Assessing Digital Library Services: Approaches, Issues, and Considerations

John Carlo Bertot <bertot@lis.fsu.edu>
Professor
School of Information Studies
Florida State University

ABSTRACT

Over the last decade, there have been significant efforts to develop digital libraries. These efforts, however, largely focused on the building of the technology, telecommunications, and standards (i.e., metadata) infrastructures that enable digital libraries to function technically. More recently, the efforts include the building of digital content that populate digital libraries. Little attention in the development of digital library initiatives – particularly those funded by through national and international funding agencies – actually assessed the extent to which the digital libraries met user information needs, service quality expectations, facilitated outcomes, or add value in other ways. This paper identifies a number of evaluation strategies to assess digital libraries through a number of approaches that can yield a variety of data regarding the efficiency, effectiveness, and quality of digital libraries. The results from these evaluation strategies can have a number of significant implications for the continued development of digital libraries.

1 Introduction

In 1999, Lakos [1] used the phrase "culture of assessment" in his discussion of the need for libraries to develop and sustain coherent and pervasive evaluation strategies regarding library service and resource provision. Briefly, Lakos argued that libraries need to create an organizational culture in which assessment is a key component to understanding the meeting space of users and libraries. This type of culture is one in which library services – electronic or traditional – are under an ongoing evaluation system so as to foster continued improvement in meeting both library and customer needs.

There appears to be a range of evaluation effort in libraries, however, from minimal (e.g., write a check/"silver bullet" generally dependent on the latest trend in assessment) and integral (e.g., constant consideration of evaluation and assessment activities dependent on library data needs that can be both internal and external). A true culture of assessment requires different librarian attitudes and perceptions of library services and resources provision, different library management and working group structures, and continual librarian training and education in a number of areas.

Moreover, librarians need to build evaluation activities and strategies as they are developing the services, resources, and programs that they intend to provide to their customers. Post program evaluation strategies often fall short of providing useful and meaningful data, as they are an after thought, do not provide pre-program baseline data, and occur at a later point in time without clear objectives and goals [2]. In short, researchers and practitioners will get out of their assessment activities what they are willing to put into such activities.

Library practitioners and researchers are only just beginning to develop evaluation strategies and approaches for digital libraries and network-based services and resources, particularly from a user perspective. Evaluating network services requires the same commitment to assessment and learning as does the evaluation of traditional library services/resources on the part of library managers, staff, and researchers. Indeed, the networked environment is such that libraries need to develop new measures, methods, and approaches for assessing digital services – even if the approaches adopt familiar evaluation frameworks.

2 Evaluation Frameworks

The primary purpose of conducting evaluation research in a library setting is twofold: 1) to understand user interaction with library resources and services; and 2) to capture data that inform the planning, management, and implementation of library resources and services. Integrated research, evaluation, and planning efforts, through various measurement approaches and methods, can assist libraries to efficiently and effectively develop, implement, and change resources and services that meet user information needs over time. They can also provide an understanding of the broader social context of libraries from a user perspective.

2.1 Evaluative Approaches

There are no standard definitions or approaches to library evaluation approaches, strategies, or practices. Each evaluative approach offers potential information – based on collected data – particular to a specific area of focus within a library. The area of focus may be broad in scope such as a library's digital collections and presence or defined narrowly for a specific task within a specific setting such as a
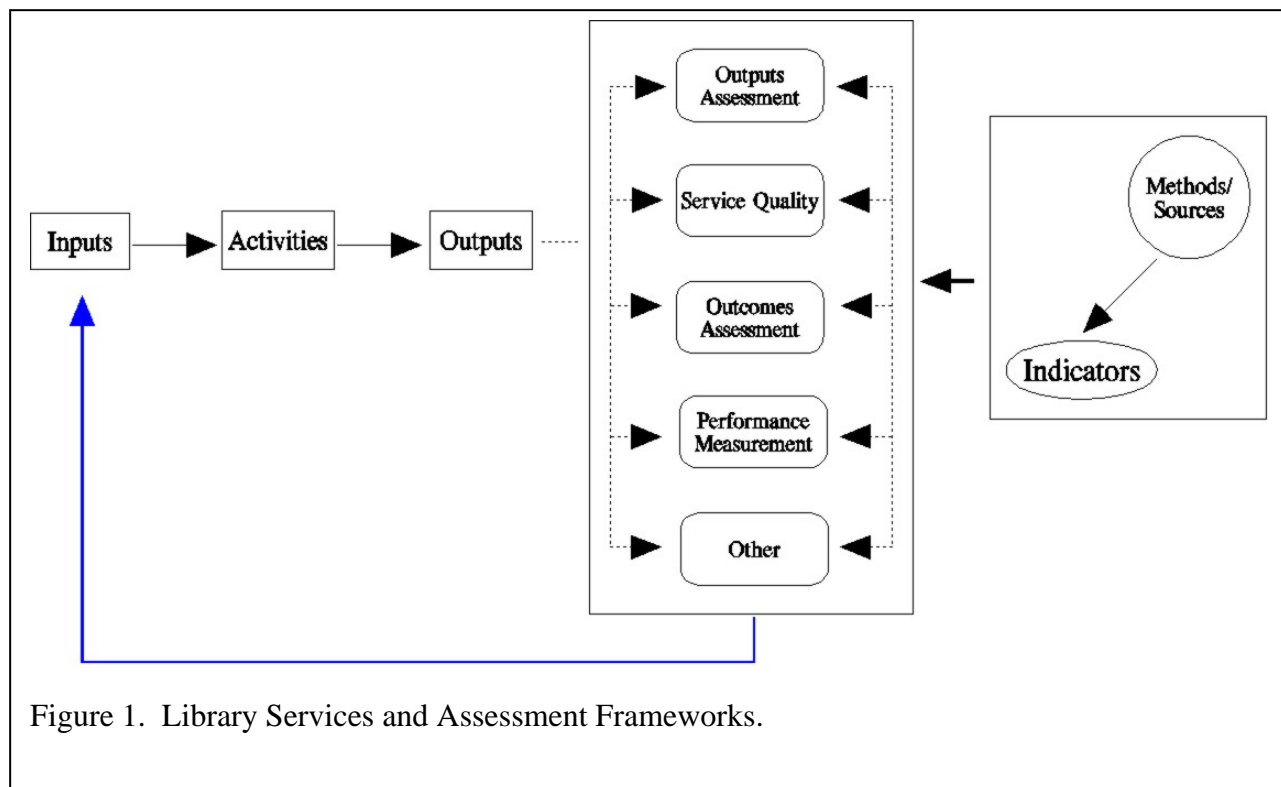
Figure 1. Library Services and Assessment Frameworks.

special digital collection, feature, or service. Also, evaluative approaches tend to be tailored to particular needs of an organization; linked to available time and funding; limited by scope and breadth of application due to funding, planning, etc.; and determined by direction that ranges from what libraries collect historically to immediate learning needs regarding library services, resources, and programs.

Library evaluation activities can be complex and include a number of factors throughout the process. In general, however, library evaluation centers on three key components (see Figure 1):

- **Inputs** – the resources that libraries invest (e.g., money, staff, workstations, online commercial databases);
- **Activities** – the library services/resources that the inputs actually generate (e.g., licensed resources availability, story hours, training sessions); and
- **Outputs** – the number of services/resources generated from library investments (e.g., number of workstations, number of database licensed, print material purchased, number of training sessions).

These three components form the basis for essentially all library service and resource evaluation strategies, as they provide vital baseline data regarding costs, investments, services, and resources.

Librarians and researchers engage in a number of methodologies (e.g., surveys, log analysis, focus groups, other) and use multiple indicators (e.g., number of sessions, number of full text downloads)

to determine the scope and breadth of their services. These indicators and methods tend to reflect the evaluation approach that the researchers and librarians use as part of their assessment activities. Ultimately, the results of the evaluation efforts feed back into the planning and resource allocation decisions of the library.

2.2 Types of Evaluative Approaches

Researchers and practitioners may use a number of approaches to evaluate library resources and services from multiple perspectives and the approaches may be library-centered or user-centered. Each approach is part of an evaluative process that includes planning, data collection, and evaluation as components of the evaluation. For purposes of this paper, evaluative approaches will be presented briefly within four broad areas of outputs assessment, performance measures, service quality, and outcomes assessment (see Table 1).

Each of these areas represents multiple approaches. One can modify each approach to fit specific needs and to answer questions based on decision making needs, in the presentation of resources or services, and upon particular perspectives. Library-centered approaches evaluate the presentation of resources and services with an emphasis on efficiency and effectiveness. User-centered approaches evaluate the quality of the presentation of resources and services; the inclusion of the needs of users, or patrons in accessing resources and services; and may include the library

Table 1: Overview of Selected Evaluative Approaches.			
Evaluation Approach	**Description**	**Examples of Evaluative Practices**	**Applications**
Outputs Assessment	Collection of counts from use of resources, services, or programming	1. Collection of traditional counts, e.g. Reference, circulation, etc. 2. Focus Groups 3. Interviews	1. Evaluation of resources, services, and programs 2. Planning
Performance Measures	Developed for evaluating the presentation and/or delivery of specific library resources, services, or programs	1. Availability studies 2. Usability studies. 3. Web-page analysis 4. Content analysis 5. Functionality analysis	1. Usability of resources, services, or programs 2. Availability of resources 3. Determine efficiency or effectiveness of presentation of resources, services, or programs, etc.
Service Quality	Developed to determine the overall quality of both traditional and network-based resources and services	1. SERVQUAL 2. LibQUAL+ 3. Balanced Scorecard 4. Best Practices 5. Benchmarking	1. User satisfaction with library services 2. Library staff/management appraisal and/or desired level of service provision
Outcomes Assessment	Developed to assess the effects of programming on patrons in terms of benefits to patrons	1. Outcomes as a product of programming 2. Outcomes as the effect of programming	1. As a product, initial measurable affect of programs on patrons 2. As an effect, initial to long-term evaluation of measurable benefits for patrons based on specific programming

community in general within the approach, or target specific stakeholder groups [3].

2.2.1 Outputs Evaluation

Outputs assessment typically involves the identification of the number of library activities that patrons engage (e.g., number of database sessions, number of database items examined, number of training sessions conducted, etc.) (see Table 1). They may also, however, include data that are qualitative in nature such as the results of focus groups and interviews. Outputs are collected to determine the usage of library resources and services and as part of a library's planning process.

It is important to note that entire library data collection systems center on this approach to library use, uses, and performance. In the U.S., for example, the Federal State Cooperative System (FSCS) managed by the National Center for Education Statistics (NCES) collects annual public library data focused on approximately 50 data elements; NCES also manages data collection activities for academic and school libraries through its library statistics program; the Association of Research Libraries (ARL) collects annual statistics from its members and so too does the Association of College Libraries (ACRL); and, as a final example, the Public Library Association collects annual public library statistics from a sample of public libraries through its Public Library Data Service (PLDS) program.

Traditionally, libraries have collected output data in many forms from reference and circulation departments, facility usage, etc. With the advent of measure the effectiveness and efficiency of library resources and services from both the library and the user perspectives. These forms of measures are versatile and changeable, determine efficiency and effectiveness, and are used to determine needs of users.

Performance measures include availability and usability studies, Web page analysis, and functionality testing. Performance measures developed for measuring impacts of physical resources and services are often adapted to the networked environment and may include outputs assessment counts and approaches. Examples of performance measures include cost per item circulated, cost per full text download, percentage of virtual visits, correct answer fill rate, etc.

2.2.3 Service Quality

In general, service quality is an evaluation of how well a library provides a service, resource, or program. Service quality approaches include evaluation from the library and user perspectives, of

libraries as a field, and of the library as an institution.

Approaches used to determine service quality saw rapid growth during the 1980s with the shift to user-centered research and with the development of electronic resources and services. Approaches developed in the business environment such as best practices, benchmarking, SERVQUAL, and others were adapted for use in a library environment in the measuring of service quality.

SERVQUAL was developed specifically to measure quality of business services based on perceptions, desires, and a minimal level of expectation of consumers. LibQUAL+ was developed specifically to measure satisfaction in the quality of library services based on perceptions, desires, and a minimal level of expectation of users [4].

Just recently, ARL announced an initiative to develop a LibQUAL+-based approach to measuring library electronic resources and services. This initiative is in the beginning phases and will develop over time (see http://www.libqual.org for additional information).

2.2.4 Outcomes Assessment

Outcomes assessment is relatively new to library evaluation. This evaluation approach seeks to determine the impact of library resources and services on the users of library resources and services. In particular, outcomes assessment seeks to determine the benefits or changes in knowledge or skill level, behavior, attitudes, or any change that may be seen as a benefit for the participants. As a product, outcomes are immediate measurable benefits for users with potential long-term effects. As an evaluative approach, a large body of work is available from the fields of education and sociology in which outcomes are typically measurable results of interventions or programs.

When viewed as an effect, the measurable change is part of the assessment process in determining the initial to long-term impact of specifically developed programs on participants of the programs. Developed programs are designed in terms of institution goals or standards to create benefits for participants, and performance measures are used

within the programs to measure the differences produced by the programs.

2.2.4 Balanced Scorecard

Libraries are also applying the balanced scorecard approach to evaluating library services [5]. The balanced scorecard, developed by Kaplan and Norton [6], is a method to determine overall organizational performance along four dimensions – financial, customer, internal, and innovation/learning dimensions. This approach provides libraries with an overall assessment technique that involves multiple evaluation frameworks – outputs, service quality, and outcomes.

2.2.5 Summary of evaluative approaches

In summary, one can describe the assessment approaches in the following manner:

- **Outputs assessment** involves the identification of the number of library activities that patrons use (e.g., number of database sessions, number of database items examined, number of training sessions conducted, etc.);
- **Performance measurement** involves the use of efficiency and effectiveness measures (e.g., cost per loan, cost per item downloaded);
- **Quality assessment** involves determining the degree to which users find the library services/resources (outputs) to be satisfactory or meets developed service quality standards;
- **Outcomes assessment** seeks to determine the impact of the library's services/ resources (again, outputs) on the library service and resource users, or benefits, changes in skill/knowledge that library users derive from library services/resources; and
- **Balanced scorecard** uses aspects of outputs, performance measurement, quality, and outcomes assessment approaches to develop an overall sense of the service along financial, customer, internal, and innovation/learning dimensions.
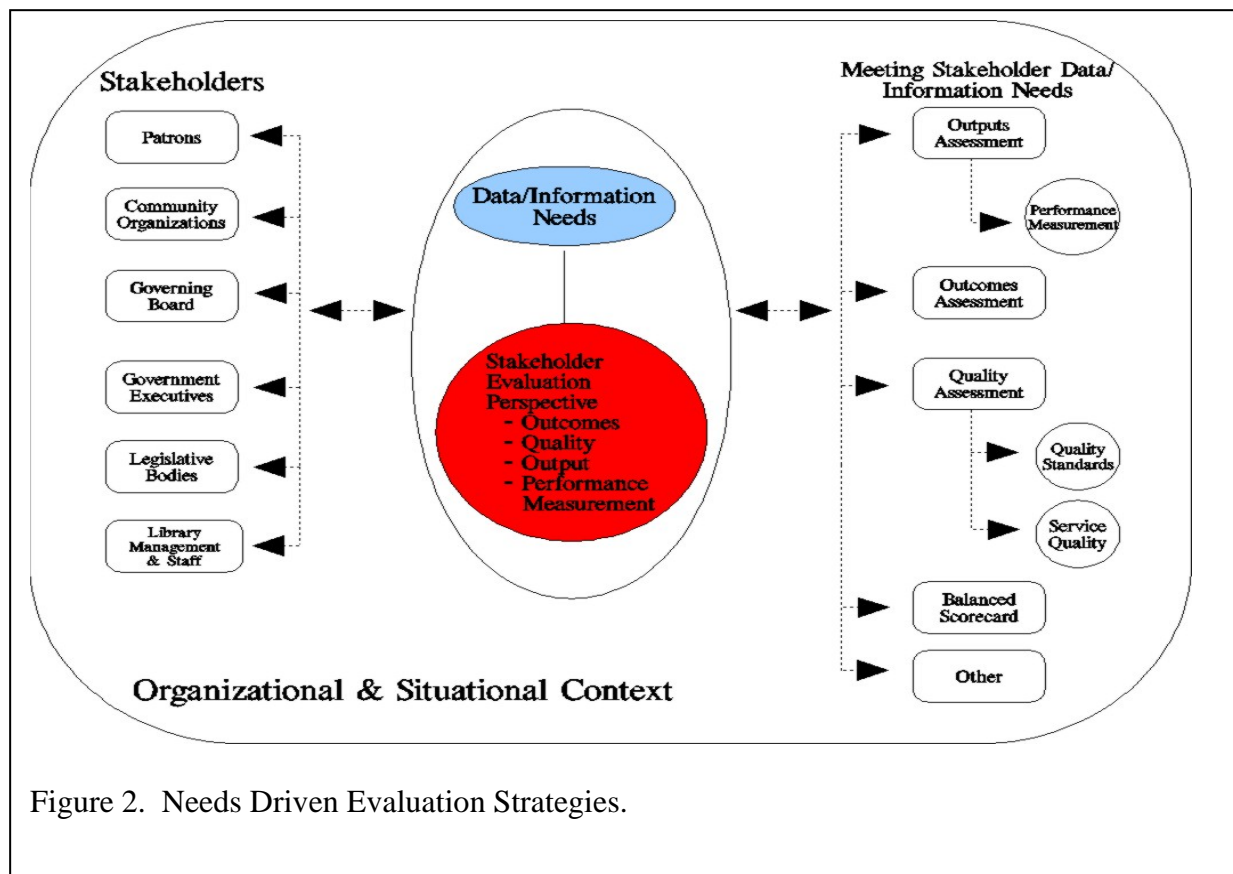
Figure 2. Needs Driven Evaluation Strategies.

Evaluative approaches are developed to answer the questions of what libraries need or want to know regarding their resources and services (see Figure 2). A number of factors may influence a library's need/desire to know, such as internal and external stakeholders such as library management, governing boards, or other community-based stakeholders. To meet these data needs, evaluation activities within libraries often use multiple assessment strategies and approaches, as no one approach will likely provide all the answers for library management and decision making purposes.

The approaches are versatile, and as presented here within four broad categories, often overlap and may serve multiple purposes. Performance measures may use aspects of outputs approaches within the evaluative scheme and the results may serve as indicators of service quality. Approaches developed for physical settings are adapted and utilized in the networked environments, though increasingly there is a need for separate and distinct measures of library electronic services and resources as they have no corollary in the traditional library operating environment.

3 Different Framework, Different Results

Each evaluative framework presented (admittedly selectively and briefly) yields a different type of data regarding a library service, resource, or program. For example:

- **Outputs** provide data regarding quantity – how much and how often;
- **Service Quality** approaches provide data regarding how users (and librarians) find the library's services, resources, or programs – satisfaction, meet or exceed expectations, perceived value.
- **Outcomes** approaches provide data regarding the extent to which library resource, service, and programs created a change in skills, knowledge, or behavior of users – information literacy, ability to use technology, etc.

Thus, researchers and practitioners who desire data regarding how much, is the service any good, and what was the impact of the service will need to engage in evaluation activities that deploy multiple assessment approaches. No one evaluation framework will provide the evaluators with all the necessary data – one size does not fit all. networked resources and services, however, outputs today include usage statistics of the networked environment that include data and usage statistics from digital reference services, database article downloads, remote visits to library websites, and others. Typically, these online data and usage statistics are collected using electronic means.

Electronic means of collection include data from log files that track and document the information seeking behavior of users; network traffic statistics that offer data on where users originate, browsers
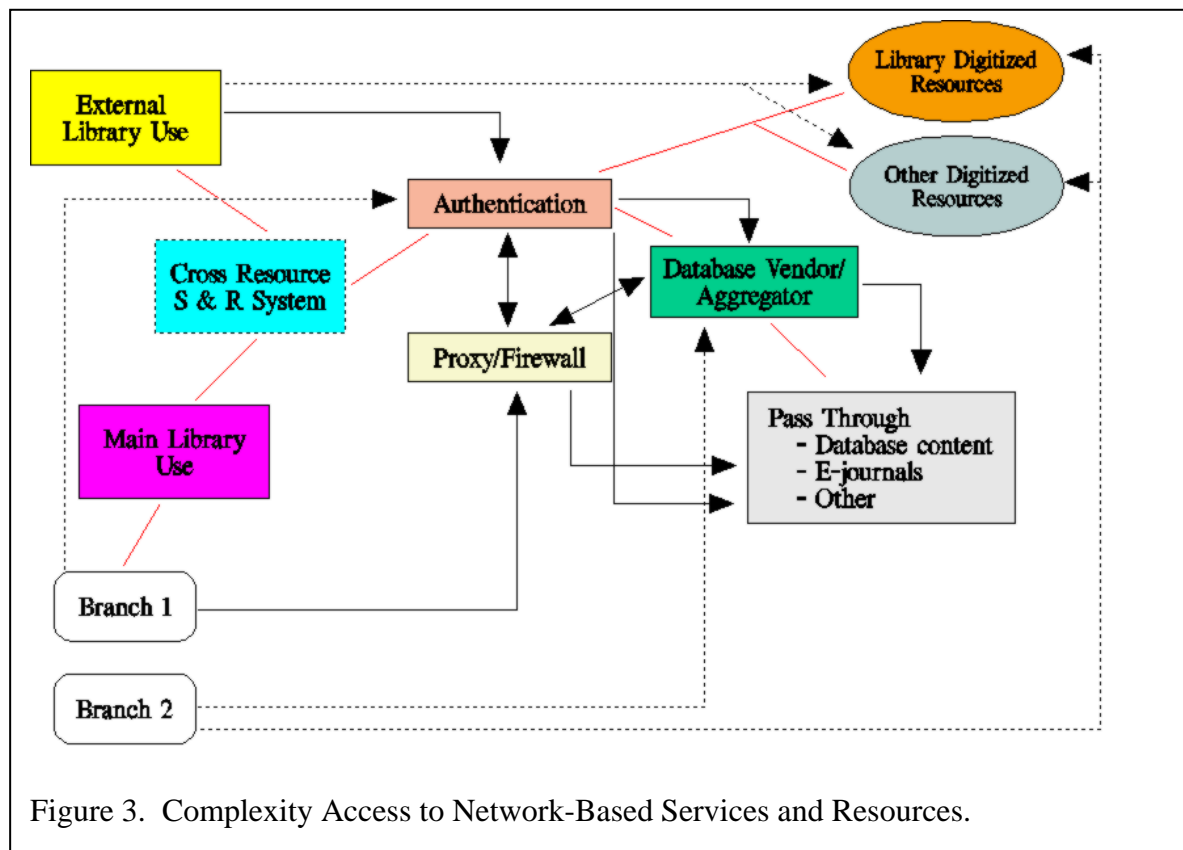
Figure 3. Complexity Access to Network-Based Services and Resources.

used, files accessed, and other data; and use of databases that can be tracked from statistics gathered directly from library developed databases and from statistics supplied by vendors of databases. As technology continues to evolve, the development of electronic means to collect outputs will as well.

2.2.2 Performance Measures

Performance measures (also known as performance indicators) have a long history within the tradition of library research. They are used to

4 Issues with Evaluation Frameworks and Measuring Electronic Services

There are several issues associated with service quality and outcomes assessment in general and in the networked environment in particular. A more detailed discussion of these issues is available in Bertot & McClure [7]. This paper, however, focuses on a high level discussion of service quality and outcomes assessment in the networked environment. Figure 3 demonstrates the complexity of library network-based service and resource provision. At their core, service quality and outcomes assessments focus on user-based perceptions of a) the quality of library services/resources, and b) the impacts of those services/resources on users.   However, as

Figure 3 shows, a vast majority of network-based services/resources that libraries provide are not under the control of the library. For example:

- Libraries are often not the content creators/managers for network-based services and resources
  - OPACS and other internal operational software are most often purchased/leased from specific vendors and are proprietary,
  - Licensed content (e.g., databases, e-books, and the interfaces used to access vendor content) are the property of the vendor(s), and libraries typically lease that content through annual licensing agreements (through libraries can in fact purchase permanent access to e-book holdings and other resources), and
  - A new, and likely to increase in use, vendor-based product is that of a cross-resource search and retrieval interface (think Google) that libraries can purchase for the purpose of enabling customers to search across vendor, web, and library online resources through a single interface. This interface,

which sits in between the user and various other resources, is a proprietary vendor product not under the control of libraries; and

- Various technology infrastructures are not part of the library network/equipment. Customers can access "library content" from a number of locations (e.g., office, home, dorm room, other) with a wide range of computing technology and connectivity (including wireless connectivity and mobile devices). Moreover, external library connectivity has many parties involved from leased-line providers (e.g., academic computing, county information technology services, bell operating companies) to ISPs, to phone lines and wireless technologies.

To summarize, then, libraries do not control a vast majority of their network-based services and resources. Therefore, any service quality and outcomes assessment techniques will need to take that into account.

This is a particularly problematic issue with currently promoted service quality and outcomes assessment products. For example, ARL's LibQUAL+ initiative [8], and the outcomes assessment approach promoted by Hernon & Dugan [9], use or recommend the use of survey instruments and other data collection techniques that mix online and print/traditional library services and assume library ownership of collections, services, and resources. These approaches can be quite useful at gauging library service quality/outcomes in the aggregate. Research indicates, however, that the print/ traditional and electronic environments differ dramatically in important key areas such as user information seeking behavior [10] and the ability of users to engage and extract content [11]. Lumping together traditional and networked services, therefore, leads to confounded variables, data, and results – and potentially erroneous conclusions regarding customer perceptions of outcomes and service quality. As mentioned earlier, however, ARL is embarking on an electronic service/resource LibQUAL+-based approach.

There is a substantial need for service quality and outcomes assessment tools to probe deeper into the specifics of the services/resources they are assessing rather than continue to ask generalized questions. The general questions are helpful to provide libraries, at a glance, successful and less than successful areas of library services according to users. They do not, however, provide specific reasons for the success or lack of success of such services. Thus, libraries need to consider what the subsequent evaluation effort will be to enable in-depth probing into particular service/resource areas.

Moreover, it is likely the case that customers may actually provide feedback regarding a "library service" that is not actually provided by the library,

such as online leased content. In most instances, libraries simply serve as gateways to content that resides with, and is owned by, external entities. This begs the question: upon what, exactly, would libraries measure service quality and outcomes? For example, when a user provides feedback regarding the level of satisfaction with an online journal, is that user assessing the connectivity that leads to the journal, the interface that leads to the journal, the authentication system for access to the journal, the search interface for journal content, the journal content's format (e.g., HTML, PDF), etc.? Almost none of the above are actual services/resources provided by the library. Rather, they are particular to the various vendor systems to which the library subscribes. Asking users what they "think about a library service," therefore, is quite complex in the networked environment and points to a number of methodological problems that require resolution. Simply put, the outcomes and service quality evaluation tools of today are not adequate to engage in meaningful assessment activities for library network-based services and resources. There is much research required in this area.

5 Connecting the Dots

This paper identified four key evaluation frameworks that researchers and practitioners can use to assess their digital library resources and services. Librarians and researchers, however, tend to engage in single-focus evaluation strategies – or, when they do engage in multiple assessment techniques, do not combine the different forms of data that these evaluation efforts provide to develop a comprehensive view of the digital library.

Within the strategies there is also a need to develop and/or continue to develop measures, methods, and indicators that address specifically digital library resources and services. While there are various initiatives that continue to do so within certain evaluation frameworks – e.g., outputs and performance measurement [12] – there is an overall lack of techniques in the service quality and outcomes assessment approaches for network-based services and resources.

Thus, there are several key issues that librarians and researchers should consider as they move forward to develop digital library evaluation efforts:

- Move beyond output type measures so as to understand the perceived quality, value, and impact of the digital library resource/ service;
- Continue to develop measures, methods, and indicators of digital library resources and services. These may not have any counterpart in traditional librarianship, so the past may not serve as a guide for the future;

- Consider measures and approaches that include the user – not just throughput, process, and item measures;
- Develop digital library evaluation goals, objectives, and strategies as the resource/ service is developed rather than post-implementation. Evaluation goals and objectives should integrate with the overall digital library initiative goals and objectives; and
- Engage in multiple assessment activities to develop a composite view of the digital library – depending on one form of assessment will provide only a limited view of the digital library resource/service.

Considering these issues will enable librarians and researchers to better understand and make resource investment decisions regarding their digital library efforts.

6 References

[1] Lakos, A. The missing ingredient: Culture of assessment in libraries. *Performance Measurement and Metrics*, 1(1), 3-7, 1999. Retrieved March 20, 2003 from http://www.aslib.com/pmm/1999/aug/opinion.pdf.

[2] Bertot, J.C. Service quality in the networked environment. *Library Trends*, *49*(4): 758-775, 2001.

[3] Bertot, J. C. & Snead, J. T. Social measurement in libraries. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement.* San Diego, CA: Academic Press, in press.

[4] Cook, C. & Heath, F. M. Users' perceptions of library service quality. *Library Trends*, 49 (4): 538-584, 2001; Cook, C., Heath, F.M., & Thompson, B.T. Score norms for improving library service quality: A LibQUAL+™ study. *Portals: Libraries and the Academy*, 2(1): 13-26, 2002.

[5] Self, J. From values to metrics: Implementation of the balanced scorecard at a university library. *Performance Measurement and Metrics*, *4*(2), 57-63, 2003.

[6] Kaplan, R.S. & Norton, D.P. *The Balanced scorecard: Translating strategy into action.* Boston, MA: Harvard Business School Press, 1996.

[7] Bertot, J.C., & McClure, C.R. Outcomes assessment in the networked environment: Research questions, issues, considerations, and moving forward. *Library Trends*, *51*(4): 590-613, 2003.

[8] Cook, Heath, & Thompson, 2002.

[9] Hernon, P., and Dugan, R.E. *An action plan for outcomes assessment in your library.* Chicago, IL: American Library Association, 2002.

[10] Cool, C., & Spink, A. Issues of context in information retrieval (IR): an introduction to the special issue. Information Processing & Management, 38(5): 605-611, 2002; Ke, H.R., Kwakkelaar, R., Tai, Y.M., & Chen, L.C. Exploring behavior of e-journal users in science and technology: Transaction log analysis of Elsevier's ScienceDirect OnSite in Taiwan. *Library and Information Science Research*, 24(3): 265-291, 2002.

[11] Brophy, P., Fisher, S., & Clark, Z. *Libraries without walls 4: the delivery of library services to distant users.* London, UK: Facet Publishing, 2002.

[12] A number of standards initiatives are incorporating research regarding e-metrics. These include ISO 2789 *Library Statistics*, ISO 11620 *Library Performance Indicators*, and Z39.7 *Library Statistics* (see http://www.niso.org/emetrics). The e-metrics are based largely on research by Bertot, J.C., McClure, C.R., & Ryan, J. *Statistics and performance measured for public library networked services.* Chicago, IL: American Library Association, 2001; Shim, J.W., McClure, C.R., Fraser, B.T., Bertot, J.C., Dagli, A., and Leahy, E.H. *Measures and statistics for research library networked services: Procedures and issues: ARL e-metrics phase II report.* Washington, DC: Association of Research Libraries, 2001; and Brophy, P., Clarke, Z., Brinkley, M., Mundt, S., Poll, R. *Equinox: Library performance measurement and quality management system performance indicators for electronic library services.* Retrieved January 20, 2004 from http://equinox.dcu.ie/reports/pilist.html.