

Characteristics of Successful Digital Libraries *

디지털도서관 성공도 결정요인에 관한 연구

정진택(Jin-Tack Jung)**

Contents

- | | |
|--|-------------------------------------|
| 1. Introduction | 3.2 Public Goods Approach |
| 2. The Definition of the Digital Library | 3.3 Information Transfer Approach |
| 3. Antecedents of a Successful Digital Library | 3.4 Organizational Ecology Approach |
| 3.1 Human Capital Approach | 4. Summary |

ABSTRACT

The purpose of this paper is to examine the variables that contribute to making a successful digital library with integrated conceptual framework. The central research question is what are the characteristics of most successful digital libraries? and what factors determine whether a new digital library is likely to survive? Three groups of factors that affect the success of an digital library can be extracted from previous research: (1) individual characteristics of the digital librarian; (2) group structural characteristics; (3) conditions characterizing the environment of a new digital library. provide a brief introduction to this comprehensive model.

초 록

디지털도서관의 지속적 성장과 발전을 위해서는 디지털도서관의 성공요인을 새롭고 종합적으로 측정 가능한 모델을 필요로 한다. 이는 디지털도서관 환경이 기존의 환경과는 달리 분산 지향적이며, 이질적인 멀티미디어 정보를 취급하고 점차 양방향 의사소통이 가능한 대화형(Interactive) 환경으로서 다양한 이용자의 욕구에 부응해야 하기 때문이다. 이처럼 새로운 환경에 처한 디지털도서관의 성공요인에 관한 연구는 기존의 틀로는 분석되기 어려운 점이 있다. 본 연구는 이러한 문제의식에서 출발하여, 새롭고 종합적인 접근 방식을 통해 디지털도서관의 성공요인을 측정할 수 있는 종합적인 모델을 개발하고자, 디지털도서관 이용자를 대상으로 하여 디지털도서관에 관련된 변수들을 독립변수로, 디지털도서관의 성공도를 종속변수로 하여 디지털도서관 환경에서 성공요인을 측정하는 모델을 만들고자 시도하였다.

* Thanks to the Hangeung University which provided partial support for the research

**Jin-Tack, Jung, Ph. D

Dept. of Library & Information Science Hangeung University

접수일자 1998년 12월 22일

1. Introduction

The digital library is, in broad terms, a computerized system that allows a community of users to obtain a coherent means of access to an organized, electronically stored repository of information and knowledge (Lynch and Garcia-Molina, 1995). The information resources and technologies embodied by the World Wide Web are now accepted as the primary example of the digital library. Since the first digital libraries began in early 1990s, between 10,000 and 30,000 different digital libraries with dozens to hundreds of participants each are currently in the world-wide Internet and are becoming the central source of information networking.

A considerable body of literature in the fields of library and information science and computer science has recently begun to give digital libraries serious consideration (see Drabenstott 1993, Communications of ACM April 1995); but the main thrust of this literature is the technical creation of powerful and effective digital libraries. Nonetheless, as the digital library becomes established as a component of library and information

services, and as more individuals obtain Web access and utilize digital libraries on the Web, study of the digital library will become an important item on the information research agenda. To date, little attention has been paid to the role of the digital library in the new environment and little opportunity has been available to assess variables that are contributing to a successful digital library. Thus, accepted measures for evaluation of digital library success do not exist.

The purpose of this paper is to examine the variables that contribute to making a successful digital library with integrated conceptual framework. The central research question is what are the characteristics of most successful digital libraries? and what factors determine whether a new digital library is likely to survive? Three groups of factors that affect the success of an digital library can be extracted from previous research: (1) individual characteristics of the digital librarian; (2) group structural characteristics; (3) conditions characterizing the environment of a new digital library.

The dependent variables of interest all relate to the success of digital

libraries, described in terms of contribution levels (Thorns & Connolly 1987), adoption rate (Markus, 1990), longevity and usage.

- Contribution level is defined as the ratio of files contributed in a week to the total number of weekly file transactions by users .
- Adoption rate is the ratio of regular digital library users to the total number of users .
- Longevity is the duration since the digital library started.
- Usage is the average number of page views on an average day.

To avoid a meaningless compilation of many independent variables, I use human capital theory to examine the potential effects of individuals. For group and contextual factors, public goods theory, information transfer theory and organizational ecology theory provide meaningful variables for analysis. Some variables that may be important to another researchers are not included in my analysis, but I believe that a theory-guided selection of variables yields more insight than a compilation of all available variables.

2. The Definition of the Digital Library

Broadly stated, the digital library consists of all digitized resources available to users of a computing environment (Levy and Marshall 1995). Nevertheless, the obviously heterogeneous nature of digitized information resources has led researchers to define the digital library variously, as one or another subset of information in digital format. Some writers have defined the digital library as an extension of resources accessed through traditional libraries (Levy and Marshall 1995), while other research has asserted that it consists of all resources available on the Internet, whether or not such resources are accessed within the traditional library (Adam, Halem, and Naqvi 1995). Adam, Halem, and Naqvi (1995) and Arms (1995) are the primary advocates of this narrowly bounded perspective, claiming the digital library consists exclusively of digitized information resources in the Internet.

Lynch and Garcia-Molina (1995) criticize the assumption of Adam, Halem, and Naqvi (1995) that the

digital library is based exclusively on fully digitized Web-based sources. They suggest that the exclusive emphasis on resources in distributed hypermedia format (i.e., fully digitized resources) is too limiting, and propose that the objective of the digital library is to develop integrative systems "to fully exploit the opportunities that are offered by materials in digital format." Thus, they suggest that the digital library may consist of self-contained digitized resources such as CD-ROM databases, not only resources available on the Web. In their view, the digital library is an interface system, a gateway, for accessing digitized information (Lynch and Garcia-Molina 1995; Drabenstott 1993).

Despite the disagreement concerning which information resources characterize a digital library, the research consistently recognizes that the digital library performs information search and retrieval functions (Borgman, et al. 1996). Users of the digital library are concerned to a great extent simply with gaining access to information (Adam, Halem, and Naqvi 1995), and the environment functions as a gateway or interface providing access to

information resources (i.e., video, audio, and printed information).

Borgman, et al. (1996) asserted that a digital library should be defined not only in terms of information entities that can be collected and organized into digital libraries (tangible information resources), but also in terms of the communication process used in the creation and utilization of those information entities (intangible information resources). Borgman, et al. pointed out that by focusing on the digital library as a collection of tangible resources, Adam's definition disregards the communicative aspects of digital library. Thus, Borgman, et al. (1996) proposed a digital library concept that encompasses two complementary ideas, one emphasizing that a digital library extends and enhances existing information storage and retrieval systems, incorporating digital data and meta-data in any form, and the other emphasizing that a digital library is constructed, collected, and organized by a community of users through a form of communication.

In summary, despite the disagreements concerning the definition of the digital library, information

retrieval and communication functions are recognized as common denominators of the digital library definition.

3. Characteristics of Successful Digital Libraries

3.1 Human Capital Approach

The human capital theory argues that high human capital endowment of the founder improves chances for organizational success. Several researchers support this argument. Psychologists tend to describe the founder of an organization by personality traits such as a high need for achievement (McClelland 1961). Economists see them as coordinators, risk-takers and innovators (Herbert 1989). Sociologists identify socio-demographic attributes of founders or see them as "displaced persons" (Brenner 1987)

The economist Becker (1975) distinguishes between general and specific human capital. General human capital is the sum of educational level and work experience of the founder. The notion of specific human capital must be

modified and adapted to the context of an digital library. An important indicator of digital library-specific human capital is "leadership experience," i.e. , experience in managing and directing people.

The digital librarian is usually a key person who sets the norms and procedures of the group, recognizing members' contributions, prompting them to address certain issues and summarizing the state of discussion. He is extremely important to digital library success (Kerr, 1982). Kerr (1986) also summarizes effective leadership styles and skills for moderating online meetings and facilitating electronic groups. The most effective digital librarian style is group-oriented (rather than authoritarian). The group-oriented digital librarian includes all group members interested in the formation of policy and moves them to consensus. Effective digital librarians must exert strong organizational pressures and also be familiar with the technology. The digital librarian should set the agenda and keep the group moving toward its goal.

H1:

The level of education of the

digital librarian will be positively related to digital library success.

The level of experience of the digital librarian will be positively related to digital library success.

The leadership experience of digital librarian will be positively related to digital library success .

3.2 Public Goods Approach

The main idea of this approach is the analogy of digital libraries to public goods. Digital libraries are available to all members in a community of interest, regardless of whether the individual members themselves contribute to the group. Because no single individual is required to contribute to the content of the digital libraries to benefit from it, the "rational member" might be expected to adopt the strategy of consuming the medium without making contributions. Yet, if no one contributed, there would soon be nothing to consume. This is the so-called dilemma of public goods (Ostrom, 1990). Two theoretical perspectives emerged from this approach discretionary databases theory (Thorn & Connolly, 1987) and critical mass theory (Markus, 1990).

Thorn & Connolly (1987) developed a theory of discretionary databases from the notion of public goods and used it to predict rates of contribution to information database shared within an organization. In a series of laboratory experiments, they found that the lower the perceived value of the information and the less symmetrical the benefits to users in a group, the lower the contribution levels. Furthermore, individuals contributed less as participation costs and group size increased.

Critical mass theory (Markus, 1990) takes a parallel approach. Markus (1990) attempts to explain the growing adoption of electronic mail in a community of interest until a state of near-total participation, or universal access, comes to exist. The theory predicts that the chances of attaining universal access are inversely related to the resource contributions (skill, effort or cost) required from users. On the other hand, the more heterogeneous the interests and resources found among members of a community, the greater the chance of achieving universal access. Task interdependence, centralization of resources, user group size, and geographical dispersion are hypothesized to be directly related to hetero-

geneity and hence to adoption.

Though the two theories share common underlying assumptions about human behavior, they contradict each other in some areas. For example, the critical mass theory makes the opposite prediction regarding the role of user group size. If larger user group sizes are also more heterogeneous, then critical mass theory predicts that universal access is more likely to be achieved in larger groups. According to this view, participation relates not only directly to user group size, it is exponentially related to the number of participants (Markus 1990).

In my view, the more users there are, the more each user can safely assume that someone else will contribute to the group. This could even be a positive factor when making an adoption decision. When considering whether to adopt a new digital library, potential users may find larger systems more attractive precisely because they seem to require a smaller investment of resources in terms of their own contribution than small user groups.

H2:

The user group size will be

negatively related to contributed levels but positively related to adoption measures of digital library success.

The diversity of contents will be positively related to the success of digital library.

3.3 Information Transfer Approach

Two applicable concepts can be borrowed from this approach: technical gatekeeper (Allen, 1967) and invisible college (Price, 1963; Crane, 1972).

The technical gatekeeper (TG) is a key person whom average people in an organization rely on for outside information. TG's contribute directly to a group's technical goals, in that they are typically high performers. They also tend to be highly cited, front-line supervisors. According to organizational information flow model (Allen, 1967), new information is brought in through the TG. It can then be communicated to other TGs in the organization through the TG network and disseminated outward from one or more points to other members of the organization. This is a spontaneous process, without management intervention. TG networks

tend to be highly specialized.

The invisible college (IC) is a concept which was introduced by Price (1963) in which he likened it to an "in-group". Crane (1972) embraced the view that interactions could be plotted as a social circle and that the influential figures in the group hold the informal network together. Mullins (1973) defined IC as concentrations of interest ties without clear boundaries. They encourage feedback, provide filtering function, facilitate boundary spanning and have a bonding effect. Parker (1982) suggested the use of informal communication channels and media was the single best predictor of research group productivity. The existence of IC can be elucidated from citation analysis (Griffith, 1972), sociometry (Crane, 1972), and micro-sociology (Mulkay, 1980).

H3:

The existence of technical gatekeeper and/or invisible college will be positively relate to the success of digital library.

3.4 Organizational Ecology Approach

Organizational ecology theory (Hannan and Freeman, 1989) addresses

evolutionary processes within or between populations of organizations observed over periods of time. The analysis of organizational mortality is an important topic in this field, and I believe that organizational ecology and its associated empirical research offer a background framework to derive hypotheses regarding the determinants of the survival chances of new digital libraries.

Organizations begin with a "liability of newness" (Stinchcombe 1965). Many organizations die young, and organizational ecologists have elaborated on this age-dependent pattern of mortality in considerable detail (Hannan and Freeman 1983). Young organizations have a higher risk of failure than older organizations.

Organizational ecologists often discuss the "liability of smallness" in connection with the liability of newness (Aldrich and Auster 1986). The assumption is that large, new organizations have better survival prospects than small, new organizations. Initial size may be measured by the number of people participating at the time of founding. A large pool of human resources improves the chances of a new organization to

survive the critical start-up period and to cope with random shocks from the environment.

A third feature of organizational ecology that is useful for my analysis is organizational strategies in connection with the concept of "niche". Hanan and Freeman (1977) distinguished between generalist and specialist strategies. Generalist organizations aim to occupy a broad niche, whereas specialist organizations aim for a narrow niche. In a study of small organization, they examined whether generalist and specialist organizations have various survival chances under different environmental circumstances.

H4:

The age of digital library will be positively related to the survival of digital library.

The size of digital library will be positively related to the survival of

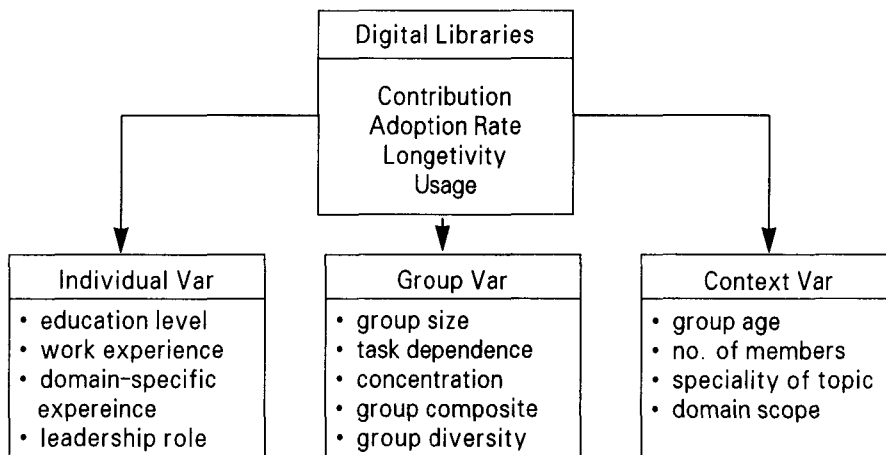
digital library.

The domain specialty will be positively related to the survival of digital library.

4. Summary

Previous research on determinants of survivor chances of new digital libraries has concentrated on individual characteristics of the digital library founder. Human capital theory captures most of these individual characteristics in a theoretically meaningful way. Public good approach and information transfer approach address the group structural attributes. For organizational and environmental determinants of survival, the organizational ecology approach promises theoretical progress. These theories yield a rich set of testable hypotheses.

Conceptual Model of Successful Digital Libraries



Preliminary Analysis of Most Successful Digital Libraries

Independent Variables		PACs-DL	STAT-DL	TQM-DL	QUAL-DL	CWIS-DL
FOUNDNRR	exp of educate	PH.D	PH.D	PH.D	PH.D	PH.D
	exp of work	17	15	14	12	10
	exp of domain	15	15	12	5	8
	exp of leader	YES	YES	YES	YES	YES
GROUP	Group size	15,000	12,000	6,500	3,000	13,000
	Task-dependence	mostly	mostly	mostly	mostly	Mostly
	Task-concentra.	mostly	mostly	mostly	mostly	Mostly
	Technical gate.	5	7	4	3	8
	Indivisible coll.	3	5	2	2	5
CONTEXT	Task-diversity	high	high	medium	medium	High
	Group age	36 mon	38	17	24	22
	No. of members	800	1200	200	600	400
	Speciality level	medium	high	medium	high	high
	Domain scope	high	high	medium	low	medium

References

- Adam, Nabil R; Milton Halem and Shamin Naqvi. "Promising Research Directions in Digital Libraries." N. Adam, B. Bhargava, and Y. Yesha (Editors), *Digital Libraries: Current Issues*, Lecture Notes in Computer Science, Springer-Verlag, Chapter 3, 21-29, May 1995.
- Aldrich, H. E. & Auster, E. (1986), "Even Dwarf s Started Small : Liability of size and Age and Their Strategic Implications," Pp.165-9E3 in *Research in Organizational Behavior vol.8*, edited by B.M. Staw and L. L. Cummings, Greenwich, CT: JAI Press .
- Allen, T.J. (1967). *Organizational aspects of information flow in technology*. *Aslib Proceedings*; 1968; 20; 433-453.
- Anis, N. (1991). Interview with Ward Christensen the first BBS operator. *BroadWatch Magazine* , pp. 14-22 .
- Arms, William Y. "Key concepts in the architecture of the digital library." *D-Lib Magazine*, Jul, 1995.
- Becker, G. (1975). *Human Capital*, 2d. ed. Chicago: University of Chicago.
- Borgman, et al. "Social Aspects of Digital Libraries." February 16-17, University of California, Los Angeles, 1996.
- Brenne, R. (1987). *Rivalry: In Business, Science, Among Nations*. Cambridge, England: Cambridge University Press.
- Crane, D. (1972). *Invisible College: Diffusion of KnowlDLe in Scientific Communications*. Chicago: University of Chicago Press.
- Drabenstott, K. M. *Analytical Review of the Library of the Future*. Council on Library Resources, Washington, DC. 1993.
- Digital Libraries*. *Communications of the ACM*. Vol. 38, No. 38, pp.22-96, April 1995.
- Griffith, B.C. (1972). *Coherent groups in scientific change*. *Science*; 1972; (15, Sep); 959-964.
- Hanan, M. T., & Freeman, J. (1989). *Organizational Ecology*. Cambridge: Harvard University Press.

- Hanan, M. T., & Freeman, J. (1977). "The Population Ecology of Organizations," *American Journal of Sociology* 82: 929-64.
- Kerr, E. B. (1982). *Computer Mediated Communications: Status and Evaluation*, Academic Press. New York.
- Kerr, E. B. (1986). *Electronic Leadership: A guide to moderating online conference*, IEEE Transactions on Professional Communication, Vol 29, 12-18.
- Levy, D. M. and Marshall, C. C(1994). "What Color was George Washington's White Horse? A Look at the Assumptions Underlying Digital Libraries," *Proceedings of Digital Libraries' 94*, pp.163-169, 1994.
- Lynch, Clifford and Garcia-Molina, Hector(1995), *Interoperability, Scaling, and the Digital Libraries Research Agenda: A Report on the May 18-19, 1995 IITA Digital Libraries Workshop*, August 22, 1995 (<http://www-diglib.stanford.edu/diglib/pub/reports/iita-dlw/main.html>), 1995.
- Markus, M. L. (1990). "Toward a "critical mass" theory of interactive media: Universal access, interdependence and diffusion. *Communication research*. 14. 491-511.
- McClelland, D. C. (1961). *The Achieving Society*. Princeton, N.J: Van-
Nostrand Herbert, R. F. (1989). "In search of the meaning of entrepreneurship," *Small Business Economics*: 1: 39-49
- Mulkay, M. J. (1980). *The Sociology of science in East and West*. *Current Sociology*: 1980: 28 (3): 1-350.
- Mullin, N. C. (1985). *Invisible colleges as science elites*. *Scientometrics*: 1985: 7 (3-6): 357-68.
- Ostrom, E. (1990). *Governing the commons*. New York: Cambridge University Press.
- Rafaeli, S. (1988). *Interactivity: From new media to communication*, *Advancing communication science: Sage annual review of communication research* (Vol. 16, pp.110-134). Newbury Park, CA: Sage.
- Swift, C. R. (1989). *Audience activity in computer mediated communication*. Unpublished doctoral dissertation. Indiana University, Bloomington.
- Thorn, B. K. & Connolly, T. (1987). *Discretionary data bases: A theory and some experimental findings*. *Communication Research*, 14, 512-528.

Price, D. (1963). *Little Science, Big Science*. New York: Columbia University Press: 1963.

Parker, R. H. (1982). *Bibliometric models for management of an information Store*: JASIS:

1982; 33 (3): 134-138.

Stinchcombe, A. L. (1965), "Social Structure and Organization," pp. 142 -193 in *Handbook of organization* edited by J. G. March. Chicago: Rand McNally .