DIGITAL LIBRARY EVALUATION: A LONGER VIEW

Rich Gazan
Department of Information Studies, University of California Los Angeles
USA

When an animal bred in captivity runs, flaps or slithers into the wilderness for the first time, scientists do not simply wave goodbye, wish it well and move on to the next project. Usually the animal is tracked and observed for a long time thereafter—the point of release is where the research really begins. However, this has been less true for grant-funded digital library projects. Administrative and institutional barriers tend to confine evaluation to the funding period when the system is being designed, and when its uses and effects can only be guessed at. In this short paper I will discuss a digital library project in environmental science where some of these evaluation issues arose, the value of eliciting narrative data from users, and some ways to take better advantage of existing but underused mechanisms for long-term evaluation of digital libraries.

Viewing information technologies and society as inextricably co-determined makes it necessary to gather data both on the design process and on how the finished systems are used by people—often in quite unexpected ways. Iterative design and formative evaluation are ways to introduce user feedback into ongoing design (Borgman et al. 2001), but while these strategies can serve as an effective translation layer between users, designers and builders, the focus is on how an evolving system *might* be used. How it is *actually* used can be more accurately determined when the system is in the wild, when the designers and their formal evaluation instruments have gone away.

Digital libraries are social entities. They tell stories about a culture, a science, a place or a time, through the items represented and organized in the collection, and in use, they help generate new stories as well. Manovich (2001) has advanced the idea of database and narrative as two ends of same continuum; both structure information, but where a database aims for access, a narrative aims for psychological immersion in the story. From an institutional standpoint, digital libraries are rarely ends in themselves. The social missions of the funding organizations—the stories they wish to tell—often drive the creation of digital libraries, but how well a finished system supports the mission can't be fully evaluated until the digital library has had some time to develop a user base.

This line of research began with a participant observation of the design of a digital library of environmental science collections. Funded by an Library Services and Technology Act (LSTA) grant, the project involved diverse designers, content, metadata and institutional participants (Gazan 2005). As one of the participants in the environmental science digital library noted, "Part of the art of grant writing is interpreting vague language in a way that lets you do what you want to do," or in other words, making project proposals and achievements dovetail with the mission and goals of the funding agency. These goals included outreach and evaluation, so a usability component was included in the grant proposal. My initial role in the project was to develop the instruments and conduct the evaluation.

Understandably, digital library designers tend to create evaluation instruments that demonstrate in a measurable way the work they've done, a tacit statement of the value produced for the grant funds received. For example, the design of the interface, the appropriateness of descriptive metadata, and user success at canned search tasks are classic evaluative measures. However, in an analysis of the Perseus Digital Library (PDL), Marchionini (2000; p. 328) writes:

> "Operational data are powerful components in a chain of inferences that address impact but the PDL evaluation illustrates the value of anecdotes and "stories" that illustrate new effects, i.e., how DLs augment existing capabilities with new ones. These augmentations garner public support for a DL and should not be underestimated in assessing impact...Integrating multiple views is more naturally done with narratives than summary statistics and integrating these forms of evidence can aid in assessing complex change."

In the evaluation, statistical and demographic data were not hard to come by. Observation, interviews, document analysis, narrative analysis and social network analysis were effective ways to construct as complete a picture as possible about the interactions of the designers. But the usability component needed to be completed at the same time as the digital library, in time for the results to be included in the final LSTA grant report (though the exact LSTA grant regulations vary by state, most LSTA grantees must submit a final grant report within 15-30 days of project completion). Participants in the usability study were evaluating a still-evolving system, and had to project potential uses into their responses to open-ended interview questions. By the time they might integrate the collections into their

professional lives, and perhaps surprise themselves with unexpected dimensions of usefulness, the evaluation would be long over.

This is certainly not to say that there is no such thing as continuing evaluation of grant-funded digital libraries. But the mechanisms for long-term evaluation are usually little more than Web forms or e-mail links, not organic components of the digital library. In practice, designers are more concerned with present and future digital library projects than with continuing evaluation of those of the past. Digital library researchers have an opportunity to conduct this longer-term research, to question and reveal the impacts of digital libraries as social entities, and to apply the resulting knowledge to future projects.

The good news is that in some situations, a mechanism for longer term evaluation already exists. For example, the State Library of Ohio's LSTA grant process makes use of a "year-after" evaluation form (http://winslo.state.oh.us/publib/lstayraft.html), which includes questions such as "Did the project produce any unexpected results?" and encourages narrative ("Please provide a success story of how your project has impacted someone's life or had a positive impact on the community."). This is precisely the sort of open-ended data collection instrument that can reveal how a digital library is actually being used—but even this document is only the length of a one-page questionnaire. Subsequent research will attempt to evaluate the usefulness of this and other longer-term evaluation instruments, and how they might be expanded.

While longer-term evaluation should include quantitative data such as transaction logs and perhaps a list of external sites that link to collection content, evaluation instruments should also be open-ended, designed to encourage narrative expressions of unexpected use, the kind of data that reveal to funding agencies the real impact of their grants. Iterative digital library design philosophies have always had at their core the sense that user input should feed back into ongoing system design. I propose here simply a wider iterative design circle, one that allows for more naturalistic data about longer-term use to be fed back into future systems.

In sum, lessons learned that will be explored in future research include an increased emphasis on:

- long-term evaluation of digital libraries
- narrative data; allowing users to tell stories
- unexpected uses
- evaluating social outcomes, not just the design process or product
- linking findings more directly to the higher-level goals of funding agencies

References

Borgman, Christine L.; Leazer, Gregory H.; Gilliland-Swetland, Anne J. & Gazan, Rich (2001). Iterative Design and Evaluation of a Geographic Digital Library for University Students: A Case Study of the Alexandria Digital Earth Prototype (ADEPT). *Proceedings of the Fifth European Conference on Digital Libraries*, 4-8 September 2001, Darmstadt, Germany.

Gazan, Rich (2005; in press). Imposing Structures: Narrative Analysis and the Design of Information Systems. *Library & Information Science Research* 27(3).

Manovich, Lev (2001). The Language of New Media. Cambridge, Mass.: MIT Press.

Marchionini, Gary (2000). Evaluating Digital Libraries: A Longitudinal and Multifaceted View. *Library Trends* 49, 304-333.