

Web-Based Tools for Learning Object Evaluation

John C. Nesbit and Jerry Li
School of Interactive Arts & Technology
Simon Fraser University
Surrey, British Columbia, V3T 2W1

ABSTRACT

The emergence of large repositories of web-based learning resources has increased the need for valid and usable evaluation tools. This paper reviews current approaches to learning object evaluation and introduces eLera, a set of web-based tools we have developed for communities of teachers, learners, instructional designers and developers. Compatible with current metadata standards, eLera provides a learning object review instrument (LORI) and other features supporting collaborative evaluation. eLera is also designed to assist researchers to gather data on evaluation processes and perceptions of learning object quality.

Keywords: e-learning, resources, reviews, assessment, eLera, LORI, convergent participation

1. NEED FOR LEARNING OBJECT EVALUATION

The primary feature distinguishing learning objects from other educational software is perhaps their ready availability through web-based repositories or collections that can be searched with standardized metadata. Because user communities such as elementary school teachers and university instructors have different requirements and expectations, specialized repositories are emerging that are interlinked by metadata and interoperability standards. Within the next five years, the U.S. National Science Digital Library (NSDL) is predicted to grow to include as many as 100,000 collections representing over a million learning objects [1].

We believe that the effectiveness of most online learning resources is severely limited because they do not follow design principles established by educational research, and have not been subjected to formative user testing. Thus, there is an immediate need for methods and tools to facilitate the development, dissemination and retrieval of high quality resources.

The scope of the quality problem is such that a variety of evaluation methods is required. For example, teachers and learners may only be induced to evaluate if the instruments are uncomplicated and demand very little additional time and effort. In contrast, professional e-learning design teams require detailed instruments that

support systematic assessment at different stages of product development.

2. APPROACHES TO EVALUATION

Our research on collaborative learning object review is one branch of a growing body of practice driven by a renewed awareness of learning resource evaluation as a critical complement to learning design and an essential factor in resource selection.

CLOE

The Co-operative Learning Object Exchange (CLOE), jointly developed by seventeen Ontario universities to facilitate the design and application of multimedia-rich learning resources, operates a structured review process [2]. A learning object submitted to CLOE is first examined by the editor-in-chief to decide if it meets specified technical requirements. The object is either returned to the author for revision, or forwarded to an instructional design reviewer and content reviewers. The instructional design reviewer gives a binary decision (go or no-go). Normally, content is reviewed by two content reviewers. When they disagree, the decision to approve the object falls to a third content reviewer. CLOE provides three broad evaluative dimensions: quality of content, effectiveness as a teaching/learning tool, and ease of use.

MERLOT

MERLOT (www.merlot.org) is a repository containing educational resources classified into seven broad subject categories: Arts; Business; Education; Humanities; Mathematics Science and Technology; Social Sciences. Each category is divided into sub-categories to form over 500 subjects in total. MERLOT provides tools for both individual member comments and peer review. In both types of evaluation, resources are rated on a five-point scale. An object may be selected for peer review by an editorial board representing one of 14 discipline-based communities. The commonly practiced peer review process in MERLOT is similar to that for CLOE, except that there is no provision for an instructional design reviewer.

DLNET

NSDL is a federated repository that includes DLNET, the Digital Library Network for Engineering and

Technology (www.dlnet.vt.edu). As a subject taxonomy, DLNET uses an adapted version of the INSPEC taxonomy of scientific and technical literature (www.iee.org/Publish/Inspec).

Like MERLOT, DLNET maintains a two-tier evaluation system allowing review by expert peers and “public review” by users at large. But it differs from MERLOT in that an object is not published in the repository until it has been approved by peer review. On the other hand, the function of public reviews is to provide an ongoing ranking of published objects by users.

DLNET reviewers fill out an instrument containing a single comment field and 11 items rated on a 5-point scale. DLNET currently allows members to publish multiple reviews on the same learning object and currently provides no statistical aggregation of rating data.

Models and Tools

These three examples (CLOE, MERLOT, DLNET) demonstrate a common model with variations. Each is formed from (a) a searchable database of learning resource metadata that more or less conform to the IEEE learning object metadata standard; (b) a subject taxonomy constituting one component of the metadata; (c) evaluation criteria in the form of guidelines or a structured instrument; (d) a process for conducting and publishing reviews including restrictions on who can review; (e) a structured form in which all reviews are published. We view such systems as socio-technical phenomena that can be analysed and empirically researched by studying the operation of different designs.

The two tiers of individual user and peer review that we see in MERLOT and DLNET mirror the two different types of consumer product evaluation systems that have proliferated on the Web. For example, at one video game review site (www.pcgamereview.com), any user can register to rate and comment on three quality dimensions (gameplay, graphics, sound) of a video game. Similarly, at a general consumer product review site (www.reviewcentre.com), any user can rate products on the two dimensions of “quality” and “value for money”, as well as record comments. In contrast to these open evaluation systems, other product evaluation sites present only expert reviews. For example, at a DVD review site (www.dvdfile.com) experts evaluate DVD movies on the quality of video, audio, supplements, interactive features, and value for money.

As with most of the product review sites, the evaluation processes of learning object repositories provide few opportunities for interaction among expert reviewers (e.g. content experts and instructional designers), and even fewer between expert and consumer reviewers (e.g., learners and teachers). Such interactions are potentially important because reviewers have been consistently observed to modify their evaluation of a learning object after being presented with reviews that differ from their own [3]. Although conclusive evidence is not yet available, we believe that experts and consumers can affect each others opinions and form

convergent evaluations demonstrating greater validity than either could achieve independently.

Interactions among reviewers also present a powerful opportunity for professional development of teachers, instructional designers and media developers. We believe that an evaluation model that educates a significant proportion of the designer population about learning object quality will raise overall quality of the resource pool, and is a much needed complement to models aiming for a high review throughput.

The major learning object repositories have not exploited the meta-evaluation and recommendation features that are now available on popular websites such as Amazon (www.amazon.com). We see a need to extend the current evaluation models and tools to incorporate these features.

3. ELERA

eLera is a website designed to support a distributed community of teachers, instructors, students, researchers, instructional designers, and media developers. Under development since September 2002, the initial version of eLera was publicly released in November 2003 at www.eleranet.net. eLera is a member of eduSource Canada, a network of interoperable Canadian repositories federally funded by CANARIE Inc.

Basic Features

Like MERLOT and DLNET, eLera maintains a searchable database of learning object metadata and reviews, and provides tools and information for learning object evaluation. eLera complies with the IEEE learning object metadata standards as interpreted by the CanCore guide [4]. With permission of the Online Computer Library Centre, it uses a modified version of the Dewey Classification System as a subject taxonomy. eLera includes evaluation forms and reports, statistical aggregation of ratings, and a “my collection” feature allowing members to assemble frequently used objects. We have recently completed localization to French and Chinese. Basic features currently under development include shareable collections, the ability to search other repositories and port metadata and reviews to and from other repositories using the eduSource Communication Language [5], and a personalized home page listing recommended reviews and objects.

Built for Research

While similar in form to the other learning object repositories, eLera has unique goals that will shape its future development. eLera is intended to facilitate research on learning object quality, evaluation, and design. Evaluation data collected through eLera will be used to test the validity and reliability of instruments and evaluation models. eLera moderators can access detailed data pages for each object that present all ratings and comments in tabular form. This feature has already been used in an investigation of Bayesian Networks in learning object evaluation [6].

eLera will be used for research on collaborative evaluation and the interrelation between design and formative evaluation in e-learning development communities. To measure the effects of collaboration, eLera allows us to easily capture the distribution of quality ratings before and after discussion sessions. We expect to create versions of eLera to support successful workflow within teams that develop learning objects. For example we may create an evaluation instrument in which items become activated or deactivated as the object passes through defined stages. This enterprise leads immediately to an examination of critical factors influencing learning object quality in design and development: What work is completed in each stage of the development process? Who should monitor quality at each stage? What information must be communicated to assure quality?

4. LEARNING OBJECT REVIEW INSTRUMENT

The eLera website allows users to evaluate resources with the Learning Object Review Instrument (LORI) [7]. LORI has been iteratively developed through reliability and validity studies with instructional developers and teachers [3]. Version 1.5 of LORI includes nine items:

- *Content Quality*: Veracity, accuracy, balanced presentation of ideas, and appropriate level of detail
- *Learning Goal Alignment*: Alignment among learning goals, activities, assessments, and learner characteristics

- *Feedback and Adaptation*: Adaptive content or feedback driven by differential learner input or learner modeling
- *Motivation*: Ability to motivate and interest an identified population of learners
- *Presentation Design*: Design of visual and auditory information for enhanced learning and efficient mental processing
- *Interaction Usability*: Ease of navigation, predictability of the user interface, and quality of the interface help features
- *Accessibility*: Design of controls and presentation formats to accommodate disabled and mobile learners
- *Reusability*: Ability to use in varying learning contexts and with learners from differing backgrounds
- *Standards Compliance*: Adherence to international standards and specifications

Figure 1 shows how LORI appears to online reviewers. For each of the nine items reviewers can enter comments and ratings on a 5-point scale. Reviewers can choose to skip items that they are unable to assess. Each review is published in eLera as a separate web page. Ratings are averaged over items and reviewers to obtain a mean rating used to sort search results.

The screenshot shows the LORI interface with the following items:

Item	Rating Scale	Comments
1. Content Quality Veracity, accuracy, balanced presentation of ideas, and appropriate level of detail >> more >>	○ ○ ○ ○ ○ ○	<input type="text"/>
2. Learning Goal Alignment Alignment among learning goals, activities, assessments, and learner characteristics >> more >>	○ ○ ○ ○ ○ ○	<input type="text"/>
3. Feedback and Adaption Adaptive content or feedback driven by differential learner input or learner modeling >> more >>	○ ○ ○ ○ ○ ○	<input type="text"/>
4. Motivation Ability to motivate and interest an identified population of learners >> more >>	○ ○ ○ ○ ○ ○	<input type="text"/>
5. Presentation Design Design of visual and auditory information for enhanced learning and efficient mental processing >> more >>	○ ○ ○ ○ ○ ○	<input type="text"/>

Figure 1. LORI as seen by a reviewer (shows only the first five items).

Manage Requests

[Edit]
[Delete]

Let's Review Algorithm Simulations
due 2004/01/28

Hi Folks: You have each been selected for the expertise you would bring to evaluating the learning objects listed here. If you decide to participate please complete your individual reviews by Jan 28. We are tentatively planning an online meeting for 4:00 PST on the 29th.

Objects
Iteration Using Do While Loop
The Animation of Recursion
Tower of Hanoi

Invited Reviewers

Kate Han	accepted
Marek Hatala	pending
Tracey Leacock	pending
Olusola Adesope	pending
Vive Kumar	accepted

Figure 2. An eLera request as viewed by the moderator who created it.

Map of the Human Heart

<http://www.pbs.org/wgbh/nova/heart/heartmap.html>

Overall rating: ★★★★★

Content Quality

score 1	score 2	score 3	score 4	score 5	N/A
0	0	2	0	5	0
John Nesbit		Accurate information with appropriate level of detail.			
Poh Chuah					
jeff may		good detail of heart parts for both science 9 and biology 12 (basic). Should have functions of heart parts.			
Robert Decman		Great for grade 9 biology			
Jimmy Wu					
Jarmila Vik		accurate for grade 9			

Figure 3. The distribution of ratings on a LORI item (Content Quality) as seen by collaborating reviewers.

5. TOOLS FOR COLLABORATIVE EVALUATION

eLera's tools for collaborative evaluation are designed to support the convergent participation model defined and tested in previous research [3, 8, 9]. In this model, small evaluation teams are formed from participants representing relevant knowledge sets and interests (e.g., subject matter expert, learner, instructional designer). A team leader or moderator chooses objects for review, schedules the review activity, and invites team members. Currently, moderators can use eLera's request feature (Figure 2) to invite members to review an object. Members may choose to accept or reject participation.

After the team members have completed individual reviews, they meet in an online, real-time conference to compare and discuss their evaluations. In the convergent participation model, reviewers first discuss the items showing the greatest inter-rater variability. The moderator can use statistics calculated by eLera to order items for discussion. To support comparison of evaluations, eLera presents an aggregated view of ratings and comments for each item of LORI (Figure 3). Team members can edit their ratings and comments during the session. When the collaborative evaluation session is completed, the moderator publishes a team review by automatically aggregating individual reviews authored by team members. The tool requires the agreement of participants before incorporating their individual reviews in the team review.

6. RECOMMENDATION AND TRUST

Through eLera we are researching models for supporting e-learning communities of practice. This research asks how online communities should be structured to foster norms of reciprocity, collective action, identity, and information flow [10]. Key questions at this stage are: How can community members recommend resources and reviews to others? How can they find, and be introduced to, other members with similar or complementary interests? How can they build the identity, interpersonal trust and reputation that are prerequisite to effective collective activity?

At present, eLera provides only rudimentary facilities for recommendation and trust. By default, search results are ordered by average rating so that the most highly rated objects are presented at the top of the list. Users can also choose to order objects by *popularity*, a metric that is incremented whenever an object is placed in a personal collection. To support trust and alliance building, eLera members can create personal profiles detailing their interests and areas of expertise. Thus, decisions about whether to trust and collaborate with a reviewer can be based on the combined knowledge of his or her profile and previous reviews.

As we build on these features we are researching more advanced models of trust and recommendation that will contribute to the nascent research base in this area [11, 12, 13, 14]. For example, we are implementing a "web of trust" for eLera in which members can create a

list of highly trusted others. eLera will be able to recommend new members for one's trust list by chaining forward through the network of trust lists.

7. ELERA WORKSHOPS

Earlier versions of eLera have been used in professional development workshops for local teachers. These have been helpful in refining the usability of the site and identifying deeper community design issues. For example, teachers expected the subject taxonomy in eLera to follow the provincial curriculum structure they work with on a daily basis. They found eLera's modified Dewey system somewhat confusing and not availing of the specific needs of their community [15]. Later versions of eLera will allow localization of subject taxonomies so that members can opt for a taxonomy already established in their community. Translation between local taxonomies will be achieved by mapping to a universal subject taxonomy, similar to the approach advocated by Hatala and Richards [16] for localization of learning object metadata standards.

We will continue to work with teachers and will extend this field-based component of our research to include instructional design and development teams. eLera will evolve as a product of the identified needs of e-learning communities and our ongoing research on learning object evaluation models.

8. REFERENCES

- [1] J.M. Saylor (2002). Why every digital science collection builder should contribute catalog information to the National Science Digital Library. Retrieved December 28, 2003 from <http://collections.comm.nsdlib.org/cgi-bin/wiki.pl?WhyJoin>
- [2] M. Clarke (2003). CLOE peer review. Retrieved January 30, 2003, from Co-operative Learning Object Exchange: <http://cloe.on.ca/documents.html>
- [3] J. Vargo, J.C. Nesbit, K. Belfer, A. Archambault (2003). Learning object evaluation: Computer mediated collaboration and inter-rater reliability. *International Journal of Computers and Applications*, 25 (3), 198-205.
- [4] N. Friesen, S. Fisher, L. Tozer, A. Roberts, S. Hesemeier, S. Habkirk (2003). *CanCore guide for LOM implementation*. CanCore Initiative. Retrieved January 7, 2004 from www.cancore.ca
- [5] M. Hatala, G. Richards, T. Eap, J. Willms (2003). *The EduSource communication language: Implementing open network for learning repositories and services*. Retrieved January 7, 2004 from www.sfu.ca/~mhatala/pubs/sap04-edusource-submit.pdf
- [6] K. Han, V. Kumar, J.C. Nesbit (2003). Rating learning object quality with Bayesian belief networks. *World Conference on E-Learning in Corp., Govt., Health., & Higher Ed.* 2003(1), 1598-1601.

- [7] J.C. Nesbit, K. Belfer, T. Leacock (2003). *Learning object review instrument (LORI)*. E-Learning Research and Assessment Network. Retrieved January 7, 2004 from www.elera.net
- [8] J.C. Nesbit, K. Belfer (in press). Collaborative evaluation of learning objects. In R. McGreal (Ed.), *Online education using learning objects*. London: Routledge/Falmer .
- [9] J.C. Nesbit, K. Belfer, J. Vargo (2002). A convergent participation model for evaluation of learning objects. *Canadian Journal of Learning and Technology*, 28 (3), 105-120.
- [10] R.D. Putnam (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.
- [11] J.C. Nesbit, P.H. Winne (2003). Self-regulated inquiry with networked resources. *Canadian Journal of Learning and Technology*, 29(3), 71-91.
- [12] D.A. Wiley, E.K. Edwards (n.d.). Online self-organizing social systems: The decentralized future of online learning. Retrieved January 4, 2004, from <http://wiley.ed.usu.edu/docs/ososs.pdf>
- [13] M. Recker, A. Walker (2003). Supporting 'word-of-mouth' social networks via collaborative information filtering. *Journal of Interactive Learning Research*, 14(1), pp. 79-98.
- [14] M. Recker, A. Walker, K. Lawless (2003). What do you recommend? Implementation and analyses of collaborative filtering of Web resources for education. *Instructional Science*, 31(4/5), 229-316.
- [15] T. Leacock, J.C. Nesbit, G. Richards (2004). Teachers need simple, effective tools to evaluate learning objects: Enter eLera's LORI. *Computers and Advanced Technology in Education*, Kauai, Hawaii.
- [16] M. Hatala, G. Richards (2002, June). Global vs. community metadata standards: Empowering users for knowledge exchange. *International Semantic Web Conference*, Sardinia, Italy.