


Database Selection: One Size Does Not Fit All

DeeAnn Allison, Beth McNeil, and Signe Swanson

With increasing costs and decreasing budgets, it is important for librarians to make sound purchasing decisions for electronic resources. What factors are important to consider in making a decision? How can librarians balance these factors, which may appear to be contradictory, and also meet the expectations of administrators, staff, and users? This article describes a strategy for making delivery decisions that address local conditions, pricing, feature options, hardware costs, and network availability. Finally, a model is presented for evaluating the decision based on dollar requirements and local issues.

ibraries share the common goals of enhancing learning and ensuring access to information. They also assist people in obtaining high-quality information that librarians have reviewed for accuracy. However, with the advent of electronic resources, these goals have become increasingly difficult for libraries to fulfill. Libraries now must be able to successfully determine the most cost-effective way to deliver information, while competing with the Internet and other nonlibrary information providers.

Over the past five years, electronic resources have undergone significant technological changes. Graphical user interfaces and the evolution of the Internet have created new options and complexities, as well as technologically savvy users who demand improved access. Determining the optimal delivery method for

these electronic resources can be difficult. Libraries must use cost-measurement techniques to evaluate alternatives or risk losing budget and staff support because libraries are judged as not meeting user needs. Local conditions, pricing, feature options, hardware costs, and network availability should be considered in the selection process.

At the University of Nebraska-Lincoln (UNL) Libraries a project was undertaken to replace the popular, but technologically obsolete, DOS-based database system with newer technology based on World Wide Web access. Although the available databases did not change significantly, the switch to the new system resulted in heavier use. For example, one of the databases, PsycLIT, dramatically increased from an annual access of 23,000 connections to more than 74,000 in just three and a half months. This article discusses the

DeeAnn Allison is a Professor and Chair of Automated Systems at the University of Nebraska-Lincoln; e-mail: decanna@unllib.unl.edu. Beth McNeil is an Assistant Professor and Executive Assistant to the Dean of Libraries at the University of Nebraska-Lincoln; e-mail: bethm@unllib.unl.edu. Signe Swanson is an Assistant Professor and Reference Librarian at the University of Nebraska-Lincoln; e-mail: signes@unllib.unl.edu.

necessary questions and criteria to consider when selecting a delivery method for electronic resources, with examples from the UNL libraries' experience. It also presents a model for assessing that decision, keeping in mind that no one answer fits all situations.

Selection Criteria

Database pricing is based on a variety of formulas determined by the vendor and/or copyright holder. Many factors complicate the pricing of electronic resources, the most basic being the delivery method. Vendors' and libraries' definitions of delivery methods vary. For instance, connections can be limited to specific buildings or geographic areas, and some databases may or may not be Internet accessible. It is important that library and vendor have a mutual understanding of these definitions when entering into licensing/contract negotiations.

For the purpose of this article, the following delivery method definitions are used throughout. A *non-networked computer* is one that is not connected to a network. It has either CD-ROM or hard-drive accessible database(s), and only one person may access any of the databases at a time. A *local area network* (LAN) is a system that includes a local server capable of handling multi-users in a multi-tasking mode over a locally or remotely accessible network, which may include Internet access. Databases are loaded on the local server, and the library is responsible for maintenance and network connectivity. A *remote Internet connection* provides access to databases that are housed on a vendor's remote server and delivered over the Internet. All of the database maintenance is carried out by the vendor, with the local institution responsible only for the connection to the vendor.

In addition to delivery method, pricing factors include the number of concurrent users/ports/passwords and other options. Pricing usually varies among different vendors providing the same database. Some vendors offer subscription pricing for titles "bundled" together, such

as OCLC's FirstSearch database. One question to ask these vendors is whether the library can determine the titles in the bundled package. Some databases have differentiated pricing based on whether the library also subscribes to the print version of the title. Other resources may have additional charges for outputs, such as printing or downloading. Licensing requirements or consortial agreements also may affect the overall charges to an individual library for a particular resource.

Outsourcing is a possibility, but it will be important to have good response time for maintenance and troubleshooting.

The American Psychological Association's (APA) electronic databases provide a good example to demonstrate the pricing situation. APA publishes *Psychological Abstracts*, the print index for the field of psychology. PsycINFO, PsycLIT, PsycFIRST, and ClinPSYC are electronic variations of *Psychological Abstracts*. Each version is available from several vendors.¹ For example, as of June 1999, PsycINFO was available as an online service from CompuServe IQuest, DataStar, DIALOG, DIMDI, HealthGate, and Ovid. Institutional leases for PsycINFO were available directly from APA, as well as from Aries Systems Corporation, Cambridge Scientific Abstracts, EBSCO Publishing, NISC, OCLC, Ovid Technologies, SilverPlatter Information, The Gale Group, and UMI. PsycLIT, a PsycINFO subset, is available only on CD-ROM from Aries Systems Corporation, EBSCO, NISC, Ovid Technologies, and SilverPlatter Information. The electronic versions of *Psychological Abstracts* provide a complex pricing model that varies depending on database vendor, delivery method, and content. Any of these may be the best choice for a particular library.

Functionality is another important consideration when selecting a database delivery method. Electronic databases usually offer several levels of functionality,

TABLE 1
Database Features: Output Options

Features	Specific Considerations
Displaying	What is default display (citation, abstract, full-text)? Can default be changed? Is scrolling through results possible from both citation list and record by record? Are search terms highlighted in results?
Printing	Can parts/sections of document/page be printed? Can print limit (number of pages, records, etc.) be set?
Downloading	Can saving be done in various file formats (ASCII, HTML, etc.)? Can results be saved and imported into bibliographic software program?
E-mailing	Can an entire marked list of articles/citations be e-mailed? Can both ASCII files and attached image files be sent?
Customization (of all output formats)	Can output be sorted (by date, author, publication, etc.)? Can one select certain fields for output other than default? Can length-of-field names be changed? Is search history included?
Document delivery	Can requests be made directly from the database (through either ILL or a fee-based service)?

ranging from base features that are built into the system to optional value-added enhancements, which may come at additional cost. Database features and options also differ among vendors, and one vendor may offer different nonnetworked, LAN, and remote Internet access versions of the same database.

Before selecting a delivery option, libraries should determine what database features are necessary for successful implementation. The inclusion of a particular database feature can never be assumed. For example, if a library wants to focus on e-mailing search results as a way to cut down on local printing costs, it is important to verify that a database supports e-mail. The following list describes database/interface features, which should be identified as either requirements or preferences for the local situation. Requirements are "must-have" features needed to satisfy user needs, while preferences are desirable, but less necessary, options. The individual features can be grouped into five categories: output options, search interface, search options, content, and local issues (see tables 1-5).

As the UNL libraries LAN server neared obsolescence, plans were made for migrating the DOS-based CD-ROM databases to new delivery methods. Of particular concern was finding a successful replacement for the heavily used PsycLIT database. The current PsycLIT SilverPlatter CD was assessed for both positive and negative features, and ideal database features for the new system were identified. Aside from pricing, hardware, and network concerns, feature requirements played an important role in influencing the final selection.

Besides licensing considerations and local requirements, the library's local situation may limit the options for delivery of electronic resources. For example, both LAN and remote access require reliable networks. The network is the wire that connects the computer to the server or the Internet and should be of sufficient speed to support graphical user interface connections. In addition, there should be minimum network downtime when relying on the network for access to important electronic databases. Moreover, for a LAN, it is necessary to have or be able to hire competent staff to set up and maintain the network and servers. Outsourcing is a possi-

TABLE 2
Database Features: Search Interface

Features	Specific Considerations
Stability	Is it a fairly static product that undergoes little change in interface?
Differences between versions (nonnetworked, LAN, Web, etc.)	Are there some things that one interface does that another does not? Can users move easily between them?
Interface customization	Can you determine which search options appear on the screen, and where? Can you enable/disable selected options? Can you add local information to the search screen(s)? Can you select the default search type (e.g., keyword vs. subject)? Can you set up different interfaces for different user groups?
Z30.50 availability	
Integration/uniformity	How important is it for the new product to look and feel similar to existing products? If the interfaces are not exactly identical, are things like truncation symbols and adjacency operators the same?
Help screens	Is help context sensitive? Is a searchable help index available?

bility, but it will be important to have good response time for maintenance and troubleshooting. In cases without a reliable network or adequate staffing, the only viable option will be databases on nonnetworked computers.

A final delivery method consideration is hardware. When planning for the development and implementation of new hardware, libraries should not anticipate technology more than five years ahead. For example, according to a research note published by Gartner Group in 1997, the technology for Web clients is predicted to change from simple browser access to more complicated hybrid clients, so it would seem unrealistic to plan beyond a five-year time frame.²

Nonnetworked computers have an initial start-up cost that is lower than the LAN cost. However, increased staff support requirements for maintaining multiple nonnetworked computers can require a large investment over a five-year time period. A LAN system where the databases are maintained on a single server can save from 30 to 40 percent over the cost of the similarly configured

nonnetworked setup where duplicate databases are running on multiple computers.

Although less expensive over time than nonnetworked computers, building a LAN still involves several cost considerations. After it is determined that the network is of sufficient reliability and staff is available for ongoing support, the system must be properly scaled to support the databases. Proper scaling will include determining the correct number of processors, hard-drive requirements, RAM requirements, and planning for upgrades over the next five years. The chart shown in table 6 addresses issues that are important in determining the size and future upgrade path for hardware.

One system that is commonly used to provide LAN access to electronic databases is Citrix WinFrame. WinFrame has been described in an article by Anna Turner as a "combination of high performance and low cost," which provides a single point of control for database maintenance with universal application access that is platform independent.³ The system has built-in support for Web access using

TABLE 3
Database Features: Searching Options

Features	Specific Considerations
Boolean and proximity operators	Can Boolean commands be typed in on the command line as well as chosen from a pull-down menu? Is "AND" assumed between search terms, or is phrase searching done automatically? Is nesting allowed? Can proximity operators be used to specify more specific relationships between search terms?
Field-specific searching	Is searching initiated through pull-down menus and/or from the command line?
Thesaurus/subject list	Does it include broader, narrower, and related terms? Can you select and execute searches directly from the thesaurus? Can you "explode" terms? Will the thesaurus suggest alternatives if your search matches nothing?
Lateral searching	Are search terms (subject headings, authors' names, etc.) hot-linked and/or searchable from the results list?
Limiting	Can limiting be done on the initial search, or do you have to limit search results? Are a variety of limit options available?
Saving/combining searches	Can you view your search history and combine and/or modify previous searches from the current search session or earlier?
Multiple search modes (basic, advanced, etc.)	Which type of search is the default? Is it difficult to distinguish or move between the two?
Multiple database searching	Can you select and search multiple databases simultaneously?
Multiple year searching	Can you search the current years and the backfile simultaneously? Do you need to select a range of years before starting the database? Can you easily switch between current and backfile?
System-assisted searching	Is weighted or relevance searching used? Is natural language querying available? Is SDI (selective dissemination of information) available? Will the system automatically look for variant spellings, plurals/singulars?

the ICA-client plug-in for Web browsers. In addition to Internet access, the system has the advantage of supporting Windows-based terminals and low-end networked computers. This can reduce the overhead for libraries that provide in-house access to electronic databases.

Using WinFrame version 1.8, a single processor (Pentium-based server) should support up to twenty concurrent users. The server requires the following RAM: 16 Mg minimum for the operating system, 4-8 Mg per user, with an additional 1 Mg for each configured WinStation. The operating system also requires 300 Mb of free disk space.

The UNL University Libraries selected and implemented the WinFrame solution for providing access to databases as a replacement for the DOS-based network. The system consists of a server farm of five multiprocessor Pentium servers, one of which is dedicated to providing Internet access to electronic databases. These databases are loaded from CD-ROMS onto hard drives on this server. A Web browser uses the ICA-client plug-in to launch a connection to the server that controls the databases. The end user is able to use the native vendor interface to access the databases. Keyserver software

TABLE 4
Database Features: Content

Features	Specific Considerations
Backfile availability	How far back does the product go?
Full-text	What is the percentage of full-text coverage? Is full-text inclusion an additional subscription? If not available, is it being considered? Can links be made from citations to full-text electronic journal subscriptions?
PDF (portable document file) format	Is PDF a viewing/output requirement or option in this product?
Graphics/images	Are graphics included? Are they part of a full-text document (as in a PDF file), or must they be handled separately?
Currency/updating	Does the updating frequency differ among the delivery methods (nonnetwork, LAN, etc.)?

is used to monitor licensing and gather statistics. The other four servers are used for in-house access to the databases and other applications.

Database Selection Model

Libraries are faced with the problem of balancing dollar requirements against local issues when selecting a delivery method. The following model conceptualizes the outcome of this decision-making

process (see table 7). Dollar requirements include the vendor's pricing model (simultaneous users, passwords, ports, etc.), access restrictions (no Internet access available, not networked, etc.), and additional charges, including add-on features. Also included as a cost is the one-time installation fee for local hardware. Local issues are those requirements deemed necessary for a successful local implementation.

TABLE 5
Database Features: Local Issues

Features	Specific Considerations
Links to local holdings	Can the database be linked to the OPAC? Can local holdings information simply be included in the database? Is this an additional cost?
ILL restrictions	Are there restrictions on lending copies of electronic full-text articles (that are not subscribed to in print)?
Authentication (remote access only)	Who is responsible for authenticating, the vendor or the local library? What is the method for authentication (proxy, passwords, etc.)?
Plug-ins/helper applications	Are extra pieces of software required for accessing the database? Will remote users require different software than local users?
Support/help desk hours	Is online or phone support available all hours the library is open? What is response turnaround time? Are both technical and search support provided?

TABLE 6
Upgrade Planning

Issue	Non-networked Computer	LAN
Total number of simultaneous users expected to use the system at the same time	One, unless additional computers/databases are owned	Scale server
Total number of concurrent users for each database	One, unless duplicate databases are owned	Scale server
Total users that will be added over the next five years	Add additional duplicate machines	Upgrade server
Application requirements	Select computer for the application	Request demo to verify that application will run on server
Memory requirements for OS and application	Base RAM plus what is needed for application	Base RAM plus what is needed for application
Disk space	Whatever system requires for OS and application	Whatever system requires for OS and application
Application requirements that are proprietary	Request demo or match system with application	Request demo to verify that application will run on server
Client requirements	Not applicable	Verify that system supports all appropriate client platforms. A client should require no more than 20 Kbps for good remote dial-in performance.

Outcomes on the left side of the model have low local value and represent risky positions that are most likely to be questioned. As a result, when a decision falls into this side of the model, it will be necessary to justify the decision to either frontline staff and/or administrators. Decisions on the right side of the model have high local value and are less likely to be questioned by frontline staff because they satisfy local needs. However, it is possible to overinvest in a high-value solution that precludes other purchases. These selections may be hard to justify to individuals with budgetary oversight. Outcomes fall into one of four categories: ideal, de-

fendable, inconsequential, or questionable.

Ideal

In an ideal situation, cost-effective investment is made in a database that provides a level of access and functionality to meet the needs of the local environment. In this situation, the library has carefully analyzed the database features, evaluated staffing support levels, and met budgetary constraints. The library has neither overinvested nor settled for less than is needed.

Defendable

In a defendable situation, the library is

TABLE 7
Decision-making Process for Database Selection

Database Selection Model		Local Issues	
		Low	High
Dollar Requirements	High	High Cost Low Value <i>Questionable</i>	High Cost High Value <i>Defendable</i>
	Low	Low Cost Low Value <i>Inconsequential</i>	Low Cost High Value <i>Ideal</i>

seem that compromising on local issues in order to get a lower price is beneficial, but users may not use the database or may use it inefficiently or ineffectively. Likewise, for locally mounted databases, staff time is wasted in maintaining underutilized databases.

Questionable

Questionable situations arise when decisions are made without regard to local requirements. In some situations, a decision may be made to purchase a high-cost database that does not

providing a high-end, high-cost solution to its electronic access needs. An example would be an expensive, full-text database that serves the library's needs but also includes features that are never used. In making this kind of selection, the library may be placed in a position to defend the decision because the high cost may have prevented purchase of other resources.

In some situations, a decision may be made to purchase a high-cost database that does not meet local needs and is not used.

Inconsequential

Inconsequential situations result from investment in a less expensive database that fails to meet local requirements. These decisions do not adversely affect the budget, but they also fail to serve user needs. It may

meet local needs and is not used. The library may be put in the position of defending the budgetary decision, particularly in cases where other purchases were sacrificed to purchase the costly database.

Conclusion

In conclusion, determining an ideal delivery method for electronic resources is a complicated process that involves many considerations. This article uses the UNL experience as one example of how this process is designed to work. Each library's situation is unique and one size will not fit all. However, when decisions on purchasing electronic resources are made with due consideration to local issues and dollar requirements, librarians can be assured that the decisions are fiscally responsible and defendable and meet the fundamental goal of providing access to information.

Notes

1. "Comparison Chart," in <http://www.apa.org/psycinfo/compare.html> (Washington, D.C.: American Psychological Association) [cited May 6, 1999].
2. Tom Austin, *Hybrid Turmoil and Vigor: Web and Traditional Clients* (Stamford, Conn.: GartnerGroup Advisory, 1997).
3. Anna Turner, "Thin Client Architecture for Networking CD-ROMs in a Medium-Sized Public Library System," *CD-ROM Librarian* 17 (Sept. 1997): 74.