, University Library, Eastern Michigan University, Ypsilanti, MI 48197.
8. The reader unfamiliar with expert system technology may want to consult the following introductory works for further information: Ralph Alberico and Mary Micco, *Expert Systems for Reference and Information Retrieval* (Westport, Conn.: Meckler, 1990); Rao Aluri and Donald E. Riggs, eds., *Expert Systems in Libraries* (Norwood, N.J.: Ablex, 1990); Louis E. Frenzel, Jr., *Crash Course in Artificial Intelligence and Expert Systems* (Indianapolis: Sams, 1987); F. W. Lancaster and Linda C. Smith, eds., *Artificial Intelligence and Expert Systems: Will They Change the Library?* papers presented at the 1990 Clinic on Library Applications of Data Processing, March 25–27, 1990 (Urbana-Champaign, Ill.: Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign); George F. Luger and William A. Stubblefield, *Artificial Intelligence and the Design of Expert Systems* (Redwood City, Calif.: Benjamin/Cummings Pubs., 1989); A. Morris, "Expert Systems for Library and Information Services—A Review," *Information Processing and Management* 27, no.6:713–24 (1991); and Ernst J. Schuegraf, "A Survey of Expert Systems in Library and Information Science," *The Canadian Journal of Information Science* 15:42–57 (Sept. 1990).
9. A detailed article about Reference Expert by Judy E. Myers, Charles W. Bailey, Jr., Jeff Fadell, Jill Hackenberg, and Thomas C. Wilson is in preparation. The working title is "Reference Expert: An Expert System for Print and CD-ROM Reference Sources."
10. James Parrott's work helped stimulate the thinking of the Knowledge Engineering Group about the structure of the knowledge base. See James R. Parrott, "Simulation of the Reference Process, Part II: REFSIM, an Implementation with Expert System and ICAI Modes," *The Reference Librarian*, no.23:153–76 (1989).
11. For further information about knowledge representation using frames, see Kamran Parsaye and Mark Chignell, *Expert Systems for Experts* (New York: Wiley, 1988), p.161–210; and John R. Walters and Norman R. Nielsen, *Crafting Knowledge-Based Systems: Expert Systems Made Easy/Realistic* (New York: Wiley, 1988), p.209–52.
12. *Expert Systems in ARL Libraries*, SPEC Kit 174 (Washington, D.C.: Office of Management Services, Association of Research Libraries, 1991), p.2 of the flyer. Also available in ERIC (ED 337 178).
13. Charles W. Bailey, Jr., "Intelligent Library Systems: Artificial Intelligence Technology and Library Automation Systems," *Advances*

in Library Automation and Networking 4:11 (1991).

14. To obtain a copy of the expert system, send a stamped, self-addressed diskette mailer and a formatted 5 1/4-inch 1.2 MB or 3 1/2-inch 1.44 MB diskette to: Charles W. Bailey, Jr., Assistant

Director for Systems, University Libraries, University of Houston, Houston, TX 77204-2091. If you reside outside of the U.S. and cannot get appropriate stamps, enclose $2 in your local currency to cover postage costs.   ■ ■

Arabic Online Catalog

Zahiruddin Khurshid

The article provides background information on the processing of Arabic materials using a combination of local and modified cataloging rules and the creation of the Arabic card catalog at the King Fahd University of Petroleum and Minerals Library (KFUPM). It also gives a brief history of KFUPM library automation and then presents various options considered for developing the Arabized version of DOBIS/LIBIS. Finally, the functions and features of the Arabic online catalog are described.

King Fahd University of Petroleum and Minerals (KFUPM, formerly UPM) was founded as a college in 1963. The status of the college was later changed to a university in 1975. KFUPM provides advanced training of students in the fields of science and engineering to prepare them for service and leadership in the Kingdom's petroleum and mineral industries.[1]

The academic programs of the university are well supported by a central library with a strong collection of more than 250,000 volumes. Because the university's main focus is on study and teaching in scientific and technical areas, the library's collection is comprised mostly of non-Arabic materials. Only about 7.5 percent of its collection is in Arabic, most of which supports the Islamic and Arabic studies programs.

PROCESSING OF ARABIC MATERIALS

The Arabic collection of the library received less attention than others in both develop-

Zahiruddin Khurshid is Manager, Cataloging Operations Division, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia.

ment and processing. Until 1976, there was only one person with a professional degree in library science, responsible for acquisitions, bibliographic control, shipment clearance, and all paperwork regarding Arabic materials. For processing, a brief description of books was provided on cards without much attention to cataloging rules. For classification, a temporary modified Dewey Decimal Classification scheme was used, according to which a book was assigned a call number composed of a general number for the class, followed by a slash and an accession number.[2]

It was soon realized that the local system for processing Arabic materials was creating access problems. For example, a second copy of a title would appear with a call number different from the first copy, whereas two separate books were at times assigned the same call number. A decision was made therefore to shift from the local system to a standard cataloging system using *AACR*, Library of Congress (LC) Subject Headings, and the LC Classification Scheme. The idea was also to make use of the LC card sets.

To save effort and time in consulting two separate catalogs for Arabic and non-Arabic materials, integration of the two was considered necessary. The transliterated LC cards enabled us to interfile them with the non-Arabic catalog cards. However, to satisfy the library patrons, who still preferred access to the collection through the Arabic alphabet, an Arabic title file was provided wherein one card in every card set was arranged alphabetically by Arabic title. On the other hand, the Arabic collection was also integrated with the non-Arabic collection so that the readers could browse the library's holdings in their subject of interest in one area of the stacks.

The decision to adopt a transliterated system for Arabic material was taken also in view of a growing backlog of Arabic books for cataloging and the shortage of Arabic catalogers. The idea was to make use of the LC catalog

records and thus reduce the amount of original cataloging. This decision ran against the general feeling of the Arabic speakers who take pride in their native language and show strong opposition to subordinating Arabic to another language in bibliographic records. It was observed that fewer and fewer people were using the card catalog. They resented this scheme and were not willing to learn it.

In 1979, the policy was modified to drop transliteration practice in favor of vernacular script records, except for subject heading and class number, which remained in English. The main entry and added entries except for subject added entries and the shelflist cards were filed in a separate Arabic catalog. The reasons to continue to use the LC Subject Headings for Arabic materials and to have an integrated subject file for both Arabic and non-Arabic materials were: (1) ease of locating an Arabic book on a specific subject with related non-Arabic books in the same file; (2) the unavailability of a reliable, periodically updated Arabic subject headings list; (3) the fact that the majority of KFUPM Library holdings are in English and it is a must to use the LC Subject Headings list; and (4) earlier attempts at translating scientific and technical headings into Arabic have proved disastrous.

LIBRARY AUTOMATION

Planning for library automation at the KFUPM library began in 1975, when five systems were investigated by the library administration. They were DOBIS, CLSI, BALLOTS, BATAB, and Hewlett Packard's 3000 System. One of the features that KFUPM wanted in a library automation system was the multiple language capability, which DOBIS/LIBIS had only for the Latin script languages, German, French, and Italian. We finally selected DOBIS/LIBIS in June 1980 with a long-term objective to modify it to handle both Latin and Arabic scripts. Modifying a program and adapting it to the local requirement was considered a better choice than developing an Arabic program from scratch. This was the time when even the Library of Congress was seriously considering complete romanization of its bibliographic records for all nonroman scripts except Chinese, Japanese, and Korean. One of the reasons for adopting this approach was that the software required for interfacing nonroman characters with the library's automated sys-

tems was not available and would entail great expense to develop.[4]

DEVELOPMENT OF THE ARABIZED VERSION OF DOBIS

Following the completion of the implementation of DOBIS/LIBIS for handling non-Arabic materials in 1986, the attention was focused on handling the Arabic script materials using DOBIS. We considered the following three options:

1. Creating records for Arabic documents in romanized form using the English version of DOBIS;

2. Entering the Arabic script data using the English version of DOBIS; or

3. Developing an Arabized version of DOBIS for handling bibliographic information in the Arabic script and keeping the Arabic files separate from the English files.

Option number one was rejected for two main reasons. First, no acceptable standard exists for romanization that has resulted in transliterating a single word into several different ways. Second, romanization is not acceptable to the users.

The second option was rejected for two reasons. First, the orientation for entering Arabic and non-Arabic data is different. Whereas the Arabic data are entered from right to left, the non-Arabic data are entered from left to right. Second, if the Arabic data are integrated with the non-Arabic data, they would confuse the user in both searching and displaying the bibliographic information.

In view of the serious problems with the first two options, the third option was found to be the most appropriate. Thus, to prepare an Arabized program, the following steps were taken:

1. Adopted ASMO 449, the standard coding system for the Arabic language developed by the Arab Organization for Standardization and Metrology (ASMO) to translate Arabic information consisting of letters and numbers to equivalent binary numbers. ASMO 449 is based on the 7-bit coding table in ISO 646 adapted for the specific requirements of Arabic and contains 128 characters (inclusive of controls, graphic symbols, and alphanumerics) with their coded representation[5];

2. Acquired an IBM X-BASIC terminal, 3192, which is most suitable to handle both English and Arabic characters and is com-