

Short Papers

Ranking Learning Objects through Integration of Different Quality Indicators

Javier Sanz-Rodríguez, Juan Manuel Dodero, and Salvador Sánchez-Alonso

Abstract—The solutions used to-date for recommending learning objects have proved unsatisfactory. In an attempt to improve the situation, this document highlights the insufficiencies of the existing approaches, and identifies quality indicators that might be used to provide information on which materials to recommend to users. Next, a synthesized quality indicator that can facilitate the ranking of learning objects, according to their overall quality, is proposed. In this way, explicit evaluations carried out by users or experts will be used, along with the usage data; thus, completing the information on which the recommendation is based. Taking a set of learning objects from the Merlot repository, we analyzed the relationships that exist between the different quality indicators to form an overall quality indicator that can be calculated automatically, guaranteeing that all resources will be rated.

Index Terms—Learning object, Merlot, ranking, quality.

1 INTRODUCTION

In the field of learning objects, financial considerations have driven technological development. The widely touted concept of the learning object was driven, at least in part, by the hope that sharable and reusable learning resources would reduce the costs needed to produce them [7]. It seems clear that the most attractive aspect of learning objects is their reusability, meaning the effective use of a learning object by different users in different technological environments and in different educational contexts [29].

However, in spite of these expectations, reuse is not currently as common as might be expected [28]. One of the factors conditioning the reuse of learning objects is how easily they can be found by users. As is common with most search engines, many searches in repositories will return a large number of hits, leaving users with the problem of deciding which resources might best suit their needs. Without some formalized process that enabled the searching algorithm to calculate relative importance, any searching process being used across such a large number of resources would always appear to have some inherent weaknesses [5], making it difficult to choose the most suitable resource and reducing reuse. In an attempt to minimize this problem, most repositories have used expert and user evaluation of educational materials. Specifically, Tzikopoulos et al. [35] found that 23 out of the 59 repositories they studied offered several mechanisms for evaluating the educational materials. However, the evaluation system used to-date is inadequate for several reasons [17].

The task of manually reviewing materials is laborious, and the quantity of educational resources is enormous and growing by the

day. For example, in October 2009 at the time of carrying out this study, the Multimedia Educational Resource for Learning and Online Teaching (Merlot) repository contained 21,399 items, of which only 2,867 or 13 percent had been peer reviewed. In this way, the unrated materials will appear at the end of the search results, as if they were poor quality items. This situation arises because existing evaluation initiatives use a time-consuming inspection of the materials as the main source of information. As Ochoa and Duval [26] point out, for a measure of the quality of learning objects to be useful, we must be able to calculate it automatically.

Furthermore, if we analyze the reliability of these explicit evaluations, we also encounter problems. Most of the evaluations carried out by experts are made individually, although the collaborative assessment process is more reliable than the individual evaluation [4]. To implement this improvement, we could develop collaborative evaluation processes for the repositories; but this would further increase the already high cost of evaluating resources.

As for user reviews, they are also subject to major limitations due to a range of problems, such as lack of user training or possible subjectivity of their preferences [14]. Besides this, only a small number of users carry out such evaluations, so their ratings might not be representative of user opinions in general [16].

In a similar vein, Akpınar [1] performed a validation study for certain areas of evaluation of the tool Learning Object Review Instrument (LORI) [25], contrasting the evaluations with surveys of students and teachers. They concluded that LORI evaluations are not sufficient to predict the educational benefits that would be obtained with the learning objects.

In addition, although there are several initiatives that allow a search to be carried out in different repositories, such as that performed in the EduSource project [23], we have a situation where the evaluation systems differ for each repository, making it difficult to sort results that involve several repositories. Just as the existence of different metadata application profiles complicates searches for materials across several repositories, and in spite of existing initiatives to solve these problems, such as the sharing of tags across repositories [37] or the representation of user feedback in a structured and reusable format so that it can be reused by different recommender systems [21], we need to develop strategies that allow us to integrate the repositories' different evaluation systems [20]. Similarly, Vuorikari et al. [38] also identify the need for a reusable and interoperable metadata model for sharing and reusing evaluations of learning resources.

Furthermore, Kelty et al. [17] claim that the educational resources are being evaluated statically, as with traditional educational materials. To reduce the effect of this shortcoming, he proposes that evaluations should not only be focused on content, but should also take into account the possible contexts of usage.

In any case, the availability of large databases with evaluations has opened up new possibilities for the development of indicators to complement existing evaluation techniques, which are based on a considerable manual inspection effort, with others that can be calculated automatically and facilitate an indicator of the quality of educational materials in a less costly manner [12].

As a possible improvement, Kelty et al. [17] propose systems similar to the "lenses" mechanism used in the Connexions repository, where each lens is created using an evaluation criterion: peer reviews, popularity, frequency of reuse, number of times it is bookmarked, etc., and the application of one or a combination of lenses allows us to filter educational materials.

Similarly, Han [14] points out that the current learning object recommendation systems lack a weighting mechanism, where

- J. Sanz-Rodríguez is with the University Carlos III of Madrid, Av. Universidad 30, 28911 Leganés, Madrid, Spain. E-mail: javier.sanz.rodriguez@uc3m.es.
- J.M. Dodero is with the University of Cádiz, C/Chile, s/n, 1003 Cádiz, Spain. E-mail: juanma.dodero@uca.es.
- S. Sánchez-Alonso is with the University of Alcalá de Henares, Ctra. Barcelona km. 33600, 28871 Alcalá de Henares, Madrid, Spain. E-mail: salvador.sanchez@uah.es.

Manuscript received 5 Feb. 2010; revised 21 Apr. 2010; accepted 2 Aug. 2010; published online 17 Aug. 2010.

For information on obtaining reprints of this article, please send e-mail to: lt@computer.org, and reference IEEECS Log Number TLT-2010-02-0011. Digital Object Identifier no. 10.1109/TLT.2010.23.

evaluations submitted by different sources can be taken into account differently. Han proposes an integrated quality rating, which brings together explicit evaluations (by experts and users), anonymous evaluations, and implicit indicators (bookmarks, number of hits).

Following in the line of these two last proposals, the aim of this study is to design an overall quality indicator that incorporates all the available quality indicators.

If we assume that all existing quality indicators constitute different views of quality that might complement one another, we can analyze how they are interrelated to form an overall quality indicator that can be calculated automatically, guaranteeing that all resources will be rated.

In order to carry out this research, we followed the phases proposed by Glass [13] reflected in the structure of the rest of the document. In the informational phase (Section 2), the quality indicators will be identified and grouped into different categories. In the propositional phase (Section 3), a measure of overall quality that incorporates all of the quality indicators will be proposed. In the analytical phase (Section 4), we will analyze the relationships between quality indicators, taking a significant sample of materials from the Merlot repository in order to study how the different quality indicators represent different views of quality, and to feed into the algorithm to calculate the overall quality indicator. In the evaluative phase (Section 5), results are revealed, and, finally, conclusions will be drawn in Section 6.

2 LEARNING OBJECT QUALITY INDICATORS

The learning object quality indicators can be classified into three categories:

- **Explicit.** Includes all explicit evaluations carried out by experts and users.
- **Implicit.** Taken from the implicit usage data for the materials, such as number of visits, number of times it is bookmarked by users, number of times it is downloaded, etc.
- **Characteristical.** Descriptive information on the characteristics of the materials obtained from the metadata.

2.1 Explicit Quality Indicators

The main reason Nesbit and Belfer [24] offer to justify evaluation is the need to help users search for and select learning objects.

Although, there are many studies on how to evaluate learning objects, such as those proposed by Kay and Knaack [15] and Kurilovas and Dagiene [19], the evaluations put into practice are those implemented in the different repositories.

In the Merlot repository, the materials are graded using a peer review process. Peer reviews evaluate three dimensions: quality of the content, usability, and effectiveness as a learning tool. Each aspect is rated on a scale from 1 to 5, rating objects from "poor" to "excellent." The weighted mean of the three dimensions will be the final value of the learning object evaluation [36]. Registered users may also rate and comment on resources.

The e-Learning Research and Assessment network (eLera) repository allows users to evaluate the materials using the LORI tool, evaluating nine aspects: content quality, learning goal alignment, feedback and adaptation, motivation, presentation design, interaction usability, accessibility, reusability, and standards compliance. As with Merlot, each feature is rated on a scale from 1 to 5. It is worth pointing out that collaborative evaluation initiatives have been developed using eLera, in which groups of experts took part [36].

Finally, the Connexions repository proposes quality evaluation, using a lenses mechanism, such that by using one or several combined lenses a user can select the best materials. Among the

possible types of lens are those based on peer reviews and those developed by users [2].

2.2 Implicit Quality Indicators

Using implicit data derived from usage, in order to recommend resources, is an idea already used for selecting web pages. Claypool et al. [6] claim the benefits of using implicit data taken from user behavior to order search results. These measures have been used to improve Internet searches, since they reflect the interests and the level of user satisfaction and are less costly than explicit evaluations [11].

In the case of learning objects, the Merlot repository holds implicit information on access to the resources or bookmarking by users. In Connexions, the lenses for recommending materials can be automatically generated based on data, such as popularity, the number of times it is reused, the number of times it is bookmarked, etc. [2]. Reinforcing this idea, Kumar et al. [18] propose that to complete the information on the quality of educational materials, usage data for the materials can be used in addition to the evaluations available in the repositories.

Similarly, Yen et al. [39] propose the use of information on references to educational materials to order them, drawing their idea from the Page Rank algorithm used by Google to return search results. Also, using the Page Rank algorithm, Duval [9] proposes LearnRank, a context-dependent ranking algorithm.

Finally, in his manifesto on learning objects, Duval [8] proposes the dynamic inclusion of all existing usage information in the metadata: when it appears in a search list, when its description is accessed, when it is downloaded, when it is assigned as part of a course, when it is used, etc. This information could be used later by users when it comes to selecting the most pertinent materials.

2.3 Characteristical Quality Indicators

The characteristical category covers indicators based on the metadata that can draw on the potential of the information describing an educational resource.

Several authors have proposed this kind of indicator:

Ochoa and Duval [27] propose the use of metadata to order the results of a search for educational materials and be able to recommend the most pertinent. To be precise, they propose a set of relevance measures for educational materials applying the ideas used to make rankings of web pages, scientific articles, etc. Knowing which materials are most relevant from different points of view, will facilitate the task of choosing which educational resource to reuse. The information needed to estimate these relevance measures is obtained from the values of the user query, from the metadata of the educational materials, from usage records for the materials, and from contextual information.

Zimmermann et al. [40] remind us that to reuse an educational resource that was conceived for a particular context, it is often necessary to adapt it to the new context and propose an evaluation of the effort this adaptation requires. This adaptation to a new learning context may involve carrying out tasks, such as adapting the material to a new learning goal or to a new group of students, different from those it was created for, extracting a part of the content, or combining it with other educational materials. As for how we can find the learning materials, which are most easily adapted to our context, Zimmermann et al. [40] propose estimating the costs of adapting existing learning resources to a hypothetical ideal resource measuring the similarity of the metadata. One limitation of this idea is that the current specification of metadata provides insufficient information to support instructional utilization decisions [30].

Finally, Sanz et al. [31], [32] propose a reusability indicator based on metadata, which can be automatically calculated. We can determine measures of the quality of learning objects focusing on reusability [33]. While reusing learning objects is an empirical and

TABLE 1
Learning Object Quality Indicators

Indicator	Quality Dimensions Covered	Applied	Scale
Overall rating Content quality Effectiveness Ease of use	Explicit	Merlot	[0-5]
Comments	Explicit	Merlot	[0-5]
LORI indicators	Explicit	eLera	[0-5]
Lenses	Explicit Implicit	Connexions	Acts as a filter
Personal Collections	Implicit	Merlot	Numerical
Exercises	Implicit	Merlot	Numerical
Used in classroom	Implicit	Merlot	Numerical
Relevance metrics	Implicit Characteristical	Experiment	Numerical
Adaptation effort	Characteristical	Experiment	Numerical
Reusability	Characteristical	Experiment	[0-5]

observable fact, Sicilia [34] affirms that reusability is an intrinsic attribute of the object, which provides an a priori measure of quality which may be proven by posterior reuse data. This concept of reusability may be defined as the degree to which a learning object can work efficiently for different users in different digital environments and in different educational contexts over time. It should always be borne in mind that there are different technical, educational, and social factors that will affect reuse [29].

The idea that underlies this proposal is to identify the factors that most influence greater reusability of a learning object and then match them with metadata that offer information on them. Depending on the value of the metadata encountered, the reusability could be quantified.

To end this section, Table 1 presents a summary of the main quality indicators analyzed, indicating the quality dimensions they cover, the scale they use to measure their results, and where the indicators have been applied.

3 INTEGRATION OF QUALITY INDICATORS IN A MEASURE OF OVERALL QUALITY

Once the different quality indicators are identified and grouped by categories, a synthesized quality indicator that can support ranking learning objects according to their overall quality is proposed. This measure of overall quality will group all information on the quality of the materials and will be automatically calculated. Where some quality indicator is absent, it allows us to obtain a measure of overall quality based on the existing indicators. This will resolve the current dilemma, where materials that do not have an expert evaluation appear at the end of any search, and are automatically eliminated, as well as increasing the reliability of recommendations.

The following is a breakdown of the different indicators that might contribute to each of the dimensions of the rating. Fig. 1 details different sources of information that might be used to determine the explicit component of educational materials.

Fig. 2 details different sources of information that might be used to determine the implicit component of educational materials.

In spite of the information that the characteristical quality indicators might provide, they have not been taken into account in the proposed measure of overall quality, since they are not currently implemented in any repository.

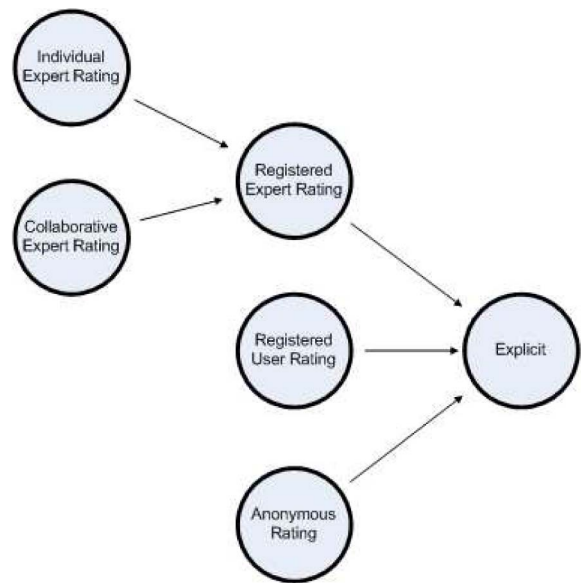


Fig. 1. Components of the explicit dimension.

We will now study two methods of integrating the information provided by the quality indicators, the Choquet Integral, which takes into account the possible redundancy derived from possible correlation between quality indicators, and a ranking algorithm for web searches called scaled footrule aggregation, where all indicators make an equal contribution, regardless of their possible correlation.

3.1 Choquet Integral

Due to the possibility that there may be some correlation between the quality indicators, Choquet's integral is the ideal candidate for modeling the aggregation process as it may be used as a generalization of the weighted arithmetic mean that takes into account correlation between criteria [22]. Two criteria $c_i, c_j \in C$ are correlated if there is a linear relationship between their values. This would introduce a certain degree of redundancy into the model.

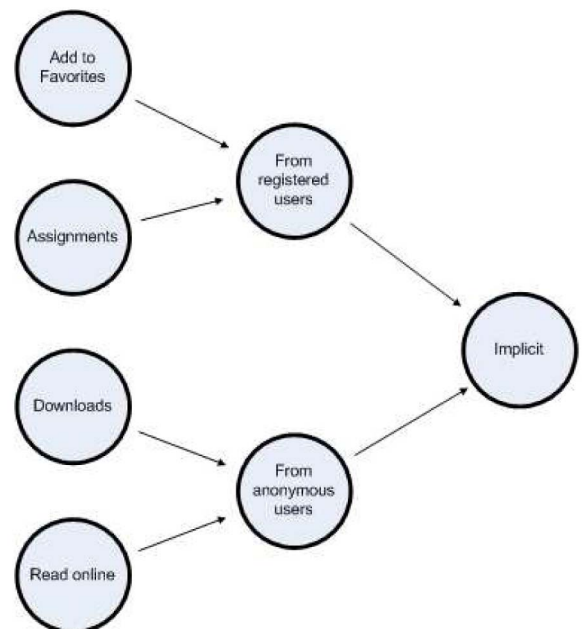


Fig. 2. Components of the implicit dimension.

TABLE 2
Merlot Quality Indicators

Dimension	Indicator
Explicit	Overall Rating
	Comments
Implicit	Personal Collections
	Learning exercises
	Used in classroom

A general expression of the integral is given in (1). The formula is a specific instance of the general form of the discrete aggregation operator on the real domain: $M_v : \mathbb{R}^n \rightarrow \mathbb{R}$, which takes an input vector $x = (x_1, \dots, x_n)$ and yields a single real value

$$C_v(x) = \sum_{i=1}^n x_{(i)} [v(\{j|x_j \geq x_{(i)}\}) - v(\{j|x_j \geq x_{(i+1)}\})], \quad (1)$$

where $x' = (x_{(1)}, \dots, x_{(n)})$ is a nondecreasing permutation of the x input n -tuple, where $x'_{(n+1)} = \emptyset$ by convention. The integral is expressed in terms of the Choquet v capacity. This measure, applied to an X set, is a monotonic set function $v : 2^X \rightarrow [0, 1]$, thus fulfilling $v(S) \leq v(T)$ when $S \subseteq T$, allowing for Choquet's capacity to assign weights not only to each criterion, but also to each subset of criteria.

To calculate the overall quality measure, all the quality indicators will be standardized in a range of values from [0-5], with average values shown where several ratings were available.

3.2 Scaled Footrule Aggregation (SFO)

Another way of determining the contribution each quality measure makes to the final ordering of resources is to use the results ranking algorithms for web searches. We could build differently ordered lists of resources for each quality indicator with partial lists, where there are objects that cannot be rated according to a particular indicator. To address the task of ranking a list of several alternatives based on many criteria, we choose the method scaled footrule aggregation because it is useful when we have partial lists [10].

Given the lists τ_1, \dots, τ_k with the positions of candidates, we define a weighted complete bipartite $graph(C, P, W)$ as follows: The first set of nodes $C = \{1, \dots, n\}$ denotes the set of learning objects to be ranked. The second set of nodes $P = \{1, \dots, n\}$ denotes the n available positions. The weight $W(c, p)$ is the total footrule distance (from the τ_i^c s) of a ranking that places element c at position p , given by (2):

$$W(c, p) = \sum_{i=1}^k |\tau_i(c) - p| \tau_i. \quad (2)$$

It can be shown that a permutation, minimizing the total footrule distance to the τ_i^c s, is given by a minimum cost perfect matching in the bipartite graph.

In contrast to the Choquet integral, this method does not consider possible correlation between indicators, and gives all quality indicators equal importance when calculating the global indicator.

4 ANALYSIS OF EXISTING CORRELATIONS BETWEEN THE DIFFERENT QUALITY INDICATORS

With the global quality measure that contemplates the different indicators now defined, relationships between them will be analyzed.

The study focuses on a set of 141 items selected from Merlot. This set of materials is the result of a query performed on October 1, 2009 to include all the materials stored in the repository between

TABLE 3
Kendall's Tau Correlation between Explicit Ratings

	Overall rating	Content quality	Effective-ness	Ease of use
Comments	0.096	0.107	0.126	0.172*

*. Correlation significant at the 0.05 level

2005 and 2008, which had been evaluated by experts and had associated user comments. The query was performed using the Merlot Material Advanced Search, and materials evaluated by users, and experts were chosen in order to have information from all dimensions. To carry out this study, we will use all of the quality indicators available in Merlot, as listed in Table 2. Personal Collections indicates the number of times materials are bookmarked, Exercises is course content that links to one or more of the materials, and Used in Classroom indicates whether the material has been used in class by the evaluating user. Obtaining values for these indicators in Merlot presents us with a series of different scenarios. For Overall Rating, Content Quality, Effectiveness, and Ease of Use are available in Merlot—for objects with a peer review—an indicator that shows a value for each of these. At the same time, in Comments, we can see the average user rating. Each object also has an accumulated value for Personal Collections and Exercises, automatically calculated by the repository. To calculate the Used in Classroom indicator, each user evaluation must be consulted in order to count those who have flagged it as used.

There followed a detailed study of the correlations between the indicators identified.

In Table 3, we can see how there is hardly any correlation between expert ratings and user ratings—only usability ease of use correlates with comments. This might be due to the fact that users do not have the knowledge needed to evaluate the material that they are analyzing, because this is an area or level they are unfamiliar with. It is also possible that users might give more importance to the ease of use in their global rating of educational materials. In addition, Han [14] points out that it is difficult to give a numerical value to user tastes in an evaluation of quality. For example, if a user prefers certain types of literature, he or she will give a better rating to the educational materials that cover those types. In any case, user evaluation can compliment that carried out by experts.

Table 4 illustrates the correlation between indicators from the explicit and implicit categories. A relationship is revealed between the tagging in favorites and expert evaluations.

TABLE 4
Kendall's Tau Correlation between Explicit and Implicit Ratings

	Personal Collections	Exercises	Used in classroom
Overall rating	0.171**	0.033	0.045
Content quality	0.145*	-0.014	0.034
Effectiveness	0.224**	0.047	0.123
Ease of use	0.146*	0.036	0.071
Comments	0.046	-0.007	0.049

** Correlation significant at the 0.01 level

* Correlation significant at the 0.05 level

TABLE 5
Kendall's Tau Correlation between Implicit Ratings

	Personal Collections	Exercises	Used in classroom
Personal Collections	1	0.227**	0.105
Exercises	0.227**	1	0.298**
Used in classroom	0.105	0.298**	1

** . Correlation significant at the 0.01 level
 * . Correlation significant at the 0.05 level

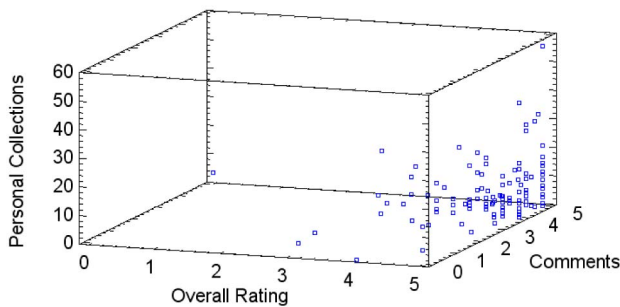


Fig. 3. Relationship between quality indicators.

In Table 5, we can see the correlation between different implicit measures. The presence in personal collections and the usage indicated by Exercises are related.

To illustrate the relationship between indicators taken from the different categories, Fig. 3 shows the relationship between Overall Rating, Personal Collections, and Comments.

The correlations detected between indicators from different categories support the idea that all are measures of quality obtained from different points of view, and which might complement one another, contributing to an indicator that rates the overall quality of a learning object.

5 RESULTS

With the correlations between different quality indicators analyzed, we now calculate the overall quality indicator using the two aggregation methods described above.

In order to avoid effects from identified correlations, we use Choquet's integral as a method of aggregation.

Choquet's capacities table is constructed, using the correlations identified, as shown in Table 6. The importance of each combination of criteria is represented by the symbol +, reflecting the presence of a criterion. The relationships between the capacities of the different criteria should satisfy certain restrictions that depend on the interactions detected between them. In this case, where only the correlation of criteria is presented, the following must be true for two correlative criteria i and j : $v(\{i, j\}) < v(\{i\}) + v(\{j\})$.

We applied the two quality indicator aggregation methods to the first 10 objects of the study sample, and found that the ordering produced by each is very similar, as shown in Fig. 4.

This could indicate that the advantage of the global quality indicator stems basically from the use of all identified quality indicators, regardless of how they are aggregated.

6 CONCLUSIONS

The correlations identified between the different quality indicators for learning objects, support the idea that they constitute different

TABLE 6
Choquet's Capacities Table

OR	+		+	+	+	+									
C	+		+				+	+	+						
PC		+					+			+	+				
LE			+					+			+	+			
UC				+					+			+	+	+	
$v(S)$	0.2	0.2	0.2	0.2	0.2	0.4	0.3	0.4	0.4	0.4	0.4	0.4	0.3	0.4	0.3
OR	+	+	+	+	+	+				+	+	+	+		
C	+		+	+			+	+	+		+	+	+	+	
PC	+	+					+	+	+	+	+	+	+	+	
LE		+	+	+				+	+	+	+	+	+	+	
UC				+	+	+			+	+	+	+	+	+	+
$v(S)$	0.5	0.4	0.5	0.6	0.6	0.5	0.5	0.5	0.6	0.4	0.6	0.5	0.7	0.7	0.6

OR=Overall Rating, C=Comments, PC=Personal Collections, LE=Learning exercises, UC=Used in classroom

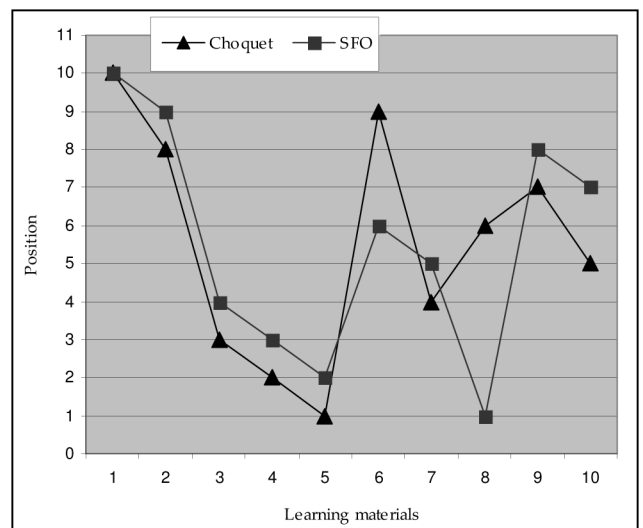


Fig. 4. Ranking learning materials.

views of their quality that might complement one another. An aggregate indicator could provide a measure of overall quality that took into account all available information, which would boost the reliability of recommendations. In addition, this measure could be calculated automatically, ensuring sustainability and allowing for all materials available in repositories to have a rating.

It should be stated that we have only studied the Merlot repository, and, for its results to be generalized, the scope of the study should be broadened. Therefore, in a forthcoming paper, the proposed procedure will be applied to the Connexions repository, where quality indicators covering all identified categories are available.

It would also be interesting to study how this overall quality indicator could be integrated into a recommendation system that contemplates other aspects relative to the context of reuse. In addition, it would be useful to incorporate some quality indicator of the characteristic dimension into a repository, in order to study how this type of indicator might enrich the overall quality measure.

ACKNOWLEDGMENTS

This work has been supported by the Comunidad de Madrid, Spain.

REFERENCES

- [1] Y. Akpınar, "Validation of a Learning Object Review Instrument: Relationship between Ratings of Learning Objects and Actual Learning Outcomes," *Interdisciplinary J. Knowledge and Learning Objects*, vol. 4, pp. 291-302, 2008.
- [2] R.G. Baraniuk, "Challenges and Opportunities for the Open Education Movement: A Connexions Case Study," *Opening Up Education: The Collective Advancement of Education through Open Technology, Open Content, and Open Knowledge*, T. Iiyoshi and M.S.V. Kumar eds., pp. 229-246, MIT Press, 2008.
- [3] T. Boyle, "Design Principles for Authoring Dynamic, Reusable Learning Objects," *Australian J. Educational Technology*, vol. 19, no. 1, pp. 46-58, 2003.
- [4] N. Boskic, "Faculty Assessment of the Quality and Reusability of Learning Objects," PhD dissertation, Athabasca Univ., 2003.
- [5] G. Brownfield and R. Oliver, "Factors Influencing the Discovery and Reusability of Digital Resources for Teaching and Learning," *Proc. 20th Ann. Conf. Australasian Soc. for Computers in Learning in Tertiary Education (ASCILITE '03)*, G. Crisp, D. Thiele, I. Scholten, S. Barker, and J. Baron, eds., pp. 74-83, 2003.
- [6] M. Claypool, P. Le, M. Wased, and D. Brown, "Implicit Interest Indicators," *Proc. Sixth Int'l Conf. Intelligent User Interfaces*, pp. 33-40, 2001, doi:10.1145/359784.359836.
- [7] S. Downes, "Models for Sustainable Open Educational Resources," *Interdisciplinary J. Knowledge and Learning Objects*, vol. 3, pp. 29-44, 2007.
- [8] E. Duval, "A Learning Object Manifesto—Towards Share and Reuse on a Global Scale," *Proc. Conf. Towards a Learning Soc. (E-Learning '05)*, pp. 113-117, May 2005.
- [9] E. Duval, "LearnRank: Towards a Real Quality Measure for Learning," *Handbook on Quality and Standardisation in E-Learning*, pp. 457-463, Springer, 2006.
- [10] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar, "Rank Aggregation Methods for the Web," *Proc. 10th Int'l Conf. World Wide Web*, pp. 613-622, 2001, doi:10.1145/371920.372165.
- [11] S. Fox, K. Karnawat, M. Mydland, S. Dumais, and T. White, "Evaluating Implicit Measures to Improve Web Search," *ACM Trans. Information Systems*, vol. 23, no. 2, pp. 147-168, 2005.
- [12] E. García-Barriocanal, Sicilia, and A. Miguel, "Preliminary Explorations on the Statistical Profiles of Highly-Rated Learning Objects," *Proc. Third Int'l Conf. Metadata and Semantic Research (MTSR '09)*, pp. 108-117, 2009.
- [13] R.L. Glass, "A Structure-Based Critique of Contemporary Computing Research," *J. Systems and Software*, vol. 28, no. 1, pp. 3-7, 1995.
- [14] K. Han, "Quality Rating of Learning Objects Using Bayesian Belief Networks," master's thesis, Simon Fraser Univ., 2004.
- [15] R. Kay and L. Knaack, "Assessing Learning, Quality, and Engagement in Learning Objects: The Learning Object Evaluation Scale for Students (LOES-S)," *Educational Technology Research and Development*, vol. 57, no. 2, pp. 147-168, 2009.
- [16] R.H. Kay and L. Knaack, "Evaluating the Learning in Learning Objects," *Open Learning: The J. Open and Distance Learning*, vol. 22, no. 1, pp. 5-28, 2007, doi:10.1080/02680510601100135.
- [17] C.M. Kelty, C.S. Burrus, and R.G. Baraniuk, "Peer Review Anew: Three Principles and a Case Study in Postpublication Quality Assurance," *Proc. IEEE '08*, vol. 96, no. 6, pp. 1000-1011, 2008.
- [18] V. Kumar, J. Nesbit, and K. Han, "Rating Learning Object Quality with Distributed Bayesian Belief Networks: the Why and the How," *Proc. Fifth IEEE Int'l Conf. Advanced Learning Technologies (ICALT '05)*, pp. 685-687, 2005.
- [19] E. Kurilovas and V. Dagiene, "Learning Objects and Virtual Learning Environments Technical Evaluation Criteria," *Electronic J. E-Learning (EJEL)*, vol. 7, no. 2, pp. 147-168, 2009.
- [20] J.Z. Li, J.C. Nesbit, and G. Richards, "Evaluating Learning Objects Across Boundaries: The Semantics of Localization," *J. Distance Education Technologies*, vol. 4, no. 1, pp. 17-30, 2006.
- [21] N. Manouselis and R. Vuorikari, "What If Annotations Were Reusable: A Preliminary Discussion," *Proc. Eighth Int'l Conf. Advances in Web-Based Learning (ICWL)*, pp. 255-264, 2009.
- [22] J.L. Marichal, "An Axiomatic Approach of the Discrete Choquet Integral as a Tool to Aggregate Interacting Criteria," *IEEE Trans. Fuzzy Systems*, vol. 8, no. 6, pp. 800-807, Dec. 2000.
- [23] R. McGreal, T. Anderson, G. Babin, S. Downes, N. Friesen, K. Harrigan, M. Hatala, D. MacLeod, M. Mattson, G. Paquette, G. Richards, T. Roberts, and S. Schafer, "EduSource: Canada's Learning Object Repository Network," *Int'l J. Instructional Technology and Distance Learning*, vol. 1, no. 3, 2004.
- [24] J.C. Nesbit and K. Belfer, "Collaborative Evaluation of Learning Objects," *Online Education Using Learning Objects*, R. McGreal, ed., RoutledgeFalmer, pp. 124-138, 2004.
- [25] J.C. Nesbit and J. Li, "Web-Based Tools for Learning Object Evaluation," *Proc. Int'l Conf. Education and Information Systems: Technologies and Applications*, pp. 334-339, 2004.
- [26] X. Ochoa and E. Duval, "Quality Metrics for Learning Object Metadata," *Proc. World Conf. Educational Multimedia, Hypermedia and Telecomm. (EDMEDIA '06)*, E. Duval, R. Klamma, and M. Wolpers, eds., pp. 1004-1011, 2006.
- [27] X. Ochoa and E. Duval, "Relevance Ranking Metrics for Learning Objects," *IEEE Trans. Learning Technologies*, vol. 1, no. 1, pp. 34-48, Jan. 2008.
- [28] X. Ochoa and E. Duval, "Measuring Learning Object Reuse," *Proc. Third European Conf. Technology Enhanced Learning: Times of Convergence*, E. Duval, R. Klamma, and M. Wolpers, eds., pp. 322-325, 2008.
- [29] K. Palmer and P. Richardson, "Learning Object Reusability Motivation, Production and Use," *Proc. 11th Int'l Conf. Assoc. for Learning Technology (ALT)*, 2004.
- [30] M.M. Recker and D.A. Wiley, "A Non-Authoritative Educational Metadata Ontology for Filtering and Recommending Learning Objects," *Interactive Learning Environments*, vol. 9, no. 3, pp. 255-271, 2001.
- [31] J. Sanz, J.M. Dodero, and S. Sanchez-Alonso, "A Preliminary Analysis of Software Engineering Metrics-Based Criteria for the Evaluation of Learning Objects Reusability," *Int'l J. Emerging Technologies in Learning (IJET)*, vol. 4, no. 1, pp. 30-34, 2009, doi:10.3991/ijet.v4s1.794.
- [32] J. Sanz, S. Sánchez-Alonso, and J.M. Dodero, "Reusability Evaluation of Learning Objects Stored in Open Repositories Based on Their Metadata," *Proc. Third Int'l Conf. Metadata and Semantic Research (MTSR)*, vol. 46, pp. 193-202, Oct. 2009.
- [33] M.A. Sicilia and E. García-Barriocanal, "On the Concepts of Usability and Reusability of Learning Objects," *Int'l Rev. Research in Open and Distance Learning*, vol. 4, no. 2, 2003.
- [34] M.A. Sicilia, "Reusability and Reuse of Learning Objects, Myths, Realities, and Possibilities," *Proc. First Pluri-Disciplinary Symp. Design, Evaluation, and Description of Reusable Learning Contents*, 2004.
- [35] A. Tzikopoulos, N. Manouselis, and R. Vuorikari, "An Overview of Learning Object Repositories," *Learning Objects for Instruction: Design and Evaluation*, pp. 44-64, Idea Group Publishing, 2007.
- [36] J. Vargo, J.C. Nesbit, K. Belfer, and A. Archambault, "Learning Object Evaluation: Computer-Mediated Collaboration and Inter-Rater Reliability," *Int'l J. Computers and Applications*, vol. 25, no. 3, pp. 198-205, 2003.
- [37] R. Vuorikari and H. Poldoja, "Comparison of Educational Tagging Systems—Any Chances of Interplay," *Proc. Second Workshop Social Information Retrieval for Technology Enhanced Learning*, 2008.
- [38] R. Vuorikari, N. Manouselis, and E. Duval, "Using Metadata for Storing, Sharing, and Reusing Evaluations in Social Recommendation: The Case of Learning Resources," *Social Information Retrieval Systems: Emerging Technologies and Applications for Searching the Web Effectively*, D.H. Go and S. Foo, eds., pp. 87-107, Idea Group Publishing, 2008.
- [39] N.Y. Yen, F.F. Hou, L.R. Chao, and T.K. Shih, "Weighting and Ranking the E-Learning Resources," *Proc. Ninth IEEE Int'l Conf. Advanced Learning Technologies (ICALT '09)*, pp. 701-703, 2009, doi:10.1109/ICALT.2009.36.
- [40] B. Zimmermann, M. Meyer, C. Rensing, and R. Steinmetz, "Improving Retrieval of Re-Usable Learning Resources by Estimating Adaptation Effort Backhouse," *Proc. First Int'l Workshop Learning Object Discovery and Exchange*, <http://fire.eun.org/lode2007/lode.pdf>, 2007.