Pre-publication draft only. Final publication may be cited as:

Planning A Digital Library Evaluation With Logic Models

Michael Khoo
The iSchool at Drexel University
3141 Chestnut Street, Philadelphia, PA 19104, USA
michael.khoo@ischool.drexel.edu

Sarah Giersch
Consultant, National Science Digital Library
Cathedral Station, PO Box 2112, New York NY 10025, USA
sgiersch@gmail.com
Planning A Digital Library Evaluation With Logic Models

INTRODUCTION

Effective digital library evaluation integrates a wide range of organizational, methodological, and technical factors. At the same time, evaluation work can be constrained by such factors as limited funds, limited time, and limited access to users. How can evaluators work within these parameters to produce a realistic evaluation work plan that yields actionable results? In this chapter we introduce a checklist that outlines some key practical steps associated with planning for digital library evaluation, with particular emphasis on developing a logic model. The checklist places evaluation planning at the intersection of four activities: identifying strategic evaluation questions, selecting appropriate evaluation methods, negotiating stakeholder buy-in, and obtaining adequate resources to support the evaluation. We will focus on the process of using strategic evaluation questions to develop a logic model and then identify the synergies between this activity and the other three activities (some of which are also covered in more detail elsewhere in this book).

An underlying assumption of the checklist is that digital library evaluation consists of a series of longitudinally linked activities, in which the success of later evaluation activities depends significantly on the preparation that went into the earlier activities. Paying close attention to evaluation during the planning stages – including identifying evaluation questions, and specifying an evaluation budget – can be crucial to supporting substantive evaluation efforts later in the life of the digital library. This chapter introduces two logic model-based evaluation methods that can be used to support for planning formative and summative digital library evaluation, especially during the initial stages of planning a digital library. The methods are presented within the context of our own digital library evaluation experiences,
and draw upon a number of existing online evaluation guides (e.g. Frechtling 2002, Reeves et al. 2006).

DIGITAL LIBRARIES ARE COMPLEX SYSTEMS, AND PLANNING A DIGITAL LIBRARY EVALUATION IS A COMPLEX ACTIVITY

Digital libraries are composed of a wide range of social, technological, organizational and other phenomena, embedded in a variety of external contexts (e.g., social, political, economic, organizational, etc.). They can be modeled as sociotechnical systems, comprised of many components linked in complex and mutually constitutive ways (Bishop et al., 2003). The external contexts in which they are embedded also interact in mutually constitutive ways, both with each other and with digital libraries. For example, factors such as a school’s bandwidth, the number and age of its computers, the presence or absence of technological support staff, and the availability of professional development, can all affect educational technology use in ways that have nothing to do with the technology itself (National Education Association, 2008).

As sociotechnical systems, digital libraries exhibit many of the properties of complex adaptive systems, including unpredictable development and emergent properties (Marchionini, 2000). While new digital libraries may begin with broadly similar aims and activities – such as developing resources, collections, metadata, catalogs, search engines, and Web interfaces – the complex local conditions in which they are situated shape their growth and development in unpredictable and unique ways. With the recent development of a range of web-based tools that support personalized services, seamless content creation and publication, and re-use and re-combination of multiple data formats, which are now being offered through digital libraries, there is no longer a ‘one size fits all’ version of the digital library (Lagoze et al., 2005; McArthur and Zia, 2008; Miller, 2006).
The complexity of digital library evaluation work is illustrated by an example from the National Science Digital Library (NSDL\textsuperscript{1}). One of the conditions for receiving NSF funding is that NSDL projects are expected to conduct evaluation work and to report the data to NSF via annual reports. A 2006 survey of evaluation practices amongst NSDL projects (Bartolo et al., 2006) asked the projects about their engagement with a range of evaluation activities, from developing evaluation questions, metrics and instruments, and identifying evaluators, to collecting data and analyzing evaluation data, and disseminating evaluation reports and sending evaluation findings to NSF. The survey assumed that evaluation activities occurred as part of an integrated workflow, with each stage of the workflow generating a foundation upon which the next stage can be developed. For example, the drafting of evaluation questions supports the development of measures for project success; project metrics support the design of data collections methods; methods enable data collection; and data collection supplies material for analysis, reports and articles. In the case of NSDL, the survey data revealed that as projects moved along this evaluation workflow, the likelihood that they would abandon their evaluation work increased, and the implementation of evaluation activities exhibited a downward attrition trend along the evaluation workflow. While almost all projects started out with an evaluation plan, fewer projects completed evaluation and reported data to NSF or to professional audiences (see Table 1, Figure 1). This decline occurred both with completed projects and also with projects that were still being funded and which had yet to complete their work. The survey also identified a number of resource barriers to carrying out

\textsuperscript{1} The NSDL acronym is applied to three complementary entities: 1) NSDL is a funding program of the National Science Foundation (NSF); 2) NSDL is a body of projects and a community of Principal Investigators who receive funding; and, 3) NSDL is a portal (NSDL.org) whereby users access the work of funded projects.

As a multi-year funding program of NSF, NSDL has a mission to develop a wide range of science, technology, engineering and mathematics (STEM) educational resources, tools and services within a distributed digital library structure, with the aim of increasing literacy and interest in STEM topics, preparing a STEM workforce, and supporting the advance of knowledge and the solving of real-world problems (NSDL, 2007). By 2006 the NSDL program had awarded over $150,000,000 to over 200 projects, which were lead by researchers and faculty at universities in the United States. Projects are funded from 1 to 4 years to develop content, infrastructure, services and outreach activities, which contribute to the overall growth and use of online STEM resources by teachers and learners of all ages. Online NSDL STEM resources can be accessed at http://nsdl.org (see Chapter X for further background on NSDL),
evaluation, including lack of budget, lack of staff, lack of evaluation capacity (including difficulties in hiring external evaluators appropriately qualified in either evaluation and/or digital library expertise), and lack of time. Each of these constraints could act on each of the evaluation stages, causing activities to stall at that particular stage.

The example of NSDL projects’ experience with completing integrated evaluation activities illustrates the importance of some of the longitudinal factors that can affect digital library evaluation. Sufficient resources are required to support each stage of the digital library evaluation workflow, and lack of resources at one workflow stage will affect not just that stage of the workflow but also subsequent stages. Given the complexity of digital libraries, evaluation works best when implemented as a holistic and integrated activity that extends throughout the duration of digital library development. While this work requires a larger initial amount of effort to prepare, a well-planned, longitudinal evaluation strategy has the potential to yield more useful data for the same amount of resource expenditure than an unplanned, ad hoc and un-integrated approach.

**An Evaluation Planning Checklist**

The rest of this chapter outlines a checklist of four key activities involved in planning and organizing a digital library evaluation initiative both during proposal writing to fund digital library project, and in the context of conducting evaluations of already-funded digital libraries. More broadly, the checklist is both a model of what digital libraries are, and also a model of what evaluators often do. Specifically, it is based on the assumption that evaluation
Evaluation of Digital Libraries

consists not just of the evaluation requirements outlined in program solicitations, but also of a series of integrated sets of practices that support the development and implementation of a holistic model of the system under evaluation. An underlying assumption of the checklist is that later evaluation work needs to be supported by the earlier identification of strategic evaluation questions, and by the pre-allocation of resources to address those questions.

The components of the checklist are:

1) Identifying key formative and summative evaluation questions (this will be the main focus of the discussion, and will include a discussion of logic models)

2) Selecting appropriate methods (such as web metrics, usability and HCI work, surveys, interviews and focus groups, ethnography, etc.)

3) Identifying important stakeholders (funders, managers, developers, users, etc.) and establishing stakeholder buy-in

4) Budgeting for evaluation work in the initial project proposal

[INSERT FIGURE 2 ABOUT HERE]

These four early stage evaluation planning activities are represented in Figure 2 as a series of overlapping circles, intended to convey the fact that while these activities involve separate considerations and tasks, they are closely linked and interdependent. For example, the drafting of strategic evaluation questions is closely related to stakeholder buy-in, shapes the methods chosen, and is affected by the budget available for evaluation activities. Again, stakeholder buy-in has a close relationship both with shaping the strategic questions, and setting an evaluation budget. Taken together, the areas of activity in Figure 2 can provide substantial support for the evaluation work carried out during the actual project, and it will be harder for projects to begin implementation of an evaluation plan late in the life of a project
without having addressed these activities (although this does not mean that commencing evaluation at during the latter stages of a project is impossible).

PLANNING FOR SUMMATIVE AND FORMATIVE EVALUATION

In the rest of this chapter we expand upon the four evaluation planning activities outlined in Figure 2, focusing mainly on the identification of strategic summative and formative evaluation goals (in the following section we also briefly cover some of the linkages between this activity and selecting methods, obtaining buy-in, and setting a budget).

Planning summative evaluation questions and goals

Evaluation resources are used most efficiently when formative and summative questions address specific strategic issues for a project, with the outcomes from these evaluation questions being used to guide project development in useful ways (Reeves et al., 2006). The development of strategic evaluation questions can be supported by integrating evaluation planning into the overall project research plan during the early stages of the project. Proposal writing and evaluation planning can be mutually supporting activities, with the description of the proposed work shaping the development of the evaluation plan, and the evaluation plan providing useful feedback on the feasibility of the project proposal. In practice, this means working with the emerging research proposal, maintaining communication with managers, developers and other project members, and eliciting from project members clear descriptions of what they are planning to do.

Using Logic Models To Plan Summative Evaluation

The identification of summative evaluation questions can be supported through the use of logic models: graphical, high-level, system-based conceptual models of a project (such as a digital library) that describe how that project is funded, what it does, and what it hopes to
achieve. A logic model does this principally by linking project inputs to activities, and then to project outcomes (Renger and Hurley, 2006; Morge et al., 2008).

Logic models generally describe four high-level sets of project activities: (1) project inputs, that is, the various funding sources and revenue streams that support project activity; (2) project activities, that is, the processes that the project engages in on a day-to-day level; (3) short-term outputs, that is, the immediate results; and (4) long-term project outcomes, “broader and more enduring impacts on the system” (Frechtling, 2002, p. 17) (Figure 3a). These activities are longitudinally articulated; for instance, project inputs support the project activities, and the project activities generate the short-term and long-term outcomes (however, there are a number of variations on these themes, e.g. Centers for Disease Control and Prevention, 2008; Frechtling, 2002; W. K. Kellogg Foundation, 2004; Taylor-Powell and Henert, 2008).

As a hypothetical example, a possible outline of a logic model for an educational digital library project is presented in Figure 3b. Here, the input is NSF funding; the project’s activities include creating exemplary resources and metadata, developing a web interface, carrying out comprehensive outreach activities, and building strategic partnerships with educational institutions; expected short-term outputs include the development of themed collections, public awareness of the library, and growing numbers of satisfied users of the library and its web site; while long-term outputs include the successful incorporation of library materials into school curricula, and an increase in students’ knowledge of the library’s subject domain.
A major benefit of a logic model approach, if it is used in conjunction with discussions with project managers and developers, especially at the beginning of a project, is that it supports stakeholders to identify and articulate some of the basic components, linkages and aims of a project. Particularly, by prompting a project to clearly define goals and outcomes, it helps stakeholders to begin defining exactly how they would to achieve those outcomes using the available resources (such as project funds, and available skills and expertises). Frechtling (2002) for example suggests using the logic model approach to work backwards from the expected outcomes, as a way of determining what activities would need to take place in order to achieve these outcomes.

**Planning formative evaluation questions and goals**

Evaluation planning should also include a number of formative evaluation measures – for instance, the number of resources generated, the quality of those resources, the quality of metadata, Web interface usability, and so on – which can both produce useful data for ongoing iterative improvement of a digital library and can provide a solid foundations upon which to build summative evaluation. While logic models do identify inputs, activities and outputs at the overall project level, they do not necessarily offer support for understanding how the internal linkages in a project support project activities in converting these inputs into outcomes. In a summative logic model, the internal functions of a project are often ‘black boxed’ and obscured from view. How therefore can evaluators identify suitable library components for formative evaluation (especially given the internal complexity of digital libraries)? In this section we describe how the logic model approach can also be adapted to guide the planning formative evaluation questions, and as an example, we describe how such an approach was implemented with the National Science Digital Library.

*Using Logic Models to Plan NSDL Evaluation*
As described above, NSDL consists of a series of distributed projects engaged in a number of digital library building activities, such as developing online STEM resources, technology architecture, and web-based services. Selecting common formative metrics for evaluation across this heterogeneous mix of components was not easy. One common unit that did link many NSDL projects across the context of the NSDL program was that of the digital, or online, resource. Particularly, many NSDL project activities were involved with creating resources, adding them to collections and cataloguing them, making them available on the Web, or supporting their use in the classroom. NSDL therefore developed a logic model-based formative evaluation approach – the model was called the ‘resource lifecycle’ – that focused on the production of digital library resources in NSDL as a way of identifying internal NSDL components and linkages for formative evaluation. Having defined digital resources as the unit of analysis, the model tracked resources through various stages of NSDL operations, from creation to the moment of use, and then beyond to the moment of redesign and improvement. Overall, the model identified several basic sequential areas of activity in which various NSDL projects acted on or modified digital resources, including

- resource creation
- collection creation
- resource retrieval
- resource use and reuse

Each of these stages was then modeled as simple logic model in itself, with inputs, activities and outputs, and then the stages were joined, with the outcome from one stage forming the input of the next stage (two example stages of the model are shown in Figure 4).

[INST FIGURE 4 ABOUT HERE]
Taken together, these stages constitute a ‘production line’ model, each stage of which involves transforming digital resources, in the process adding value and utility to those resources. For example, a resource that has been reviewed for pedagogical effectiveness, scientific accuracy, and technological functionality is more valuable than a resource that has not; a resource described by accurate metadata is more valuable than a resource that is not; a resource embedded within a powerful search and discovery tool is more valuable than one that is not; and so on. It is the cumulative outcome of all stages of the model that ensures the development of exemplary resources for NSDL; conversely, poor quality operations at any one stage of the cycle can have adverse effects downstream in the workflow. For instance, poor metadata quality affects search and retrieval; poor resource quality affects classroom use; poor usability affects the user experience (and presumably adversely affects repeat visits); and so on.

The resource life-cycle model had several advantages for the NSDL evaluation work. The model provided a unifying narrative within which to begin examining the wide range of disparate projects within NSDL. Components of the model’s stages – such as resource quality, web site usability, etc. – identified strategically useful places at which to carry out formative evaluation of individual project components, and also provided a coherent overview of how different evaluation activities at different parts of NSDL – such as webmetrics, and user interviews – might be related. In the form of documentation (in the case of NSDL, a white paper: Khoo, 2006), the model provided a useful ‘boundary object’ (Star and Griesemer, 1989) that could be used to explain the rationale for evaluation work to various stakeholders, including NSDL managers, individual NSDL projects, and the National Science Foundation (as well as to obtain buy-in and support from those groups – see below).

One disadvantage of the approach was that it could be resource-intensive, and required expertise in a wide range of digital library areas. Further, some of the potential formative
evaluation areas identified were difficult to investigate, for a variety of reasons, including access to database systems (e.g. in the case of metadata), access to servers (in the case of web metrics), access to users (in the case of classroom use), etc. Finally, as noted above, many digital libraries are now introducing services that are not resource-centric and are therefore more difficult to model using this approach and which may require newer logic models to describe them. Despite these limitations, however, the model clarified and organized some of the disparate dimensions of NSDL in a useful way, provided coherence and focus for a range of formative evaluation concepts and practices, and helped in the identification of important internal linkages within NSDL where formative evaluation resources could be targeted.

**Summary**

Logic models are useful tools, which support planning at the earliest phases of a digital library project, while also providing a means of integrating evaluation activities into the project plan. They provide a framework for identifying summative evaluation questions that are tied to project outcomes and also for identifying formative evaluation questions that are linked to digital library development activities throughout the life of the project. The planning of formative and summative evaluation questions does not take place in the abstract, however, and the following section describes how the development of logic model-based evaluation questions during the project proposal stage is closely linked with the activities of identifying data collection methods, promoting buy-in amongst stakeholders, and developing an evaluation budget for a project.

**RELATIONSHIP TO METHODS, BUY-IN, AND BUDGET**

In this section we briefly examine the relationship between identifying evaluation questions, and the three other areas of evaluation planning outlined above: selecting data collection methods, negotiating buy-in, and budgeting for evaluation activities.
Methods

Developing formative and summative evaluation questions supports the selection of appropriate tools and techniques for subsequent evaluation activities (a range of different quantitative and qualitative methodologies is discussed in the following chapters). If multiple questions have been identified, a project will ideally be able to deploy several ongoing evaluation methods in a 'mixed methods' approach (Ryan et al., 2001), or what Marchionini et al. (2003) refer to as an iterative and ‘multifaceted’ approach to digital library evaluation and development, in which “multiple data views are essential to guide design and to help us to understand the impact of digital libraries.” If one or more evaluation approaches are used to address the same question(s), confidence in evaluation data will be increased if the methods can be triangulated.

Logistical factors for methods (often related to budget issues – see below) include: the presence of evaluation skills on the project team, and whether or not to hire an external evaluator (for a task such as usability this may be expensive, but probably cheaper in the long run than paying project staff to become familiar with a new evaluation technique); comparing the types of tools to be used (choosing for instance free web metrics tools, or more expensive but also more powerful proprietary tools); choosing the mix of laboratory and naturalistic studies of users (laboratory studies can be easier and quicker to organize, however they may miss some of the more subtle data that can be obtained from observing digital library users ‘in the wild’ (e.g. Khoo and Ribes, 2005)); and, comparing time scales and cost (different methods, such as surveys and ethnography, can differ significantly in resource requirements in terms of time and money).

Buy-in

As the formal evaluation plan proceeds, it is useful to begin obtaining buy-in for the proposed work from stakeholders such as project members, funding agencies, users, and others who
may have an interest in the outcome(s) of the project. Buy-in can be supported by eliciting and addressing stakeholders’ models of and expectations for a project. Note that different stakeholders may have different priorities for a project, and may advocate for different evaluation metrics and measurements; for instance, funding agencies may have different evaluation priorities than the projects they fund. (Agencies’ general requirements, which can be used to help frame strategic evaluation goals, are usually described in program solicitations, and agencies may make available other evaluation materials elsewhere on their Web site; for instance, the NSF provides Frechtling, 2002; see chapter X.)

As was also noted in Chapter X, there is a danger that projects may see evaluation as a burden imposed by funding agencies, in which evaluation data and outcomes may be unwelcome or threatening to a project (e.g.; as evidence that a project is ‘not succeeding’); and the evaluation relationship can potentially become an adversarial one (Frechtling 2002, Reeves et al., 2006). Buy-in can be supported by drafting and circulating proposals for discussion, for instance in memo or white paper form, and by having stakeholders discuss the proposed plan. Documents generated at this stage can also be used to support the description of the evaluation in the project proposal, and also serve as the basis for discussion of ongoing evaluation as the project progresses.

Canvassing various stakeholders’ conceptions of the proposed project goals and outcomes can require significant effort, for several reasons. A project’s aims can remain in flux until the submission of the proposal (or even after). There may be different stakeholder (e.g. agency, developer, and user) understandings of a project, or no commonly articulated overall goals for the project, or disagreement amongst project members on the key strategic goals of a project, or differences between projects and funding agencies with regard to appropriate evaluation measures, or insufficient clarity on the role or purpose of evaluation for a project. Different stakeholder groups may also hold (often without realising it) different mental models of the
project’s goals, and these differences need to be untangled and reconciled (Khoo 2005). Even in these circumstances, it is still useful to continue to identify preliminary questions from various stakeholders that can be used as the basis for beginning the evaluation activities.

**Budget**

Evaluation is “a people-intensive process, and therefore, most of the money spent on evaluation usually will be for dedicated evaluation personnel and/or external consultant costs” (Reeves et al., 2006, p. 21). The proposed evaluation work must be included in the budget (and not borrowed from 'left-over' funds) or else it will be difficult to complete. The planning stages of a project proposal are a good point at which to start developing an appropriate evaluation budget since it may become harder later in the life of a project to request funds for evaluation. Opinions differ as to how much should be included in a budget for evaluation; Reeves et al. (2006) for instance recommend 5%-10% of the overall budget. Some government agencies recommend that 10-15% of the overall budget be spent.

An evaluation budget is often a miniature version of the full proposal budget, and it should be accompanied by a detailed budget justification, that includes (where appropriate) items such as staff time, external consultants’ fees, participant stipends for laboratory studies, evaluation hardware and software costs, and travel (both for any evaluation work and also for the presentation of results), etc. (Horn, 2001). The availability of evaluation budget resources can impact the selection of different types of evaluation methods (e.g. one-shot quantitative versus longitudinal qualitative evaluation), and different types of evaluation tools (such as web metrics packages).

If evaluation planning is left to the end of the proposal writing, the evaluation budget runs the risk of being minimised at the expense of other project components, or reduced as final adjustments are made to the full budget proposal. Lack of funding at this stage of the evaluation work could generate problems later in the life of the project, when it may prove
difficult or impractical to either carry out the required work, or to request additional funding. On the other hand, there are useful synergies between formulating the evaluation questions as part of a logic model (see previous section) and creating the evaluation budget. Having a sense of what is realistically achievable for evaluation in terms of the entire project's resources can help to guide the discussion on the research questions and what is hoped to be achieved overall by the project.

**SUMMARY**

Digital libraries are complex phenomena, and the evaluation of digital libraries is a complex activity that requires considerable planning if it is to be accomplished successfully. We have argued in this chapter for the importance of including digital library evaluation planning at the very earliest stages of a project, as a way of supporting the overall development of the project proposal, establishing the evaluation questions, and making the case for an appropriate allocation of resources to support the investigation of these questions. To help frame this work, we identified four key inter-related areas of evaluation planning, (i) formative and summative evaluation questions, (ii) suitable methods for investigating these questions, (iii) stakeholder buy-in for the proposed work, and (iv) a suitable budget to support the work. We have also discussed the application of a logic model approach to the identification of strategic formative and summative evaluation questions. Logic models analyze digital libraries as systems of coordinated activities, with specified inputs, goals, and outcomes, in which each activity incrementally adds value and utility, and in which the output of any one activity forms the input of one or more succeeding stages; and they can be used to identify key strategic linkages within a project and their upstream and downstream dependencies, which can then be used as loci for focused and coordinated evaluation activities.

**REFERENCES**


Table 1. % Reported Types of Evaluation Practices in NSDL (n = 28)

<table>
<thead>
<tr>
<th>Evaluation Activity</th>
<th>% Respondents Reporting Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning activities</td>
<td></td>
</tr>
<tr>
<td>Evaluation plan included in NSF proposal</td>
<td>89</td>
</tr>
<tr>
<td>Designated internal/external evaluator(s)</td>
<td>83</td>
</tr>
<tr>
<td>Developed formal metrics of project success</td>
<td>46</td>
</tr>
<tr>
<td>Developed formal methods and instruments</td>
<td>71</td>
</tr>
<tr>
<td>Implementation activities</td>
<td></td>
</tr>
<tr>
<td>Implemented evaluation plan &amp; collected data</td>
<td>75</td>
</tr>
<tr>
<td>Analyzed evaluation data</td>
<td>75</td>
</tr>
<tr>
<td>Dissemination activities</td>
<td></td>
</tr>
<tr>
<td>Presented findings at workshops</td>
<td>70</td>
</tr>
<tr>
<td>Made findings available in internal report</td>
<td>64</td>
</tr>
<tr>
<td>Presented or published findings at conferences</td>
<td>62</td>
</tr>
<tr>
<td>Sent evaluation findings to NSF</td>
<td>48</td>
</tr>
<tr>
<td>Published findings in journals</td>
<td>23</td>
</tr>
</tbody>
</table>
Figure 1: % Reported Types of Evaluation Practices in NSDL (n = 28)
Figure 2: Four key activities involved in planning and organizing a digital library evaluation initiative
Figure 3a: Outline format for a logic model format (after Frechtling, 2002)

Figure 3b: A basic logic model for evaluating an educational digital library
Figure 4: A longitudinal logic model for formative evaluation of a digital library