Service Models, Operational Decisions and Architecture of Digital Libraries

Yuehong Yuan, Stephen Roehrig
The Heinz School of Public Policy &Management
Carnegie Mellon University
Pittsburgh, PA 15213, USA
Tel: 1+412-268-2560, 8783
E-mail: yy0c, roehrig+@andrew.cmu.edu

Marvin Sirbu
Dept. of Engineering and Public Policy
Carnegie Mellon University
Pittsburgh, PA 15213, USA
Tel: 1-412-268-3436
E-mail: sirbu+@cmu.edu

ABSTRACT
A digital library simplifies the interface between information users and producers; it economizes the costs of storage, search, and telecommunications in providing access to information. Externalities in information consumption and the cost structure of information access suggest five service models of digital libraries. Optimizing models are proposed to formalize operational decisions in these service models. A managing server implementing these decision models would automate the operations of digital libraries.

KEYWORDS: Service models, economic models, digital libraries

INTRODUCTION
A networked digital environment makes it possible for any individual to become a publisher and for any user to access each author's publication. However, the sheer numbers of publishers and users impose heavy burdens of navigation on users and marketing on publishers. Providing access to information incurs costs of telecommunication, query processing and storage. Neither publishers nor users are at their best in managing these costs. Thus, there is a need for an institution to simplify the interface between publishers and users and to manage the costs of information access. That institution is the digital library. Digital libraries can reduce users' burden of navigation because there will be far fewer libraries than publishers; likewise, they reduce publishers' burden of marketing because the number of libraries is much smaller than the number of users. Digital libraries are the best at managing the costs of information access because that is what they specialize in.

SERVICE MODELS OF DIGITAL LIBRARIES
Every institution uses some model of service. What are the service models of digital libraries? The service models of digital libraries depend on how information is consumed and provided.

Information can be used for decision-making, education, entertainment, and production of new information. There are important differences between these uses. Information used for entertainment, business decision-making, and commercial information production gives utility only to the user of information. Information used for basic research, public education, and personal decision-making which may have a larger social impact, yields benefits to the user as well as external benefits bestowed to the society or community at large. Similarly, job-related information consumption produces benefit to the organization, external to the individual employee. Finally, advertising can be viewed as a type of information product which benefits primarily the advertiser.

Basic economic principles predict that information for uses with only internal benefit can be efficiently provided by commercial means; information used for ends with external benefits may not be optimally provided by purely commercial means. Thus, their service models will be different.

Information typically incurs a large cost of first-copy production and small cost of copying and distribution. This cost structure requires copyright protection of limited length to encourage production and utilization. Thus, information may be either under protection or in the public domain.

The externality or internality of information consumption and the cost structure of information suggest the following service models of digital libraries:

1. *Commercial libraries* Copyrighted information used for entertainment, business decision-making and production of commercial information is provided on a commercial basis. Commercial libraries acquire, deposit and market publications from publishers, provide indexing, cataloging, telecommunication, and query processing services, and sell copies of information to users. Users download information to personal computers for their own use but are prohibited by copyright to transfer the protected information to others. The library may charge users based on usage and/or subscription. Library and publisher share the revenue from usage and/or subscription charges.

2. *Organizational libraries and public libraries* An organizational library rents or purchases publications from publishers. The publications are only accessible to users in the organization. Usage may be free. The rent or purchase price reflects the level of use or the number of users in the organization. Similarly, a public library rents or purchases publications from publishers. Usage is free. The rent or purchase price reflects usage or number of users in the community.

3. *Advertising libraries* An advertising library stores advertisers' commercials and provides users access to them.

Usage is free. It charges advertisers for keeping commercials in the library and each time a commercial is accessed by users.

4. *National library* The national library registers publications, preserves copies of all registered publications, maintains a master-catalog of all publications, and may provide to users and libraries publications which are in the public domain. The master catalog keeps information about all publications, such as title, category, copyrights, etc. The catalog is open to all users in the nation, but its reserve of protected publications is not. Its distribution of information in the public domain is either free or carries a charge reflecting only telecommunication and request processing costs.

5. *Personal libraries* Personal libraries are clients of other libraries. They request, deposit, manage and display information from libraries or directly from publishers, and contain pointers to items in other libraries.

OPERATIONAL DECISIONS OF THE LIBRARIES

Libraries face decisions about acquisition, storage organization, pricing and so forth. The operational decisions of commercial, advertising and organizational/public libraries are of particular interest. Optimizing models of operational decisions for these types are now proposed, under simplifying assumptions.

Commercial Library

Digital information is not a traditional product, and is different from even paper-based information. A copy sold to a library means a lot of copies may be sold to final users. Thus, the revenues of publishers can not be based on copies sold to libraries alone and are to be related to the numbers of copies sold to users by the libraries.

Commercial libraries charge publishers for storing information. These charges cover costs of storage and discourage publishers from storing excessive information. Libraries also charge publishers per purchase of the publishers' information by users. These running charges cover usage-sensitive costs of search and telecommunications, and give libraries incentives to market publishers' information.

Users are charged on per-usage basis. Part of the usage revenue goes to the library as storage and running charges. The rest is the publisher's revenue.

An equilibrium is reached if a storage charge and a running charge are found where the amount of information the publisher is willing to supply equals the amount the library is willing to acquire, and a usage price is agreed upon by both publisher and library.

The payment flows described above are very different from those in traditional libraries or bookstores. The closest current analogs are on-line service providers. In many cases, the publisher pays the service provider a fee for putting the information up on a server and even a usage fee for data processing service. In return, the publisher collects the full value of revenue charged to the user, from which these fees are deducted.

The above service model are formalized under simplifying assumptions: 1) the library is a monopoly; 2) the publisher is a monopoly; 3) selling through the library is the only distribution channel of the publisher; 4) all publications are of the same category; 5) the transaction period covers the entire copyrighted life of the information. It is possible to relax these assumptions, causing only complications in the algebra.

We use the following notation:

S: size of collection of library or supply from publisher;

R: storage charge per unit of information to publisher;

p_1: running charge per unit of usage to publisher;

p: usage charge;

L: number of mirrors the library keeps for its collection;

$d(S, p)$: users' demand, increasing in S and decreasing in p;

$c(S)$: cost for the publisher to produce S unit of information.

Assume costs of the library are as follows:

Storage cost:

$$C_1 = b_1 L S^{\alpha}$$

with $0< \alpha <1$, reflecting economy of scale in storage.

Searching cost:

$$C_2 = b_2 L \left[\frac{d(S,p)}{L}\right]^{\beta} S^{\gamma} = b_2 L^{1-\beta} d^{\beta}(S,p) S^{\gamma}$$

where demand for information is equally split among mirrors, and number of searches is proportional to information demand; $0< \beta <1$ represents economy of scale in search; $\gamma >0$ reflects increasing search difficulty with collection size.

Telecommunications cost is:

$$C_3 = b_3 L^{-\frac{1}{2}} d(S,p)$$

where the average distance from users to their respective mirror is inversely related to the square root of the number of mirrors.

The revenue of the library is $RS + p_1 d(S,p)$.

The profit is:

$$\Pi_1 = RS + p_1 d(S,p) - \left(b_1 L S^{\alpha} + b_2 L^{1-\beta} [d(S,p)]^{\beta} S^{\gamma} + b_3 L^{-\frac{1}{2}} d(S,p)\right)$$

Maximizing Π_1, assuming continuous values for S and L, first order conditions $\frac{\partial \Pi_1}{\partial S} = 0$, $\frac{\partial \Pi_1}{\partial p} = 0$ and $\frac{\partial \Pi_1}{\partial L} = 0$ give the following equations:

$$R + p_1 \frac{\partial d}{\partial S} = b_1 \alpha L S^{\alpha - 1} + b_2 \beta L^{1-\beta} d^{\beta - 1} \frac{\partial d}{\partial S} S^{\gamma} + b_2 \gamma L^{1-\beta} d^{\beta} S^{\gamma - 1} + b_3 L^{-\frac{1}{2}} \frac{\partial d}{\partial S} \quad (1)$$

$$p_1 \frac{\partial}{\partial p} d(S, p) - \left(b_2 \beta L^{1-\beta} d^{\beta - 1} \frac{\partial}{\partial p} d(S, p) S^{\gamma} + b_3 L^{-\frac{1}{2}} \frac{\partial d}{\partial p} \right) = 0 \quad (2)$$

$$b_1 S^{\alpha} + b_2 (1 - \beta) L^{-\beta} d^{\beta} (S, p) S^{\gamma} = \frac{1}{2} b_3 L^{-\frac{3}{2}} d(S, p) \quad (3)$$

The publisher's profit is:

$$\Pi_2 = (p - p_1) d(S, p) - c(S) - RS$$

Maximizing Π_2, first order conditions give the following equations:

$$\frac{p - p_1}{p} = \frac{1}{\varepsilon} \quad (4)$$

$$(p - p_1) \frac{\partial}{\partial S} d(S, p) = mc(S) + R \quad (5)$$

where $\varepsilon = -\frac{p}{d(S, p)} \frac{\partial}{\partial p} d(S, p)$ is the elasticity of users' demand for information. Given parameters b_1, b_2, b_3, α, β, γ, demand $d(S, p)$ and cost $c(S)$, equations (1)-(5) collectively solve R, p_1, p, S, L.

Assume $d(S, p) = d_o S^{\mu} \exp(-\eta p)$ and $c(S) = b_4 S^{\rho}$, where $\eta > 0$ is related to the elasticity of users' demand, ρ represents the change of marginal cost of information production. Let $b_1 = 4$, $b_2 = 0.01$, $b_3 = 0.5$, $b_4 = 5$, $\alpha = 0.8$, $\beta = 0.7$, $\gamma = 1.1$, $d_0 = 1000$, $\mu = 0.3$, $\eta = 0.3$, $\rho = 1.1$ be the base values of the parameters. At the base values, one gets the solution 3.77, 0.63, 3.96, 96, 1.02 for R, p_1, p, S and L, respectively.

We varied several parameters to examine their influence on the solution. We looked at the change in the size of the library collection with respect to the cost of information production, change in usage price with elasticity of user demand, and change in number of collection mirrors with change in telecommunications cost. Changing ρ and fixing other parameters at their base values, resolving equation (1)-(5) gives results shown in Figure 1. Figure 1 shows that the size of the library collection decreases with the cost of information production.

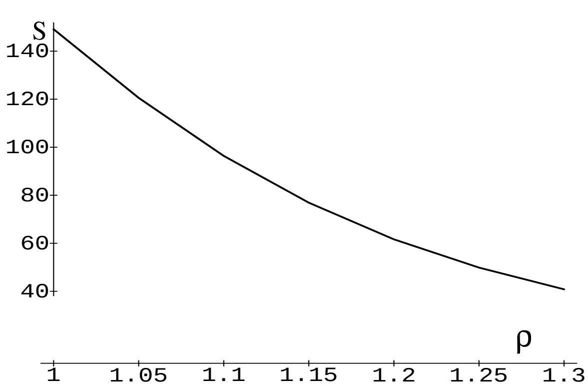

Figure 1: Library size and information production cost

Similarly, changing η and b_3, one gets Figure 2 and Figure 3, showing that the usage charge decreases when user's demand becomes more elastic and that the library creates more mirrors when telecommunication is more expensive.

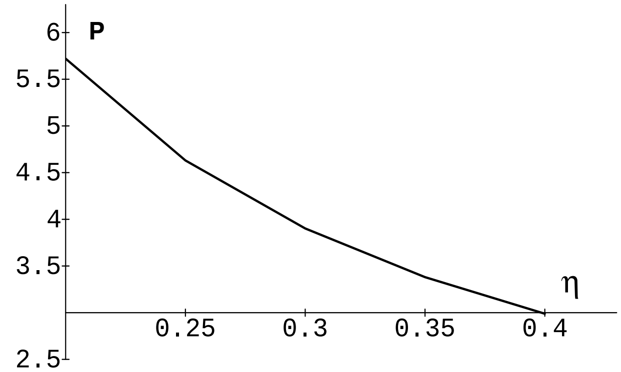

Figure 2: Usage charge and demand elasticity

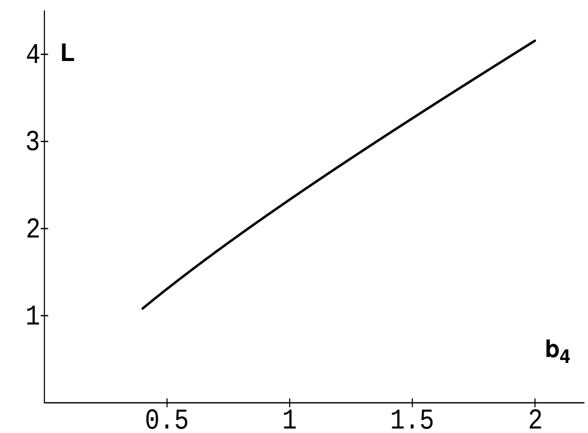

Figure 3: Mirrors and telecommunication cost

Advertising Library

Advertising libraries collect storage charges and running charges from advertisers. Usage of the library is free. Modify the earlier notation as follows:

p: running charge per unit of commercial downloaded by users;

$d(S)$: demand for commercials by users per period, increasing in S;

$D(R, p)$: Advertisers' demand to store commercials in the library, decreasing in R and p. $D(R, p)$ partially depends on $d(S)$.

The costs of an advertising library parallel those of commercial libraries. Total cost of storage, search and telecommunications is:

$$C = b_1 L S^{\alpha} + b_2 L^{1-\beta} [d(S)]^{\beta} S^{\gamma} + b_3 L^{-\frac{1}{2}} d(S)$$

The library's profit is:

$$\Pi = RS + pd(S) - \left(b_1 L S^{\alpha} + b_2 L^{1-\beta} [d(S)]^{\beta} S^{\gamma} + b_3 L^{-\frac{1}{2}} d(S) \right)$$

The advertising library determines the library size S, storage charge R and running charge p to maximize profit, that is:

$$\text{Max} \qquad \Pi = RS + pd(S) - \left(b_1 L S^{\alpha} + b_2 L^{1-\beta} [d(S)]^{\beta} S^{\gamma} + b_3 L^{-\frac{1}{2}} d(S) \right)$$

$$\text{S.t.} \quad S = D(R, P)$$

The first order conditions on the Lagrangian lead to the following equations:

$$\frac{R - b_1 \alpha L S^{\alpha-1} - b_2 \beta L^{1-\beta} d^{\beta-1} \frac{\partial d}{\partial S} S^{\gamma} - b_2 \gamma L^{1-\beta} d^{\beta} S^{\gamma-1} - b_3 L^{-\frac{1}{2}} \frac{\partial d}{\partial S}}{R} = \frac{1}{\xi^{D}{}_{R}} - \frac{p \frac{\partial}{\partial S} d}{R} \qquad (6)$$

$$\frac{p - b_1 \alpha L S^{\alpha-1} \frac{\partial S}{\partial d} - b_2 \beta L^{1-\beta} d^{\beta-1} S^{\gamma} - b_2 \gamma L^{1-\beta} d^{\beta} S^{\gamma-1} \frac{\partial S}{\partial d} - b_3 L^{-\frac{1}{2}}}{p} = \frac{1}{\xi_p \xi_S^d} - \frac{R \frac{\partial S}{\partial d}}{p} \qquad (7)$$

$$b_1 S^{\alpha} + b_2 (1-\beta) L^{-\beta} d^{\beta} (S, p) S^{\gamma} = \frac{1}{2} b_3 L^{-\frac{3}{2}} d(S, p) \qquad (8)$$

$$S = D(R, p) \qquad (9)$$

where $\xi^{D}{}_{R} = -\frac{R}{D} \frac{\partial}{\partial R} D(R, p)$ is the storage-charge elasticity of demand of advertisers; $\xi_p = -\frac{p}{D} \frac{\partial}{\partial p} D(R, p)$ is the running-charge elasticity of demand of advertisers; and $\xi_S^d = \frac{S}{d} \frac{\partial}{\partial S} d(S)$ is the library-size elasticity of users' demand. The product $\xi_p \xi_S^d = -\frac{p}{d} \frac{\partial}{\partial p} d(S)$ is the indirect elasticity of users' demand to running charge.

Given parameters b_1, b_2, b_3, α, β and γ, and demand functions *D(R, p)* and *d(S)*, one can solve for the four variables *R*, *p* and *S*, *L* from the four equations (6)-(9).

Organizational/Public Library

A key question of an organizational/public library is whether to charge for usage. Use of an organizational library serves organizational goals. Use of a public library benefits not only users themselves but also the community or society at large. It is difficult to measure these external benefits. Digital technology makes it easy to measure usage, but may not help to measure external benefits. A typical way to avoid the difficulty of measuring external benefits is simply not to charge for use.

Another problem of an organizational/public library is to determine collection size. The obvious rule is to require that marginal cost of acquiring and serving the collection equals the marginal benefit to the organization/community. Let *U(d(S))* be the total benefit of usage *d(S)* of collection S. Fixing the number of mirrors at one for simplicity, the cost of acquiring and serving *S* is:

$$C = pS + b_1 S^{\alpha} + b_2 d^{\beta}(S) S^{\gamma} + b_3 d(S)$$

The marginal cost is:

$$MC = p + b_1 \alpha S^{\alpha-1} + b_2 \beta [d(S)]^{\beta-1} S^{\gamma} \frac{\partial}{\partial S} d(S) + b_2 \gamma [d(S)]^{\beta} S^{\gamma-1} + b_3 \frac{\partial}{\partial S} d(S)$$

The marginal benefit to the organization/community from S is:

$$MU = \frac{\partial}{\partial d} U(d(S)) \frac{\partial}{\partial S} d(S)$$

From $MU = MC$, one has:

$$p = \frac{\partial d}{\partial S} \left[\frac{\partial U}{\partial d} - \left(b_1 \alpha S^{\alpha-1} \frac{\partial S}{\partial d} + b_2 \beta [d(S)]^{\beta-1} S^{\gamma} + b_2 \gamma [d(S)]^{\beta} S^{\gamma-1} \frac{\partial S}{\partial d} + b_3 \right) \right] \qquad (10)$$

This in effect gives demand S of the library at each p.

Equation (10) suggests that publishers could charge organization/public libraries by unit of information usage (p/$\frac{\partial d}{\partial S}$). The per-usage charge on the library would be related to the per-usage benefit ($\frac{\partial U}{\partial d}$) and the marginal cost of serving S.

Per-usage benefit to the organization/community is as difficult to measure as external benefit, because the per-usage benefit includes the external benefit. However, one can not avoid this difficulty in this case, because the acquisition charge can not be unilaterally determined by the library. Unlike the case of charging users, the market relationship between organizational/public libraries and publishers gives persistent incentive for the libraries to measure benefit.

Because the usage per unit of collection ($\frac{\partial d}{\partial S}$) is related to the number of users in the organization/community, Equation (10) suggests that publishers may alternatively charge libraries by the number of users in the organization/community.

IMPLICATIONS FOR THE TECHNICAL ARCHITECTURE

The technical architecture of a digital library includes specifications about the client (browser), communication protocols, and servers for access control, indexing and searching, depositing and retrieving, billing and so on. What are the implications of service models and the formalization of operational decision-making on the architecture?

Above analyses show that service models of digital libraries can be identified. Their operational decisions can be formalized. The optimizing models studied are examples of such formalization under simplifying assumptions. However, extension and refinement of these models for real decision-making are possible. Furthermore, the operation of a digital library will generate statistics on usage, pricing and acquisition etc. These statistics can be used to estimate demand and supply of users and publishers. Therefore, it seems feasible and beneficial to implement decision models into the architecture of digital libraries as a managing server. The managing server would make use of the operational statistics of the library, estimate relevant parameters, automate some operational decisions and assist the librarian on others. For example, it could 1) decide automatically the optimum number of mirrors for each document; (2) compute and suggest optimal responses of demand for publications and usage pricing to the publisher's proposal of storage charge and running charge, in the case of the commercial library; (3) in the case of the advertising library, estimate publishers' demand for advertising and users' demand for commercials and compute optimal storage and running charges.

Different service models have different decision models. A general managing server could implement several models.

One specific library would configure the server to the intended model in which it operates.

In order to retain flexibility for refinement and changes in the decision models, the managing server would be built as modularly as possible. Particularly, the modularity would allow the library administrator to adjust decision variables, objective functions etc.

A managing server would improve decision-making and operations of the digital library and reduce the requirement for human intervention. For digital libraries, as the next stage in the automation of information access, to include managing servers to automate operating decisions seems a natural step.

CONCLUSIONS

Digital libraries are not only technical facilities but also economic enterprises. As economic enterprises, digital libraries simplify the interface between information consumers and information producers, and relieve the users' burden of navigation and the publishers' burden of marketing. They economize the costs of storage, search and telecommunication in providing access to information.

Properties of information consumption, namely internality and externality, and features of information cost structure suggest that commercial, organizational/public and advertising libraries and the national library are likely service models.

Commercial libraries and advertising libraries face decisions of acquisition, mirroring and pricing; organizational and public libraries face decisions of acquiring and mirroring. We have modeled and analyzed these decisions under simplifying assumptions. A distinct feature of the economics of a commercial library is that publishers can be isolated from final users; usage price has to be agreed upon by both the publisher and the library.

One implication of the service models and formalization of operational decisions is the feasibility and usefulness of including a managing server in the technical architecture of digital libraries. Based on models of operating decisions, statistics of past usage, acquisitions and prices, and other available parameters, a managing server may automate and assist decision-making about acquisition, mirroring, and pricing library services, thus improving the efficiency of the digital library.

REFERENCES

1. Baumol, W. J. and M. Marcus (1973): *Economics of Academic Libraries, American Council on Education*, Washington, D.C.

2. Bookstein, A. (1984): "An Economic Model of Library Service", In S. A. Roberts (ed.) *Costing and the Economics of Library and Information Services*, Aslib, London.

3. Casper, C. A. (1979): "Pricing Policy for Library Services", *J. of the American Society for Information Science*, September, 1979.

4. Davis, J. R. and C. Lagoze (1994). "A Protocol and Server for a Distributed Digital Technical Report Library", Cornell University Computer Science Dept. Technical Report 94-1418, June 1994.

5. Fox, E. (ed.) (1993): *The Source Book of Digital Library,* Dept. of Computer Science, Virginia Tech University, TR-93-35.

6. Gladney, H. M., N. J. Belkin, Z. Ahmed, E. Fox, R. Ashany and M. Zemankova (1994): "Digital Library: Gross Structure and Requirements (Report from a Workshop)", IBM Research Report RJ 9840, May 1994. (Available at ftp://fox.cs.vt.edu/pub/DigitalLibrary/RJ9840.ps)

7. Kahn, R. and R. Wilensky (1995): "Accessing Digital Library Services and Objects: A Frame of Reference", Draft 4.4. (Available at http://www.cnri.reston.va.us/home/cstr/arch/k-w.html.)

8. Knight, D. M. and E. S. Nourse (ed.) (1969): *Library at Large: Tradition, Innovation and the National Interest*. R. R, Bowker Company, New York and London.

9. Lunin, L. F and E. Fox (1993), "Perspective on Digital Libraries: Introduction and Overview", *J. of the American Society for Information Science*, Vol. 44, No. 8, Sep, 1993, 440 - 445.

10. Zahray, A. P. and M. Sirbu (1991), "The provision of scholarly journals by libraries via electronic technologies: an economic analysis", *J. of Information Economics and Policy*, No.4, 127-154.

Service Models, Operational Decisions and Architecture of Digital Libraries

Yuehong Yuan, Stephen Roehrig
The Heinz School of Public Policy &Management
Carnegie Mellon University
Pittsburgh, PA 15213, USA
Tel: 1+412-268-2560, 8783
E-mail: yy0c, roehrig+@andrew.cmu.edu

Marvin Sirbu
Dept. of Engineering and Public Policy
Carnegie Mellon University
Pittsburgh, PA 15213, USA
Tel: 1-412-268-3436
E-mail: sirbu+@cmu.edu

ABSTRACT
A digital library simplifies the interface between information users and producers; it economizes the costs of storage, search, and telecommunications in providing access to information. Externalities in information consumption and the cost structure of information access suggest five service models of digital libraries. Optimizing models are proposed to formalize operational decisions in these service models. A managing server implementing these decision models would automate the operations of digital libraries.

KEYWORDS: Service models, economic models, digital libraries

INTRODUCTION
A networked digital environment makes it possible for any individual to become a publisher and for any user to access each author's publication. However, the sheer numbers of publishers and users impose heavy burdens of navigation on users and marketing on publishers. Providing access to information incurs costs of telecommunication, query processing and storage. Neither publishers nor users are at their best in managing these costs. Thus, there is a need for an institution to simplify the interface between publishers and users and to manage the costs of information access. That institution is the digital library. Digital libraries can reduce users' burden of navigation because there will be far fewer libraries than publishers; likewise, they reduce publishers' burden of marketing because the number of libraries is much smaller than the number of users. Digital libraries are the best at managing the costs of information access because that is what they specialize in.

SERVICE MODELS OF DIGITAL LIBRARIES
Every institution uses some model of service. What are the service models of digital libraries? The service models of digital libraries depend on how information is consumed and provided.

Information can be used for decision-making, education, entertainment, and production of new information. There are important differences between these uses. Information used for entertainment, business decision-making, and commercial information production gives utility only to the user of information. Information used for basic research, public education, and personal decision-making which may have a larger social impact, yields benefits to the user as well as external benefits bestowed to the society or community at large. Similarly, job-related information consumption produces benefit to the organization, external to the individual employee. Finally, advertising can be viewed as a type of information product which benefits primarily the advertiser.

Basic economic principles predict that information for uses with only internal benefit can be efficiently provided by commercial means; information used for ends with external benefits may not be optimally provided by purely commercial means. Thus, their service models will be different.

Information typically incurs a large cost of first-copy production and small cost of copying and distribution. This cost structure requires copyright protection of limited length to encourage production and utilization. Thus, information may be either under protection or in the public domain.

The externality or internality of information consumption and the cost structure of information suggest the following service models of digital libraries:

1. *Commercial libraries* Copyrighted information used for entertainment, business decision-making and production of commercial information is provided on a commercial basis. Commercial libraries acquire, deposit and market publications from publishers, provide indexing, cataloging, telecommunication, and query processing services, and sell copies of information to users. Users download information to personal computers for their own use but are prohibited by copyright to transfer the protected information to others. The library may charge users based on usage and/or subscription. Library and publisher share the revenue from usage and/or subscription charges.

2. *Organizational libraries and public libraries* An organizational library rents or purchases publications from publishers. The publications are only accessible to users in the organization. Usage may be free. The rent or purchase price reflects the level of use or the number of users in the organization. Similarly, a public library rents or purchases publications from publishers. Usage is free. The rent or purchase price reflects usage or number of users in the community.

3. *Advertising libraries* An advertising library stores advertisers' commercials and provides users access to them.

Usage is free. It charges advertisers for keeping commercials in the library and each time a commercial is accessed by users.

4. *National library* The national library registers publications, preserves copies of all registered publications, maintains a master-catalog of all publications, and may provide to users and libraries publications which are in the public domain. The master catalog keeps information about all publications, such as title, category, copyrights, etc. The catalog is open to all users in the nation, but its reserve of protected publications is not. Its distribution of information in the public domain is either free or carries a charge reflecting only telecommunication and request processing costs.

5. *Personal libraries* Personal libraries are clients of other libraries. They request, deposit, manage and display information from libraries or directly from publishers, and contain pointers to items in other libraries.

OPERATIONAL DECISIONS OF THE LIBRARIES

Libraries face decisions about acquisition, storage organization, pricing and so forth. The operational decisions of commercial, advertising and organizational/public libraries are of particular interest. Optimizing models of operational decisions for these types are now proposed, under simplifying assumptions.

Commercial Library

Digital information is not a traditional product, and is different from even paper-based information. A copy sold to a library means a lot of copies may be sold to final users. Thus, the revenues of publishers can not be based on copies sold to libraries alone and are to be related to the numbers of copies sold to users by the libraries.

Commercial libraries charge publishers for storing information. These charges cover costs of storage and discourage publishers from storing excessive information. Libraries also charge publishers per purchase of the publishers' information by users. These running charges cover usage-sensitive costs of search and telecommunications, and give libraries incentives to market publishers' information.

Users are charged on per-usage basis. Part of the usage revenue goes to the library as storage and running charges. The rest is the publisher's revenue.

An equilibrium is reached if a storage charge and a running charge are found where the amount of information the publisher is willing to supply equals the amount the library is willing to acquire, and a usage price is agreed upon by both publisher and library.

The payment flows described above are very different from those in traditional libraries or bookstores. The closest current analogs are on-line service providers. In many cases, the publisher pays the service provider a fee for putting the information up on a server and even a usage fee for data processing service. In return, the publisher collects the full value of revenue charged to the user, from which these fees are deducted.

The above service model are formalized under simplifying assumptions: 1) the library is a monopoly; 2) the publisher is a monopoly; 3) selling through the library is the only distribution channel of the publisher; 4) all publications are of the same category; 5) the transaction period covers the entire copyrighted life of the information. It is possible to relax these assumptions, causing only complications in the algebra.

We use the following notation:

S: size of collection of library or supply from publisher;

R: storage charge per unit of information to publisher;

p_1: running charge per unit of usage to publisher;

p: usage charge;

L: number of mirrors the library keeps for its collection;

$d(S, p)$: users' demand, increasing in S and decreasing in p;

$c(S)$: cost for the publisher to produce S unit of information.

Assume costs of the library are as follows:

Storage cost:

$$C_1 = b_1 L S^{\alpha}$$

with $0< \alpha <1$, reflecting economy of scale in storage.

Searching cost:

$$C_2 = b_2 L \left[\frac{d(S,p)}{L}\right]^{\beta} S^{\gamma} = b_2 L^{1-\beta} d^{\beta}(S,p)\, S^{\gamma}$$

where demand for information is equally split among mirrors, and number of searches is proportional to information demand; $0< \beta <1$ represents economy of scale in search; $\gamma >0$ reflects increasing search difficulty with collection size.

Telecommunications cost is:

$$C_3 = b_3 L^{-\frac{1}{2}} d(S,p)$$

where the average distance from users to their respective mirror is inversely related to the square root of the number of mirrors.

The revenue of the library is $RS + p_1 d(S,p)$.

The profit is:

$$\Pi_1 = RS + p_1 d(S,p) - \left(b_1 L S^{\alpha} + b_2 L^{1-\beta} [d(S,p)]^{\beta} S^{\gamma} + b_3 L^{-\frac{1}{2}} d(S,p) \right)$$

Maximizing Π_1, assuming continuous values for S and L, first order conditions $\frac{\partial \Pi_1}{\partial S} = 0$, $\frac{\partial \Pi_1}{\partial p} = 0$ and $\frac{\partial \Pi_1}{\partial L} = 0$ give the following equations:

$$R + p_1 \frac{\partial d}{\partial S} = b_1 \alpha L S^{\alpha-1} + b_2 \beta L^{1-\beta} d^{\beta-1} \frac{\partial d}{\partial S} S^{\gamma} + b_2 \gamma L^{1-\beta} d^{\beta} S^{\gamma-1} + b_3 L^{-\frac{1}{2}} \frac{\partial d}{\partial S} \quad (1)$$

$$p_1 \frac{\partial}{\partial p} d(S,p) - \left(b_2 \beta L^{1-\beta} d^{\beta-1} \frac{\partial}{\partial p} d(S,p) S^{\gamma} + b_3 L^{-\frac{1}{2}} \frac{\partial d}{\partial p} \right) = 0 \quad (2)$$

$$b_1 S^{\alpha} + b_2 (1-\beta) L^{-\beta} d^{\beta}(S,p) S^{\gamma} = \frac{1}{2} b_3 L^{-\frac{3}{2}} d(S,p) \quad (3)$$

The publisher's profit is:

$$\Pi_2 = (p - p_1)\, d(S,p) - c(S) - RS$$

Maximizing Π_2, first order conditions give the following equations:

$$\frac{p - p_1}{p} = \frac{1}{\varepsilon} \quad (4)$$

$$(p - p_1) \frac{\partial}{\partial S} d(S,p) = mc(S) + R \quad (5)$$

where $\varepsilon = -\frac{p}{d(S,p)} \frac{\partial}{\partial p} d(S,p)$ is the elasticity of users' demand for information. Given parameters b_1, b_2, b_3, α, β, γ, demand $d(S, p)$ and cost $c(S)$, equations (1)-(5) collectively solve R, p_1, p, S, L.

Assume $d(S,p)) = d_o S^{\mu} \exp(-\eta p)$ and $c(S) = b_4 S^{\rho}$, where $\eta > 0$ is related to the elasticity of users' demand, ρ represents the change of marginal cost of information production. Let $b_1=4$, $b_2=0.01$, $b_3=0.5$, $b_4=5$, $\alpha=0.8$, $\beta=0.7$, $\gamma=1.1$, $d_0=1000$, $\mu=0.3$, $\eta=0.3$, $\rho=1.1$ be the base values of the parameters. At the base values, one gets the solution 3.77, 0.63, 3.96, 96, 1.02 for R, p_1, p, S and L, respectively.

We varied several parameters to examine their influence on the solution. We looked at the change in the size of the library collection with respect to the cost of information production, change in usage price with elasticity of user demand, and change in number of collection mirrors with change in telecommunications cost. Changing ρ and fixing other parameters at their base values, resolving equation (1)-(5) gives results shown in Figure 1. Figure 1 shows that the size of the library collection decreases with the cost of information production.

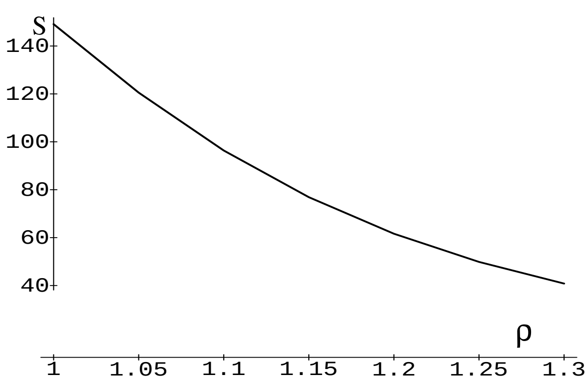

Figure 1: Library size and information production cost

Similarly, changing η and b_3, one gets Figure 2 and Figure 3, showing that the usage charge decreases when user's demand becomes more elastic and that the library creates more mirrors when telecommunication is more expensive.

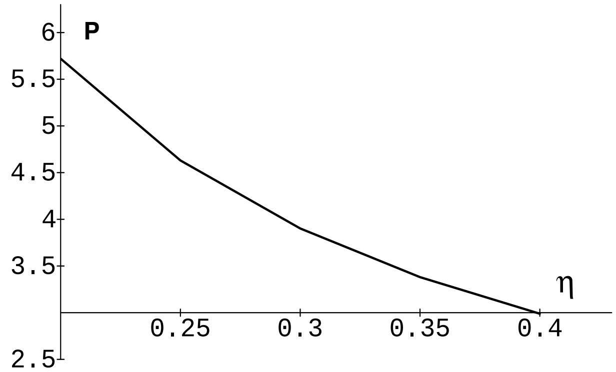

Figure 2: Usage charge and demand elasticity

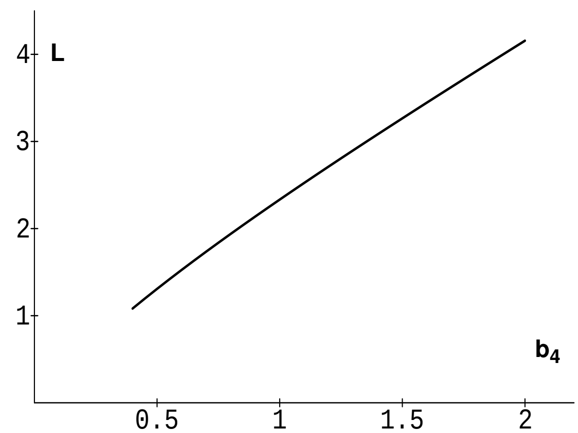

Figure 3: Mirrors and telecommunication cost

Advertising Library

Advertising libraries collect storage charges and running charges from advertisers. Usage of the library is free. Modify the earlier notation as follows:

p: running charge per unit of commercial downloaded by users;

$d(S)$: demand for commercials by users per period, increasing in S;

$D(R, p)$: Advertisers' demand to store commercials in the library, decreasing in R and p. $D(R, p)$ partially depends on $d(S)$.

The costs of an advertising library parallel those of commercial libraries. Total cost of storage, search and telecommunications is:

$$C = b_1 L S^{\alpha} + b_2 L^{1-\beta} [d(S)]^{\beta} S^{\gamma} + b_3 L^{-\frac{1}{2}} d(S)$$

The library's profit is:

$$\Pi = RS + pd(S) - \left(b_1 L S^{\alpha} + b_2 L^{1-\beta} [d(S)]^{\beta} S^{\gamma} + b_3 L^{-\frac{1}{2}} d(S) \right)$$

The advertising library determines the library size S, storage charge R and running charge p to maximize profit, that is:

$$\text{Max} \quad \Pi = RS + pd(S) - \left(b_1 LS^{\alpha} + b_2 L^{1-\beta} [d(S)]^{\beta} S^{\gamma} + b_3 L^{-\frac{1}{2}} d(S) \right)$$

$$\text{S.t.} \quad S = D(R, P)$$

The first order conditions on the Lagrangian lead to the following equations:

$$\frac{R - b_1 \alpha LS^{\alpha-1} - b_2 \beta L^{1-\beta} d^{\beta-1} \frac{\partial d}{\partial S} S^{\gamma} - b_2 \gamma L^{1-\beta} d^{\beta} S^{\gamma-1} - b_3 L^{-\frac{1}{2}} \frac{\partial d}{\partial S}}{R} = \frac{1}{\xi^D{}_R} - \frac{p \frac{\partial}{\partial S} d}{R} \quad (6)$$

$$\frac{p - b_1 \alpha LS^{\alpha-1} \frac{\partial S}{\partial d} - b_2 \beta L^{1-\beta} d^{\beta-1} S^{\gamma} - b_2 \gamma L^{1-\beta} d^{\beta} S^{\gamma-1} \frac{\partial S}{\partial d} - b_3 L^{-\frac{1}{2}}}{p} = \frac{1}{\xi_p \xi_S^d} - \frac{R \frac{\partial S}{\partial d}}{p} \quad (7)$$

$$b_1 S^{\alpha} + b_2 (1-\beta) L^{-\beta} d^{\beta} (S, p) S^{\gamma} = \frac{1}{2} b_3 L^{-\frac{3}{2}} d(S, p) \quad (8)$$

$$S = D(R, p) \quad (9)$$

where $\xi^D{}_R = -\frac{R}{D} \frac{\partial}{\partial R} D(R, p)$ is the storage-charge elasticity of demand of advertisers; $\xi_p = -\frac{p}{D} \frac{\partial}{\partial p} D(R, p)$ is the running-charge elasticity of demand of advertisers; and $\xi_S^d = \frac{S}{d} \frac{\partial}{\partial S} d(S)$ is the library-size elasticity of users' demand. The product $\xi_p \xi_S^d = -\frac{p}{d} \frac{\partial}{\partial p} d(S)$ is the indirect elasticity of users' demand to running charge.

Given parameters b_1, b_2, b_3, α, β and γ, and demand functions $D(R, p)$ and $d(S)$, one can solve for the four variables R, p and S, L from the four equations (6)-(9).

Organizational/Public Library

A key question of an organizational/public library is whether to charge for usage. Use of an organizational library serves organizational goals. Use of a public library benefits not only users themselves but also the community or society at large. It is difficult to measure these external benefits. Digital technology makes it easy to measure usage, but may not help to measure external benefits. A typical way to avoid the difficulty of measuring external benefits is simply not to charge for use.

Another problem of an organizational/public library is to determine collection size. The obvious rule is to require that marginal cost of acquiring and serving the collection equals the marginal benefit to the organization/community. Let $U(d(S))$ be the total benefit of usage $d(S)$ of collection S. Fixing the number of mirrors at one for simplicity, the cost of acquiring and serving S is:

$$C = pS + b_1 S^{\alpha} + b_2 d^{\beta}(S) S^{\gamma} + b_3 d(S)$$

The marginal cost is:

$$MC = p + b_1 \alpha S^{\alpha-1} + b_2 \beta [d(S)]^{\beta-1} S^{\gamma} \frac{\partial}{\partial S} d(S) + b_2 \gamma [d(S)]^{\beta} S^{\gamma-1} + b_3 \frac{\partial}{\partial S} d(S)$$

The marginal benefit to the organization/community from S is:

$$MU = \frac{\partial}{\partial d} U(d(S)) \frac{\partial}{\partial S} d(S)$$

From $MU = MC$, one has:

$$p = \frac{\partial d}{\partial S} \left[\frac{\partial U}{\partial d} - \left(b_1 \alpha S^{\alpha-1} \frac{\partial S}{\partial d} + b_2 \beta [d(S)]^{\beta-1} S^{\gamma} + b_2 \gamma [d(S)]^{\beta} S^{\gamma-1} \frac{\partial S}{\partial d} + b_3 \right) \right] \quad (10)$$

This in effect gives demand S of the library at each p.

Equation (10) suggests that publishers could charge organization/public libraries by unit of information usage ($p / \frac{\partial d}{\partial S}$). The per-usage charge on the library would be related to the per-usage benefit ($\frac{\partial U}{\partial d}$) and the marginal cost of serving S.

Per-usage benefit to the organization/community is as difficult to measure as external benefit, because the per-usage benefit includes the external benefit. However, one can not avoid this difficulty in this case, because the acquisition charge can not be unilaterally determined by the library. Unlike the case of charging users, the market relationship between organizational/public libraries and publishers gives persistent incentive for the libraries to measure benefit.

Because the usage per unit of collection ($\frac{\partial d}{\partial S}$) is related to the number of users in the organization/community, Equation (10) suggests that publishers may alternatively charge libraries by the number of users in the organization/community.

IMPLICATIONS FOR THE TECHNICAL ARCHITECTURE

The technical architecture of a digital library includes specifications about the client (browser), communication protocols, and servers for access control, indexing and searching, depositing and retrieving, billing and so on. What are the implications of service models and the formalization of operational decision-making on the architecture?

Above analyses show that service models of digital libraries can be identified. Their operational decisions can be formalized. The optimizing models studied are examples of such formalization under simplifying assumptions. However, extension and refinement of these models for real decision-making are possible. Furthermore, the operation of a digital library will generate statistics on usage, pricing and acquisition etc. These statistics can be used to estimate demand and supply of users and publishers. Therefore, it seems feasible and beneficial to implement decision models into the architecture of digital libraries as a managing server. The managing server would make use of the operational statistics of the library, estimate relevant parameters, automate some operational decisions and assist the librarian on others. For example, it could 1) decide automatically the optimum number of mirrors for each document; (2) compute and suggest optimal responses of demand for publications and usage pricing to the publisher's proposal of storage charge and running charge, in the case of the commercial library; (3) in the case of the advertising library, estimate publishers' demand for advertising and users' demand for commercials and compute optimal storage and running charges.

Different service models have different decision models. A general managing server could implement several models.

One specific library would configure the server to the intended model in which it operates.

In order to retain flexibility for refinement and changes in the decision models, the managing server would be built as modularly as possible. Particularly, the modularity would allow the library administrator to adjust decision variables, objective functions etc.

A managing server would improve decision-making and operations of the digital library and reduce the requirement for human intervention. For digital libraries, as the next stage in the automation of information access, to include managing servers to automate operating decisions seems a natural step.

CONCLUSIONS

Digital libraries are not only technical facilities but also economic enterprises. As economic enterprises, digital libraries simplify the interface between information consumers and information producers, and relieve the users' burden of navigation and the publishers' burden of marketing. They economize the costs of storage, search and telecommunication in providing access to information.

Properties of information consumption, namely internality and externality, and features of information cost structure suggest that commercial, organizational/public and advertising libraries and the national library are likely service models.

Commercial libraries and advertising libraries face decisions of acquisition, mirroring and pricing; organizational and public libraries face decisions of acquiring and mirroring. We have modeled and analyzed these decisions under simplifying assumptions. A distinct feature of the economics of a commercial library is that publishers can be isolated from final users; usage price has to be agreed upon by both the publisher and the library.

One implication of the service models and formalization of operational decisions is the feasibility and usefulness of including a managing server in the technical architecture of digital libraries. Based on models of operating decisions, statistics of past usage, acquisitions and prices, and other available parameters, a managing server may automate and assist decision-making about acquisition, mirroring, and pricing library services, thus improving the efficiency of the digital library.

REFERENCES

1. Baumol, W. J. and M. Marcus (1973): *Economics of Academic Libraries, American Council on Education*, Washington, D.C.

2. Bookstein, A. (1984): "An Economic Model of Library Service", In S. A. Roberts (ed.) *Costing and the Economics of Library and Information Services*, Aslib, London.

3. Casper, C. A. (1979): "Pricing Policy for Library Services", *J. of the American Society for Information Science*, September, 1979.

4. Davis, J. R. and C. Lagoze (1994). "A Protocol and Server for a Distributed Digital Technical Report Library", Cornell University Computer Science Dept. Technical Report 94-1418, June 1994.

5. Fox, E. (ed.) (1993): *The Source Book of Digital Library,* Dept. of Computer Science, Virginia Tech University, TR-93-35.

6. Gladney, H. M., N. J. Belkin, Z. Ahmed, E. Fox, R. Ashany and M. Zemankova (1994): "Digital Library: Gross Structure and Requirements (Report from a Workshop)", IBM Research Report RJ 9840, May 1994. (Available at ftp://fox.cs.vt.edu/pub/DigitalLibrary/RJ9840.ps)

7. Kahn, R. and R. Wilensky (1995): "Accessing Digital Library Services and Objects: A Frame of Reference", Draft 4.4. (Available at http://www.cnri.reston.va.us/home/cstr/arch/k-w.html.)

8. Knight, D. M. and E. S. Nourse (ed.) (1969): *Library at Large: Tradition, Innovation and the National Interest*. R. R, Bowker Company, New York and London.

9. Lunin, L. F and E. Fox (1993), "Perspective on Digital Libraries: Introduction and Overview", *J. of the American Society for Information Science*, Vol. 44, No. 8, Sep, 1993, 440 - 445.

10. Zahray, A. P. and M. Sirbu (1991), "The provision of scholarly journals by libraries via electronic technologies: an economic analysis", *J. of Information Economics and Policy*, No.4, 127-154.