Quality Evaluation of Open Educational Resources[*]

Mirette Elias[1], Allard Oelen[2], Mohammadreza Tavakoli[2],
Gábor Kismihok[2], and Sören Auer[2]

[1] Fraunhofer IAIS and University of Bonn, Germany {melias@uni-bonn.de}
[2] TIB Leibniz Information Centre for Science and Technology, Germany
{allard.oelen,reza.tavakoli,gabor.kismihok,soeren.auer}@tib.eu

Abstract. Open Educational Resources (OER) are free and open-licensed educational materials widely used for learning. OER quality assessment has become essential to support learners and teachers in finding high-quality OERs, and to enable online learning repositories to improve their OERs. In this work, we establish a set of evaluation metrics that assess OER quality in OER authoring tools. These metrics provide guidance to OER content authors to create high-quality content. The metrics were implemented and evaluated within SlideWiki, a collaborative Open-CourseWare platform that provides educational materials in presentation slides format. To evaluate the relevance of the metrics, a questionnaire is conducted among OER expert users. The evaluation results indicate that the metrics address relevant quality aspects and can be used to determine the overall OER quality.

1 Introduction

OpenCourseWare (OCW) is defined as free and open digital publication of educational and learning content [9]. OCW platforms organize education materials, known as Open Educational Resources (OER), in the form of online courses. These courses generally provide a learning plan and evaluation tools. Many OCW platforms exist (e.g., MIT[3]) with various OER representations, such as videos, audio and slides. Finding high-quality OERs becomes increasingly cumbersome due to the growing amount of published resources [10]. However, selecting high-quality resources is crucial to ensure the quality of an online course. In this work, we propose evaluation metrics to assess the quality of OERs. The metrics are implemented within SlideWiki[4], a collaborative OCW platform focused on presentation slides. The implementation demonstrates how the metrics can be integrated within OER authoring tools. Although the presented metrics are evaluated on presentation slides, they can be applied to other OER representations as well.

This article addresses two research question: 1) *how to evaluate the quality of OER material?* and 2) *how to use this evaluation to guide OER authors and*

---

[3] https://ocw.mit.edu
[4] https://slidewiki.org

learners? In order to define the quality metrics and to develop the implementation, we investigate related work to OER quality assessment. Accordingly, we select and extend the dimensions that are related to content representation, and we define a set of metrics for each dimension. Finally, we evaluate our work by conducting a questionnaire with OER expert users (i.e., instructors and PhD students) and by implementing a set of the metrics in an authoring tool.

The remainder of this paper is structured as follows: Section 2 discusses and analyzes the state-of-the-art of evaluation approaches used for assessing the quality of online educational systems and contents. Section 3 defines a list of quality metrics and describes each of them. Section 4 explains the implementation of the metrics and evaluates the results. Finally, Section 5 concludes this work.

2 Related Work

This section reviews the state-of-the-art of the quality evaluation approaches for OER repositories. We focused on approaches that address quality aspects related to the OER content and representation. We analysed dimensions found in the literature and categorized them based on the quality aspects: 1) feature quality (i.e., quality related to functionalities provided by the OER repository), 2) technological quality (i.e., quality related to the technology and implementation of the OER repository), and 3) content quality (i.e., quality related to the OER material and content representation). Table 1 shows the dimensions that were extracted and categorised as per our analysis.

Table 1: Summary of quality evaluation dimensions

References / Dimensions		[13]	[7]	[2]	[6]	[5]	[1]	[11]	[8]	[4]	[3]
Features quality	Availability	✓	✓	✓							✓
	Multiliguality	✓	✓	✓		✓					
	Reusability	✓		✓		✓		✓	✓		
	Provenance	✓						✓			✓
	Recency	✓						✓			✓
	Openness	✓	✓	✓	✓		✓	✓		✓	
Technological quality	Accessibility		✓		✓	✓	✓	✓	✓	✓	
	Alignment to standards		✓	✓	✓		✓		✓		✓
	Usability		✓		✓	✓		✓	✓	✓	
	Compatability		✓			✓		✓	✓		
Content quality	Structure		✓					✓	✓	✓	
	Accuracy		✓		✓			✓	✓	✓	
	Comprehensiveness		✓		✓		✓	✓			✓
	Discoverability	✓	✓	✓							
	Multimodality				✓	✓	✓	✓	✓	✓	✓
	Self-assessment	✓	✓		✓		✓				

From the analysis, we found that most of the evaluation approaches that were studied in Table 1, evaluate the dimensions and metrics either conceptually or by providing a checklist to experts or users. These checklists are either filled out manually or in the form of online surveys [14]. Automatic OER quality assessment and author quality guidance were not addressed. Since this study is

focusing on evaluating the quality of *OER materials*, we focus on the dimensions defined in the content quality part from Table 1 and extended them in Table 2. We also use accessibility and compatibility from the technological aspect because they address OER content as well.

3 Proposed OER Quality Evaluation Metrics

The Open Education Consortium (OEC) defines OERs as materials that are composed of course planning, thematic content, and assessment tools [9]. Accordingly, we divided our evaluation approach of OERs into three components: content structure, learning content and self-assessment. Content structure defines the organization and navigation of the educational resource. Learning content refers to representation of the learning material. Self-assessment is related to the availability of questions to evaluate the learning process. Table 2 lists the dimensions and metrics to assess the quality of OER materials.

Table 2: OER quality metrics

Dimension	Metrics	Description
Content Structure (CS)	CS1. Clearness of the taxonomies	CS1.1 Short and descriptive name (i.e., characters limit) CS1.2 Coherence with content title (i.e., consistent file name with the content title) CS1.3 Progress inference from title (i.e., consistent coding scheme)
	CS2. Easiness of navigation	CS2.1 Hierarchical design (i.e., well-organized structure) CS2.2 Depth of the taxonomy (i.e., less scrolling)
	CS3. Adaptability of the structure	CS3.1 Availability of adaptability mechanism (e.g. smaller chunks design)
	CS4. Discoverability of the content [12]	CS4.1 Availability of Standardized Metadata (i.e., sum of the normalized importance scores of metadata) CS4.2 Adherence to Standardized Metadata (i.e., including a Rating function)
Learning Content (LC)	LC1. Quality of text	LC1.1 Correctness of text spelling and grammar LC1.2 Comprehensiveness of text (i.e., using readability meters)
	LC2. Adaptability of content	LC2.1 Availability of various content formats (e.g., based content, web media, interactive media, video, audio) LC2.2 Availability of multiple content representation (e.g., multiple themes for learning slides) LC2.3 Consistency between the content types (i.e., synchronized maintenance and versioning management)
	LC3. Compatibility of content on multiple devices	LC3.1 The number of supported devices (e.g., mobile phone, tablet, laptop, assistive technologies) LC3.2 Availability of compatibility checking mechanisms (e.g., validating responsiveness of web pages)
	LC4. Accessibility of content representation	LC4.1 Compliance to guidelines of content representation (e.g., WCAG 2.1 guidelines) LC4.2 Availability of validation approach of content representation (e.g., validating that an image contains alternative description to support accessibility)
	LC5. Multilinguality of content	LC5.1 Availability of resources in more than one language (i.e., other than English) LC5.2 Existence of translation approach (i.e., automatic translation, expert-revised) LC5.3 Availability of synchronization of material translation
Self-assessment (SA)	SA1 Availability of self-assessment	SA1.1 Existence of self-assessment content SA1.2 Availability of answers SA1.3 Average number of question covering the content (i.e., number of questions per each learning object) SA1.4 Existence of question generation approach (e.g., automatic generation or author entry)
	SA2. Variety of self-assessment questions	SA2.1 Available type of questions (e.g., multiple choice, close text, sorting). SA2.2 Average number of question per assessment type

4   Implementation and Evaluation

For the implementation, a set of eight metrics (i.e., CS1.1, CS2.1, CS4.1, CS4.2, LC4.1, LC5.1, SA1.1, SA1.2) has been selected from Table 2. The set of quality metrics was selected based on relevance, appropriateness and technical viability within the SlideWiki platform. Figure 1 shows a quality report within the SlideWiki user interface. The figure displays a quality report from a deck that is available via SlideWiki[5]. The quality report is displayed on deck-level, and is visible for all users. There are several reasons for making the quality report public. Firstly, there is an extra incentive for OER creators to ensure that their presentation meets a certain quality standard. Secondly, learners can decide to use an OER based on its quality. And finally, due to the collaborative nature of SlideWiki, learners can help improving the slide deck based on the report. For each metric, the amount of detected issues is listed. In case no issues are found, the text "All good" is displayed. For metrics CS4.1 and CS4.2 a quality score is shown. Listed metrics can be expanded to view more details about a particular metric, including why adhering to this metric is important. In case an issue is detected, more information about this issue is displayed.
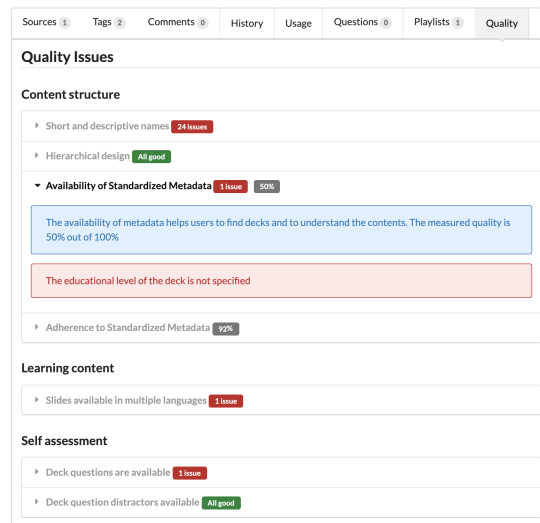

Fig. 1: Quality scoring report within SlideWiki

To evaluate our quality dimensions and metrics, we invited OER expert users (either university instructors or PhD students) and asked them about the importance (1: less important, 5: very important) of our metrics in each dimension with the help of a qualitative questionnaire[6]. Moreover, the participants provided opinions about the overall quality of existing OERs (as *Current Quality*

---

[5] http://slidewiki.org/deck/90789/02-rdf-data-model/deck/90789
[6] https://forms.gle/2Y4bhzbEK3LTY5y78

column, 1: lowest quality, 5: highest quality), and overall usefulness (1: not useful, 5: very useful) of our metrics in each dimension. We collected the feedback of ten participants who had experience with OERs as author (2 participants), learner (5 participants), and teacher (5 participants). The evaluation results of each dimension and metric are: 1) *Content Structure* is considered useful by 100% of the participants, 2) *Learning Content* is considered useful by 60% of the participants, and 3) *Self-assessment* is considered useful by 80% of the participants. Regarding the usefulness and coverage of the proposed dimensions and metrics, 70% of the participants find our dimensions and metrics useful and 50% of the participants agreed that the proposed dimensions and metrics cover the important metrics needed for evaluating the quality of OER materials, while 30% of the participants provided a neutral response.

5   Conclusion

This paper proposes quality evaluation metrics for OERs to help learners and teachers to find high-quality OERs and guide OER repositories to improve their content. Two research questions were addressed in this article. To answer the first question "how to evaluate the quality of OERs", we established and distributed quality evaluation metrics covering three aspects of OER quality assessment: content structure, learning content and self-assessment. For the second question "how to use the evaluation metrics to guide authors and learners of OERs", we selected seven of these metrics and implemented them in SlideWiki. Quality reports are publicly visible for all users in order to help learners find high quality content, and encourage authors to improve their materials. We evaluated our metrics by collecting feedback from OER users and creators via a questionnaire. As future work, we will continue implementing the metrics in SlideWiki and study the effect of the evaluation reports on the learners' selection of OERs, and authors' ways of creating and editing OERs.

References

1. A. Achieve. Rubrics for evaluating open education resource (oer) objects. *Washington, DC: Achieve, Inc. Retrieved January*, 9:2013, 2011.
2. J. Atenas and L. Havemann. Questions of quality in repositories of open educational resources: a literature review. *Research in Learning Technology*, 22, 2014.
3. M. Custard and T. Sumner. Using machine learning to support quality judgments. *D-Lib Magazine*, 11(10):1082–9873, 2005.
4. M. Haughey and B. Muirhead. Evaluating learning objects for schools. *E-Journal of Instructional Science and Technology*, 8(1), 2005.
5. E. Kurilovas et al. Methodology for Evaluating Quality and Reusability of Learning Objects. *Electronic Journal of e-Learning*, 9(1):39–51, 2011.
6. H. Leary et al. Developing and using a guide to assess learning resource quality in educational digital libraries. *Digital libraries-methods and applications*, pages 181–196, 2011.

7. G. Moise et al. MASECO: A multi-agent system for evaluation and classification of OERs and OCW based on quality criteria. In *E-Learning Paradigms and Applications*, pages 185–227. Springer, 2014.
8. J. Nesbit and J. Li. Web-based tools for learning object evaluation. In *International conference on education and information systems: Technologies and Applications*, pages 21–25, 2004.
9. OEC. What is Open Courseware? https://www.oeconsortium.org/faq/what-is-open-courseware/.
10. S. Palvia et al. Online Education: Worldwide Status, Challenges, Trends, and Implications. *Journal of Global Information Technology Management*, 21(4):233–241, 2018.
11. M. Pérez-Mateo et al. Learner generated content: Quality criteria in online collaborative learning. *EU Journal of Open, Distance and E-Learning*, 14(2), 2011.
12. M. Tavakoli, M. Elias, G. Kismihók, and S. Auer. Quality prediction of open educational resources a metadata-based approach. *arXiv:2005.10542*, 2020.
13. S. Vahdati et al. Opencourseware observatory: does the quality of opencourseware live up to its promise? In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, pages 73–82, 2015.
14. M. Yuan and M. Recker. Not all rubrics are equal: A review of rubrics for evaluating the quality of open educational resources. *International Review of Research in Open and Distributed Learning*, 16(5):16–38, 2015.