# Developing and Using a Guide to Assess Learning Resource Quality in Educational Digital Libraries

Heather Leary[1], Sarah Giersch[2], Andrew Walker[1] and Mimi Recker[1]
*[1]Utah State University,*
*[2]Columbia University*
*United States of America*

## 1. Introduction

The democratization of content creation via ubiquitous Internet tools and infrastructure (Anderson, 2006) has fueled an explosion of user-generated digital content in the commercial and educational markets. Federal agencies in the United States, such as the National Science Foundation, are actively seeking ways to integrate teachers and learners into the *education cyber-infrastructure* (Computing Research Association, 2005; Pea et al., 2008). The overall purpose of this education cyberinfrastructure is to connect people with information and with web-based tools to improve productivity, teaching, and learning. By connecting teachers and learners with tools, they become co-creators of educational content. Educational digital libraries, part of this cyberinfrastructure, have been created as places to deposit and disseminate educational content, which often takes the form of online learning resources of varying levels of granularity. Digital library content can be created easily and deposited rapidly, placing it outside the peer review processes typically employed by publishers and professional societies, but it is unclear whether teachers are using high quality online learning resources in the classroom. To date, educational digital library developers, catalogers, teachers and school administrators have depended on one or a combination of the following proxies to establish an imprimatur of content quality: the reputation and oversight of a funding organization (e.g., the National Science Foundation), the credentials of the content creator (e.g., the National Science Teachers Association), or the collection development policies of specific digital libraries (e.g., the National Science Digital Library or the Digital Library for Earth System Education). However, the definition of "quality"for organizations like those listed above often reflects internal policies and goals, resulting in reviews or judgments that are not comparable across institutions.

Further blurring the boundaries between creator-reviewer, teacher-learner and publisher-consumer, many sites employing user ratings and comment tools, such as YouTube (http://www.youtube.com/), Flickr (http://www.flickr.com/), iTunesU (http://www.apple.com/education/itunes-u/), and ccMixter (http://ccmixter.org/) provide an alternative to the evolving education cyber-infrastructure, creating a rich and diverse environment for disseminating user-generated educational content. Most educational digital libraries do not rely on even a quasi-review for assessing the quality of the content created for,

or by, their users as the above sites do. However, in the omnipresent climate of accountability within the U.S. K-12 education system at the federal, state and local levels, education digital libraries are being challenged to prove their value. For these reasons, it is useful, if not necessary, to develop a method to review the quality of online education resources. And, since social networking tools make sharing a by-product of content creation, it is necessary to develop a standardized set of measures that can be employed across a range of education digital library environments while leveraging existing and emerging social and technical networks to enrich, facilitate, and automate the review process.

This chapter describes the development of a guide to assess the quality of educational learning resources within the context of the Instructional Architect (http://ia.usu.edu), a web-based content-authoring tool for K-12 teachers in the United States. We describe the motivation for developing this Quality Guide, the process for creating it by synthesizing the rubrics of other education digital libraries, and the results of testing and using the Guide with K-12 teachers in the context of professional development workshops. Analyses of its usability and reliability indicate that the Quality Guide influences how teachers design instructional activities using online learning resources. But, defining "quality" remains a difficult task since the perception of quality is dependent on a reviewer's, or user's, purpose for the learning resources and context of their use.

## 2. Background

Over the past 10-15 years, as part of the many global initiatives created to provide access to online educational resources, educational digital libraries in the United States have been developed for K-12 and higher education audiences with support from state and federal agencies: (e.g., the Department of Education; the National Science Foundation), from private foundations (e.g., the William and Flora Hewlett Foundation; the George Lucas Education Foundation), and from universities (e.g. MIT). These institutions funded projects create, or provide access to, curricula that integrates online educational resources and training programs to empower teachers to incorporate learning technologies into their daily preparation and practice and, ultimately, to design their own learning resources. Educational digital library developers have gathered online learning resources of varying levels of granularity (e.g., from images to entire lessons) and of varying sources of authorship (e.g., grant-funded subject matter experts; K-12 teachers; graduate students) into central online portals to enable discovery and re-use by other educators.

While traditional libraries use collection development plans to guide acquisitions that are based on knowledge of their users, educational digital libraries acquire resources created as a result of projects with museums, practicing teachers, researchers and from Internet users. The challenge is to balance collecting and providing access to many online learning resources while maintaining a level of resource quality, cataloging, and curation that distinguishes educational digital libraries from generic Internet search engines and non-educational social software.

However, it is often as difficult for teachers and learners to define quality as it is for the courts to define pornography. As U.S. Supreme Court Justice Potter Steward once noted, "I know it [pornography] when I see it." The problem (or the solution?) lies in the perspective of users who are confronted with a variety of resources of indeterminate origin. For school administrators, teachers, and learners, a high quality resource could consist of an entire

lesson that meets several state standards; a web-based animation that accurately depicts a science concept; or, an essay that provides an answer in a quiz.

Grappling with the idea of "quality", several education digital libraries have already created rubrics to assess their content. As we tried to apply these criteria to the content in the Instructional Architect, as described below, we observed that portions of each rubric were applicable in different settings, and other portions were specific to each digital library with little room for re-use outside of the original context (Martin, 2004). As such, our goals were 1) to synthesize the various dimensions of existing rubrics in order to identify a standardized set of criteria that could potentially be used by any digital library with online educational resources (Giersch, Leary, Palmer & Recker, 2008a), and 2) to create a guide that could be used to assess the quality of materials in the Instructional Architect projects (http://ia.usu.edu). This chapter briefly describes our process for developing the standardized set of assessment criteria; evaluating its utility and usability with middle school science and math teachers; developing a Quality Guide for use with Instructional Architect projects; testing its reliability; and, exploring how the Guide could foster teachers' skills in designing learning resources.

## 3. Previous research

Many education digital library builders have developed rubrics to help assess the quality of online educational resources generated by teachers, faculty and learners. However, the motivations and methods for implementing the rubrics vary as do the implied assumptions about the value of the results for each education digital library.

Established publishing houses and professional societies institute peer review processes to maintain a reputation for producing quality publications thereby increasing subscriptions and membership. For newly-funded education digital libraries, under pressure to prove their value, rubrics and a process to measure quality, were created to first *establish* a reputation with users (Robertshaw, Leary, Walker, Bloxham, Recker, 2009). Similarly, other education digital libraries used rubrics to support access to high quality resources in a digital library or repository (Liu & Ward, 2007; McMartin, 2004) because use of a collection development policy for inclusion of quality resources adds value and buy-in from peer reviewers and users (Sumner et al, 2003). Still other sites developed a rubric to guide authors in creating high quality online resources by gathering feedback from target users (McMartin, 2004).

The methods by which education digital libraries collected data using rubrics and how they used that information also varied. Muramatsu and Agogino (1999) used a rubric to gather feedback from *targeted users* to inform collection development, while Recker, Walker, and Lawless (2003) gathered data through automated mechanisms to provide teachers with suggestions based on resources they had previously viewed and used. Establishing peer-review panels and processes was not a *de facto* choice for every education digital library (Fitzgerald, Lovin, & Branch, 2003; ORC, McMartin, 2004). Enlisting users-as-builders, or reviewers-as-builders, is a characteristic of the grass-roots environment in which many education digital libraries evolved over the last decade in the U.S. However, many digital library developers found that gathering reviews from users was difficult, though providing incentives, in the form of awards (Muramatsu & Agogino, 1999) or other methods (McMartin, 2004), was one way to increase acceptance and adoption of the rubric and review process and of the education digital library in general.

The process of creating and testing an assessment rubric can be time consuming. Many rubric creators believe it is important to map out stakeholders and end-users, review previous work by other digital libraries (Knox et al, 1999), and learn how to gather data (Recker, Walker, Lawless, 2003; Sumner et al., 2003). Testing a rubric with users, for formative evaluation, further refines the rubric, yielding usability information (Fitzgerald, Lovin, & Branch, 2003) as well as real time feedback from target users (Recker, Walker, Lawless, 2003; Sumner et al., 2003), which ultimately assists with the process of implementing the rubric with a larger group of users.

## 4. Context

### 4.1 The Instructional Architect

The Instructional Architect is a simple, web-based authoring tool designed to help K-12 teachers find, design, and use online learning resources in their classrooms (Recker, 2006). When using the Instructional Architect, teachers search for and save links to online learning resources from the Web and educational digital libraries such as the National Science Digital Library (http://nsdl.org). Resources that can be linked include online content, RSS feeds and podcasts. Once resources are gathered users then create 'Instructional Architect projects' (in the form of web pages) that contain instructional objectives, activities, and assessments. The resources are either embedded into the project or link out to web-based resources. In this way, teachers create Instructional Architect projects that customize resources to their local context. Figure 1 shows an Instructional Architect project and an accompanying resource.



Fig. 1. Example of an Instructional Architect project created by Mr. B and related resource.

Initially the majority of users were K-12 teachers in the U.S. who participated in professional development workshops from 2005-2009. Recently, the Instructional Architect system has

garnered more widespread usage and spread 'virally' to teachers not directly attending these workshops. Table 1 shows statistics on Instructional Architect usage and growth in the number of Instructional Architect projects and resources over the last 12 months.

| Data | N | 12-month growth |
|---|---|---|
| Registered users | 5,300 | 23% |
| Instructional Architect projects created | 11,500 | 37% |
| Online learning resources used | 52,100 | 67% |
| Visits to Instructional Architect project (since 8/2006) | 1 million | 64% |

Table 1. Instructional Architect usage data (to August 2010).

Before any of the abundant Instructional Architect projects created by users can be ingested into educational digital libraries, like the National Science Digital Library, to become part of the education cyberinfrastructure, they need to be evaluated for quality. To support this outcome, the Instructional Architect projects must be reviewed with a rubric that is reliable across multiple sites, a process that is scalable across the thousands of IA projects, and that has useful results for end-users, reviewers, and digital library developers.

## 5. Method

### 5.1 Synthesizing existing rubrics & creating the Quality Guide

Following the recommendations above for establishing a rubric to assess content (Knox et al, 1999; Recker, Walker, Lawless, 2003; Sumner et al., 2003), we began with a literature review in the disciplines of computer science, library and information science, education (specifically online learning objects), and digital libraries. We selected only articles that included or referenced rubrics used to review online educational resources. The literature review yielded articles with descriptions of 12 rubrics (Liu, & Ward, 2007; Fitzgerald, Lovin, & Branch, 2003; Nesbit, Belfer, Leacock, 2003; Ohio Resource Center; Knox et al., 1999; Recker, Walker, & Lawless, 2003; Sumner, Khoo, Recker, & Marlino, 2003; Custard & Sumner, 2005; Kastens, DeFelice, Devaul, DiLeonardo, Ginger, Larsen, et al., 2005; McMartin, 2004; Muramatsu & Agogino, 1999).

In keeping with our first goal (to synthesize the various dimensions of existing digital library rubrics in order to identify a standardized set of criteria that could potentially be used by any digital library with online educational resources), our initial plan was to identify review criteria at the most granular level and then do a frequency analysis to identify broad topics that applied across criteria (e.g., pedagogy or usability). These would then become the foundational criteria. Accordingly, we un-bundled the rubrics and identified over 200 review criteria, some of which included only one or two words and sometimes no definition. However, when conducting our frequency analysis, we found it difficult to group similar review criteria because:

- None of the rubrics used a standard vocabulary for their review criteria;
- The definitions for criteria contained multiple concepts that defied easy categorization; and,

- Some criteria (with and without definitions) were so contextualized to their site that they became virtually meaningless when viewed out of context from the whole rubric.

These factors also made it difficult to create topic areas before and after the frequency analysis. Given the challenges of our initial approach, we revised our methods and used a card sort technique. Card sorts are typically employed as a user-centered design approach for assessing web site structure (Maurer & Warfel, 2004). This process has participants sort a series of cards, each labeled with a piece of content or functionality into groups that make sense to them (Lamantia, 2003). This technique allows input from several experts, with the consensus between them resolving ambiguities in and across the review criteria.

In preparing the initial list of over 200 review criteria for sorting, we realized that some of the most granular criteria, especially those without definitions, would prove difficult to sort, appearing as so much noise to study participants. Therefore, we culled the list to include only criteria from the literature that included a definition and that would make sense when read outside the context of its original use. Our intention was not to represent criteria as they had been used within a particular review process, but to use the criteria as presented in the literature. The result was that 104 review criteria remained for use in the card sort exercise.

The card sort methodology recommends using seven to ten participants. Our sorting exercise included 10 participants who were graduate students, professors from instructional technology and education, and academic librarians. We presented the participants with a spreadsheet containing the review criteria, criteria definitions, and the source rubric or citation. We chose the "open sort" method and asked participants to sort the review criteria into groups of their own choosing within a spreadsheet and to provide their own titles for the groups. Participants were allowed to place criteria in more than one group as long as they indicated which criteria-group pairing represented their highest priority.

Each of the ten participants created an average of nearly 13 groups into which they individually sorted the 104 review criteria. One participant sorted criteria into multiple groups but did not indicate a top priority, so rather than guessing at the intention, the results were not included. Ultimately, we analyzed the nine result sets containing an average of 11 groups per participant.

A detailed analysis (Giersch, Leary, Palmer, Recker, 2008a; Giersch, Leary, Palmer, Recker, 2008b) yielded six elements from the 11 groups. These became the foundational set of criteria, which were standardized to the point where they could be used by any digital library with online education resources, thereby accomplishing our first goal. Table 2 shows the six elements, or standardized criteria, developed by study participants in the card sort exercise (e.g., Content). Corresponding definitions and the references for them are also displayed. The definitions (e.g., Resource content is accurate) are statements from the original 12 rubrics that participants associated with the standardized criteria.

Using these foundational criteria, we then began to address the second goal of creating a guide that could be used to assess the quality of materials in the Instructional Architect projects. Figure 2 shows the prototype of the foundational criteria and definitions in beta form before it was designed to be more user-friendly and visually appealing. This subsequently became the first version of the Instructional Architect Quality Guide, and included five Likert scale star ratings for each criterion, as these are widely used by rating systems, and less-detailed definitions of criteria.

| Standardized criteria | Definitions and references |
|---|---|
| Interface Design & Accessibility | The resource is attractive (Liu & Ward, 2007)<br>The resource is easy to navigate (Ohio Resource Center)<br>The resource contains no advertising (Custard & Sumner, 2005; Sumner, Khoo, Recker, Marlino, 2003)<br>The resource contains links that work (Liu & Ward, 2007)<br>The resource is designed to accommodate disabled and mobile learners (Nesbit, Belfer, Leacock, 2003) |
| Technical Reliability (Muramatsu & Agogino, 1999) | The resource uses multimedia (Flash, sound clips, videos, or applets) that work (Custard & Sumner, 2005)<br>The resource clearly identifies the supporting technical resources required (Knox et al., 1999)<br>The resource contains help features (Nesbit, Belfer, Leacock, 2003) |
| Content | Resource content is accurate (Muramatsu & Agogino, 1999; Kastens et al., 2005; Nesbit, Belfer, Leacock, 2003; Recker, Walker, Lawless, 2003; Ohio Resource Center)<br>Resource content is complete (Fitzgerald, Loving, & Branch, 2003)<br>Resource content is impartial (Fitzgerald, Loving, & Branch, 2003)<br>Resource content is clearly written (Ohio Resource Center)<br>Resources content is maintained (Custard & Sumner, 2005; Liu & Ward, 2007) |
| Pedagogy | The resource is engaging (Liu & Ward, 2007)<br>The reource is motivating (Fitzgerald, Lovoing, & Branch, 2003)<br>The resource is interactive (Muramatsu & Agogino, 1999)<br>The resources includes assessment(s) (Ohio Resource Center)<br>The resource provides feedback (Nesbit, Belfer, & Leacock, 2003)<br>The resource supports learners proceeding at their own pace (Muramatsu & Agogino, 1999)<br>The resource supports introductory, reinforcing, or summative activities |
| Administrative | The resource contains direct and explicit links to state or national educational teaching standards (Custard & Sumner, 2005; Ohio Resources Center)<br>The resource contains information about ist author or creator, email (Recker, Walker, Lawless, 2003), site domain (Custard & Sumner, 2005), role (Custard & Sumner, 2005)<br>The resource is described by current metadata (Custard & Sumner, 2005) |
| Other | The resource requires a fee for access (Custard & Sumner, 2005)<br>The resources overall rating/confidence level (Recker, Walker, Lawless, 2003) |

Table 2. Standardized set of criteria with definitions.

Quality Guide
While viewing an IA project, rate the following six items as found in the project.

| ☆☆☆☆☆ | **Interface design**<br>Definition: dynamic, visually pleasing aesthetics, standards, accessibility, navigation, good user interface |
|---|---|
| ☆☆☆☆☆ | **Accessibility**<br>Definition: technical reliability, reusability, usability design, reliability, adaptability, ADA |
| ☆☆☆☆☆ | **Content**<br>Definition: accuracy, organization, quality, completeness, currency, value, interdisciplinary |
| ☆☆☆☆☆ | **Pedagogy**<br>Definition: instruction methods, design methods, learning goals, anticipated outcomes, motivation, engagement |
| ☆☆☆☆☆ | **Administrative**<br>Definition: authority, authorship, metadata, advertising, credibility |
| ☆☆☆☆☆ | **Overall rating** |

Fig. 2. Prototype of foundational criteria and definitions

**IA project ID:_____**          **Reviewer Name:_____**

Circle the star you feel best describes the IA project, and provide any comments about your rating.

| Number | Criteria | Rating |
|--------|----------|--------|
| 1 | Content accuracy | ☆ Very Inaccurate   ☆ Somewhat Inaccurate   ☆ Not enough Information   ☆ Somewhat Accurate   ☆ Very Accurate<br><br>Comments: |
| 2 | Text clarity | ☆ Very Unclear   ☆ Somewhat Unclear   ☆ Not enough Information   ☆ Somewhat Clear   ☆ Very Clear<br><br>Comments: |
| 3. | Links in project | ☆ No links work   ☆ Some links don't work   ☆ Not enough Information   ☆ Some links work   ☆ All links work<br><br>Comments: |
| 4 | Project completeness: includes the state standard learning goal, assessment, etc. | ☆ Very Incomplete   ☆ Somewhat incomplete   ☆ Not enough Information   ☆ Somewhat complete   ☆ Very complete<br><br>Comments: |
| 5 | The project: (circle one) | a.  Provides a resorce list<br>b.  Teaches a concept<br>c.  Reinforces a concept<br>d.  Provides a summary of content  with a learning activity<br>e.  Don't Know<br>f.  Does something else (describe below) |
| 6 | Overall rating of the project | ☆ Meets no criteria   ☆ Meets very little criteria   ☆ Meets some criteria   ☆ Meets criteria   ☆ Exceed criteria<br><br>Comments: |

Fig. 3. First version of the Instructional Architect Quality Guide.

### 5.2 Testing the Instructional Architect Quality Guide

Twenty-eight participants were recruited to test the first version of the Instructional Architect Quality Guide. They were part of a cohort of U.S. K-12 teachers in an online graduate program, and they completed required activities as part of an online course. Complete data were received from 17 participants. Testing the Quality Guide was directed by the following research questions:

1.  What is the student's view of using technology in a classroom?
2.  How do teachers assess an online educational resource before and after using the Quality Guide?
3.  Using the Quality Guide, how helpful are conducting and receiving peer reviews for the student?
4.  What changes should be made to the Instructional Architect Quality Guide based on student use and feedback?

### 5.3 Data sources

The participants took part in an online learning module in the context of learning how to use the Instructional Architect and the Quality Guide. They completed pre-and post-online surveys, which used a combination of open-ended and Likert-scale questions. Surveys measured the extent of participants' experience and classroom practice in using online educational resources; their attitudes about online educational resources and technology in general and their use in a classroom; and, strategies for evaluating an online educational resource. Within the learning module, participants were placed in small groups of 4-5 participants, and they evaluated one another's Instructional Architect projects using the Quality Guide. Conversations about these evaluations were posted in a BlackBoard discussion forum and collected for analysis by the researchers. Lastly, participants wrote a reflection paper describing their experience using the Instructional Architect, how they used online educational resources in an instructional situation, difficulties and successes in designing and implementing their project, what they learned by reviewing their peers' Instructional Architect projects and by using the Quality Guide, and how they could improve their Instructional Architect project.

## 6. Results

Research question one used data from the pre-and post-survey Likert items. Effect sizes were calculated on questions relating to participants' attitude, experience with and use of online educational resources and technology in their classrooms in order to understand changes in participants' pedagogical choices and activities in their classrooms. An effect size ($d_w$) shows a magnitude of change and is considered small at 0.2 detectable by an expert looking closely at the phenomenon, medium at 0.5, and large at 0.8 detectable by an untrained observer (Cohen, 1988). Larger effect sizes equate to larger impacts due to the intervention. Small, but positive changes, statistically significant at $p < 0.05$, were shown from the pre-test to the post-test as teachers reported that they knew how to effectively teach with technology in the classroom ($d_w = 0.28$, $p < 0.05$), and they knew how to effectively use technology in their classroom ($d_w = 0.42$, $p < 0.05$). From this we know that after participating in the learning module, teachers began to overcome initial barriers of using technology in the classroom and could now focus on making better pedagogical decisions about the type of learning resources they could create.

We then used a qualitative analysis to address research question two with the understanding that participants had learned how to use online educational resources. Using a constant comparative analysis (Glaser & Strauss, 1967), data from the reflection papers, discussion boards, and one question in the pre-and post-survey were mapped into themes and analyzed. Participants reported in the pre-survey that when evaluating online educational resources they looked for fit with the curriculum, accuracy, ease of use, currency, text readability, and recommendations by others. After completing this course module, participants added in their post-survey comments that they looked for: content quality, distractions on the resource pages, credibility of the site, and engagement. Many of the criteria they added were items listed in the Quality Guide. Thus, it appeared that use of the Quality Guide helped refine participants' approach to designing online learning resources, in the form of Instructional Architect projects.

To answer research question three, the comments from the discussion boards in Blackboard where participants posted their evaluations of their peers' Instructional Architect projects were used. Some participants commented on the readability and text clarity of the Instructional Architect project, and many caught spelling errors. One participant commented after receiving some feedback, "I have learned that having a peer read through my projects can be very valuable. They catch mistakes that I do not see." Overall, 53% of the participants reported that providing and receiving feedback from their peers via reviews was valuable.

Two additional themes surfaced from analyses of the discussion boards during the process of peer evaluation that related to re-use of other teachers' Instructional Architect projects and ideas. First, participants asked if they could use an Instructional Architect project they reviewed, or said they planned to use it in their classroom. The second theme was that participants learned what others had done and wanted to implement the same idea in their project. A participant commented that, "Completing a peer review gives us an opportunity to see other creations and improve our programs as we see fit."

The final data collection point included participants' reflection papers. Participants were asked to report on their instructional situation, successes and difficulties, what they learned from their peers and using the Quality Guide, and how they could improve their Instructional Architect project. Three areas of learning were repeatedly reported in the reflection papers: what they learned by reviewing peer work; what they learned for their own work; and, the value of re-using Instructional Architect projects. Similar ideas were also expressed throughout the discussion boards and survey answers.

### 6.1 Reliability and improvements

Version one (see Figure 3) of the Quality Guide was tested for reliability in order to answer research question four. Peer ratings of Instructional Architect projects using the Guide, provided by the students in the online graduate course, were used for this analysis. The Quality Guide criteria (5 stars) were scored on a scale of 0-4, where zero or one star is low (e.g., very unclear) and four or five stars is high (e.g., very clear). See Figure 3 above. Means and standard deviations were run for each question.

Criteria one through four and six of the Quality Guide were analyzed using an intra-class correlation (scale of -1 to 1, high agreement is found as it approaches 1, with less agreement as it approaches -1). Criteria five was not analyzed as it asks for a classification on what the Instructional Architect project is doing (e.g., teaching a new concept, reinforcing a concept, etc). Table 4 shows the intraclass correlation values for each question analyzed.

A close look at the data (see Table 3) shows that the teachers generally rated their peers at an average of 2 or more (3 or more stars on the Quality Guide) for any of the criteria on the Guide. This could possibly be attributed to the fact that educators as a whole tend to provide high ratings for their peer, a phenomenon observed in prior work (Walker, Recker, Lawless, & Wiley, 2004). As a result of the high ratings, the full range of the scale was not used, and slight departures were magnified in the intraclass correlation. The negative values indicate that there was more variation between projects than there was between raters. As is shown in the low intraclass correlation (Table 4), agreement on those ratings was never high, and in most cases was non-existent.

| Criteria | Mean | Standard Deviation |
|---|---|---|
| 1 – Content accuracy | 2.50 | 1.13 |
| 2 – Text clarity | 2.34 | 0.98 |
| 3 – Links in project | 3.30 | 0.83 |
| 4 – Project completeness | 2.14 | 1.50 |
| 6 – Overall rating | 3.22 | 1.84 |

Table 3. Means and standard deviations for criteria 1-4, 6 of the Instructional Architect Quality Guide.

| Criteria | Intraclass Correlation |
|---|---|
| 1 – Content accuracy | 0 |
| 2 – Text clarity | -0.129 |
| 3 – Links in project | 0.031 |
| 4 – Project completeness | -0.89 |
| 6 – Overall rating | -0.129 |

Table 4. Intraclass correlation (ICC) values for criteria 1-4, 6 of the Instructional Architect Quality Guide.

The reliability analysis suggests that the scale needs more explanation for each item and that the scale should encompass either a broader range or a dichotomous rating (yes or no). Improvements made to the Quality Guide as a result of this testing included adding more detailed definitions of the review criteria, re-arranging the order of criteria, and modifying scoring instructions. As a result of this testing, improvements were made to create the final version of the Quality Guide (see Figure 4), which is currently used to inform and support individual teachers as they create Instructional Architect projects.

**INSTRUCTIONAL**
**ARCHITECT**™

Use the following indicators to inform your choices when designing or reviewing IA projects.

## Accuracy
Information is credible, truthful, reliable, current

| Very inaccurate | Inaccurate | Accurate | Very accurate |
|---|---|---|---|
| Comments: | | | |

## Text Clarity
Appropriate grade level, length, and amount (as concise as needed)

| Very unclear | Unclear | Clear | Very clear |
|---|---|---|---|
| Comments: | | | |

## Links in Project
Work and go to the correct page

| All links fail | Most links fail | Most links work | All links work |
|---|---|---|---|
| Comments: | | | |

## Project Completeness
Includes the state standard, objectives, learning goal, assessment, or an example

| Very Incomplete | Incomplete | Complete | Very complete |
|---|---|---|---|
| Comments: | | | |

## The Project

| Provides a resource list | Teaches a concept | Reinforces a concept | Provides a summary of content | Other |
|---|---|---|---|---|
| Comments: | | | | |

## Overall

| Meets no criteria | Meets some criteria | Meets all criteria | Exceeds criteria |
|---|---|---|---|
| Comments: | | | |

UtahState UNIVERSITY    DL DIGITAL LIBRARY CONNECT    NSF    NSDL

Fig. 4. Final version of the Instructional Architect Quality guide.

## 7. Conclusion

The foundational list of criteria (see Table 2), which was derived from rubrics of 12 different education digital libraries, can be used by other digital library developers to evaluate the quality of online learning resources, or it can be modified to fit the need of local contexts, such as the Instructional Architect. We created the Instructional Architect Quality Guide to encourage change in teachers' perspective and behaviors around designing and using online educational resources. The results of our evaluation indicate that participants found value in the Quality Guide as a means to improve their own projects through completing and receiving reviews. Unfortunately, we found that the Instructional Architect Quality Guide does not scale well when it is used to determine which of the 11,500 Instructional Architect projects are of sufficient quality to be ingested by the National Science Digital Library.

## 8. Future research

Future work is focused on addressing these scalability issues through using an already-developed and -tested automated and scalable quality assessment method, which will better support and facilitate teacher co-creation of online content. Similar to the work described in this chapter, this approach relies upon distilling the elusive notion of quality into a set of concrete indicators. We are also testing a previously developed machine learning algorithm, called Opera (Bethard et al., 2009), to assess whether its quality ratings along these indicators match teacher assessments (Recker et al., 2011). Results from these ongoing studies will help determine if Opera can serve as a proxy for the laborious and expensive task of having teachers or peer reviewers assess the quality of online resources and IA projects. We are still motivated by having online learning resources and Instructional Architect projects reviewed so that quality content can be ingested by educational digital libraries, such as the National Science Digital Library and ultimately included in the education cyberinfrastructure.

## 9. Acknowledgements

## 10. References

Anderson, C. (2006). *The Long Tail: Why the future of business is selling less of more.* Hyperion, New York.

Bethard, S., Wetzler, P., Butcher, K., Martin, J. H., Sumner, T. (2009). Automatically Characterizing Resource Quality for Educational Digital Libraries. In *Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries.* New York: ACM.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Vol. 2). Hillsdale, NJ: Lawrence Earlbaum Associates.

Computing Research Association. (2005). *Cyber-infrastructure for Education and Learning for the Future: A vision and research agenda*. Washington, D.C.

Custard, M., & Sumner, T. (2005). Using Machine Learning to Support Quality Judgments. D-Lib Magazine, 11, 11.

Fitzgerald, M.A., Lovin, V., & Branch, R.M. (2003). A Gateway to Educational Materials: An Evaluation of an Online Resource for Teachers and an Exploration of User Behaviors. *Journal of Technology and Teacher Education*, 11, 1, 21-51.

Glaser, B. G. and A. L. Strauss (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago, Aldine Publishing Company.

Giersch, S., Leary, H., Palmer, B., Recker, M. (2008)a. Supporting Meaningful Learning with Online Resources: Developing a Review Process. In *Proceedings of the Annual Open Education Conference*, Logan, Utah, September 24-26.

Giersch, S. Leary, H., Palmer, B., & M., Recker (2008)b. Developing a Review Process for Online Resources. In *Proceedings of the Joint Conference on Digital Libraries,* Pittsburgh, PA, June 15-20.

Hanson, K. & Carlson, B. (2005). *Effective Access Report*. Education Development Center Instructional Architect Quality Guide, latest version:
http://digitalcommons.usu.edu/itls_research/5/

Kastens, K., DeFelice, B., Devaul, H., DiLeonardo, C., Ginger, K., Larsen, S., et al. (2005). Questions & Challenges Arising in Building the Collection of a Digital Library for Education: Lessons from Five Years of DLESE. *D-Lib Magazine*, 11, 11.

Knox, D., et al (1999). The Peer Review Process of Teaching Materials: Report of the ITiCSE'99 Working Group on Validation of the quality of teaching materials. *Annual Joint Conference Integrating Technology into Computer Science Education*, Cracow, Poland.

Lamantia, J. (2003). *Analyzing Card Sort Results with a Spreadsheet Template*. Retrieved January 18, 2008, from Boxes and Arrows Web site:
http://www.boxesandarrows.com/view/analyzing_card_sort_
results_with_a_spreadsheet_template

Leary, H., Giersch, S., Walker, A., Recker, M. (2009). Developing a Review Rubric for Learning Resources in Digital Libraries. *ITLS Faculty Publications*. Paper 17.
http://digitalcommons.usu.edu/itls_facpub/17

Liu, K. & Ward, V. (2007). Access Excellence @ the National Health Museum. *D-Lib Magazine*, 13, 11/12.

Maurer, D. & Warfel, T. (2004). *Card Sorting: a definitive guide*. Retrieved January 18, 2008, from Boxes and Arrows Web site:
http://www.boxesandarrows.com/view/card_sorting_a_defini tive_guide

McMartin, F. (2004). MERLOT: A Model for User Involvement in Digital Library Design and Implementation. *Journal of Digital Information*, 5, 3.

Muramatsu, B. & Agogino, A. (1999). The National Engineering Education Delivery System: A Digital Library for Engineering Education. *D-Lib Magazine*, 4, 5.

Nesbit, J., Belfer, K., Leacock, T. (2003). *Learning Object Review Instrument (LORI 1.5)*, User Manual. Retrieved January 15, 2008.

Ohio Resources Center (ORC): http://ohiorc.org

Pea, R., with Christine L. Borgman (Chair), Hal Abelson, Lee Dirks, Roberta Johnson, Kenneth R. Koedinger, Marcia C. Linn, Clifford A. Lynch, Diana G. Oblinger, Katie

Salen, Marshall S. Smith, Alex Szalay (2008). *Fostering learning in the networked world – the cyberlearning opportunity and challenge: A 21st century agenda for the National Science Foundation (Report of the NSF Task Force on Cyberlearning)*. Arlington, VA: National Science Foundation.

Recker, M., Leary, H., Walker, A., Diekema, A. R., Wetzler, P., Sumner, T., Martin, J. (2011, April). Modeling Teacher Ratings of Online Resources: A Human-Machine Approach to Quality. Paper *presentation at the American Educational Research Association*, New Orleans.

Recker, M., Walker, A., Giersch, S., Mao, X., Halioris, S., Palmer, B., Johnson, D., Leary, H., Robertshaw, M.B. (2007). A study of teachers' use of online learning resources to design classroom activities. *New Review of Hypermedia and Multimedia*, 13, 2, 117-134.

Recker, M. (2006). Perspectives on teachers as digital library users: Consumers, contributors, and designers. *D-Lib Magazine*, *12*, 9.

Recker, M., Walker, A,. & Lawless, K. (2003). What do you recommend? Implementation and analyses of collaborative information filtering of web resources for education. *Instructional Science*, 31, 299–316.

Robertshaw, M. B., Leary, H., Walker, A., Bloxham, K., Recker, M. (2009). Reciprocal mentoring "in the wild": A retrospective, comparative case study of ICT teacher professional development. In E. Stacey (Ed.) *Effective Blended Learning Practices: Evidence-Based Perspectives in ICT-Facilitated Education*, Melbourne: IGI Global Press.

Sumner, T., Khoo, M., Recker, M., & Marlino, M. (2003). Understanding educator perceptions of "quality" in digital libraries. In *Proceedings of the Joint Conference on Digital Libraries*, Houston, Texas, May 27-31.

Walker, A., Recker, M., Lawless, K., & Wiley, D. (2004). Collaborative information filtering: A review and an educational application. *International Journal of Artifical Intelligence and Education*, 14, 1-26.
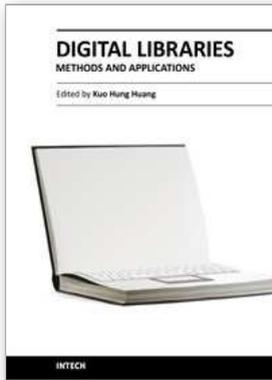
**Digital Libraries - Methods and Applications**

Edited by Dr. Kuo Hung Huang

Digital library is commonly seen as a type of information retrieval system which stores and accesses digital content remotely via computer networks. However, the vision of digital libraries is not limited to technology or management, but user experience. This book is an attempt to share the practical experiences of solutions to the operation of digital libraries. To indicate interdisciplinary routes towards successful applications, the chapters in this book explore the implication of digital libraries from the perspectives of design, operation, and promotion. Without common agreement on a broadly accepted model of digital libraries, authors from diverse fields seek to develop theories and empirical investigations that to advance our understanding of digital libraries.

**How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Heather Leary, Sarah Giersch, Andrew Walker and Mimi Recker (2011). Developing and Using a Guide to Assess Learning Resource Quality in Educational Digital Libraries, Digital Libraries - Methods and Applications, Dr. Kuo Hung Huang (Ed.), ISBN: 978-953-307-203-6, InTech, Available from: http://www.intechopen.com/books/digital-libraries-methods-and-applications/developing-and-using-a-guide-to-assess-learning-resource-quality-in-educational-digital-libraries

# INTECH
open science | open minds