Quality Assurance for Open Educational Resources: The OERTrust Framework

Douglas Almendro
Postgraduate Program in Electrical Engineering and Computing
Mackenzie Presbyterian University
São Paulo, Brazil

Ismar Frango Silveira
Postgraduate Program in Electrical Engineering and Computing
Mackenzie Presbyterian University
São Paulo, Brazil

Abstract. Learning Objects have met some barriers to their development and effective adoption, which varied from the lack of quality assurance mechanisms to the impossibility of editing and adapting most of them to real teaching and learning contexts. However, with the advent of OER (Open Educational Resources), if the later problem – to retain, reuse and even remix learning content – was meant to be solved, the same could not be said for the first one. Quality assurance is still an unsolved problem in this context, even more complex due to the possibility of versioning and collaborative design brought by OER. Thus, it is necessary to propose validation mechanisms for them, at least establishing some guarantees about their functionality and quality. In this sense, this work aims to discuss OERTrust, a proposal of supporting framework for OER validation and testing process, considering both versioning and remixing features. OERTrust is based on the principles of validation and testing that come from Software Engineering area and relies on fuzzy logic to define the importance and influence of different tests to each kind of OER.

Keywords: Open Education Resources; Quality; Testing; Software Engineering.

Introduction
In the context of Education, the discussion about OER has been gaining more and more space in recent years. OER broader definition (UNESCO, 2002) can cover complete courses, parts of courses, modules, textbooks, research articles, videos, questionnaires, simulators and other materials that can support education, provided they are designed under the principles of openness.
According to the principles of openness, OER may undergo different types of modifications, including revisions and remixes. These modifications are not, \textit{a priori}, prerogative of OER authors, since they could be carried out in practice by any individual or organization. Such flexibility could generate success or failure, depending on how and why modifications were made, which could have a significant impact on the quality of OER. Quality, even though being a concept laden with subjectivity, is some objective pursued in different areas of human life, including education (KAWACHI, 2014).

When designing an OER, a set of technical and didactic requirements should be followed, including content coverage on a given topic and adequacy to some pedagogical standards, for instance. At the time of the revision or remixing of an OER, even though the requirements have already been outlined, any modification made may affect the quality of the original OER, in addition to being prone to defects, failures or even errors – from the technical, didactic or content’s points of view. For this, all modifications must be tested to ensure the minimum satisfaction of some quality criteria that could have been previously established. According to Pressman (2005), any change in a product, although simple, should be tested. The activity of testing modifications in an artifact of a new or revised / remodelled OER could subside customizers by indicating them some limits about what could be modified and what are the consequences of the modification.

Thus, this paper presents OERTrust, a framework that is aimed to help OER builders – from professional ones, like instructional designers, to teachers or other educational agents – to verify and assure the quality of a given OER. In this way, one of the goals of this work is to provide, together with a framework, also an indicator of reliability in relation to the types of tests that must be performed to guarantee the quality of an OER.

**Theoretical framework**

**Open Educational Resources (OER) state-of-art**

In this paper, topics for the understanding the progresses of the research on OER will be addressed. As this field has multidisciplinary characteristics, it was necessary to carry out research in the following areas of knowledge: education, distance education, open educational resources, software engineering, tests and quality. Thus, following lines discuss the main concepts around Open Educational Resources, fundamental for this work.

According to UNESCO (2017), OER, "are any kind of material used in education, which is in the public domain or is introduced with an open license." They are open materials and can be freely copied, adapted, used, remixed and even re-shared. The possibility of reuse of open educational resources, especially digital resources, which tend to be disseminated more easily, may include "curriculum maps, course materials, textbooks, streaming videos, multimedia
applications, podcasts, and any other materials that have been designed for use in teaching and learning” (KANWAR and STAMENKA, 2015, p. 5).

Wiley (2007) explains that the OER context has its origins in efforts to standardize and conceptualize Learning Objects (LO). With the increasing evolution and use of LO, Wiley defined the Open Content concept in 1998 and created the Open Content License / Open Publication License, aiming at the massive popularization of Free Software movement’s concepts, applied to the development of educational contents. The same author points out that with the rapid spread of the idea of open educational contents, Larry Lessig and other members of Harvard Law School, founded Creative Commons in 2011 and with it, a versatile set of licenses.

In 2005, Albright (2005) sponsored an event to discuss the availability of educational resources in a universal way, highlighting the term “Open Educational Resources”, with the following definition brought by UNESCO in 2002, during the Forum on Open Courseware: "teaching, learning and research materials in any medium, digital or otherwise, that reside in the public domain or have been released under an open license that permits no-cost access, use, adaptation and redistribution by others with no or limited restrictions" (UNESCO, 2012).

The 2012 Paris Declaration (UNESCO 2012) notes that the Open Educational Resources (OER) promote the objectives of international declarations, including the Universal Declaration of Human Rights (Article 26.1), which states that "Everyone has the right to education". The same document provides a series of recommendations to stakeholders on teaching, research and policy issues. It particularly encourages governments to make available educational materials that are funded through public funds with open licenses.

Recently, the 2nd OER World Congress also held by UNESCO (2017), whose theme was "OER for inclusive and egalitarian quality education: from commitment to action", focused on identifying strategies to take advantage of the potential of OER to achieve an inclusive and equal quality education for lifelong learning for all by 2030. The objectives of this congress focused on examining solutions to address the challenges of integrating OER practices into educational systems around the world, showing the world's best practices in OER policies and initiatives, and generating recommendations for the adoption of OER with focus on best practices. Still in this document, it is stressed out that OER offer the potential to improve the benefits of learning, lower costs and improve teaching quality, enabling effective knowledge sharing, which is fundamental to support the implementation of SDG 4 Sustainable Development Goals.

On the other hand, Wiley (2014) points out the need for OER to follow the so-called 5R permissions – Retain, Reuse, Revise, Remix and Redistribute. In this sense, the same principles of openness that give greater flexibility to OER also raise questions about their quality. Reflections on the need for quality standards and testing mechanisms become more and more important. As long as several customizations and versioning of an OER are to be performed, the need to create mechanisms for OER quality assessment becomes clear, since some version of it may not execute, not install, perform inefficiently, present errors, defects or failures, from many different points of view.
Big and little OER

Weller (2010) comments that much of the focus on OER has been around well-known, privately-funded OER projects such as MIT Open Courseware and Open University’s Open Learn projects. These projects helped to popularize the Open Education movement, generating a considerable amount of OER downloads. By broadening the definition of OER, resources can be produced by individuals and shared on sites outside the commonly used repositories. The same author cites as example, YouTube, Slideshare and Flickr as examples of sites that provide resources that can vary in granularity, quality and potential use for learning. Also according to the same author, the "big" OER are generated institutionally, arising from projects like Open Learn. These types of OER generally strive to achieve high quality, following explicit teaching purposes, as well as a methodology and systematization for the production of OER. As a counterpoint, the author points to the existence of "small" OER, which are produced individually, usually having a low cost of production, not necessarily involving sophisticated production teams or media resources. They can be potentially produced by anyone, not only educators.

In assuming this categorization of OER as large and small, whether produced institutionally or individually, one must look at some of the issues that arise in this context. One of the main differences between the two types of OER relies on their intentionality, with large OER being created for a specific learning purpose, whereas small OER being created to fulfil a variety of requirements or motivations, but one cannot deny their potential for learning. In addition, small OER have a greater theoretical remix potential, being able to be remixed by any other individual or institution. On the other hand, such kind of OER usually lack of contextual information, which could make its reuse or remix harder. This leads to a discussion about granularity of OER, which can be observed with more detail in Silveira (2004).

Camilleri (2014) comments that there are significant differences between how different types of OER are used and interpreted by the public. Such differences relate to aspects of quality, author’s reputation and ease of production. It may be that a mix of both types of OER, big and small, could be the best route to the realization of an open education setting. Bigger OER are a useful means of increasing access to open education and an initial way of approaching reuse, as they overcome many of the objections based on quality and reliability, since they usually have some sort of institutional endorsement. Small OER, in their turn, represent a more dynamic model that encourages the participation of various players in the educational process and may be more sustainable in the long term. For students, a mixture of both can thus result in a varied and interesting learning experience.

Quality and Testing in Software Engineering

According to Hoyle (2006) and ISO 9000 (2000), quality is "the degree to which a set of characteristics inherent to a product, process or system meets the requirements initially stipulated for them". From the general concept of quality, one can reach several definitions that can be applied in the context of Software – or extended in order to embrace software-generated content/media.
Quality is an element that must be present during all production process, whether of physical or intangible objects, such as digital artefacts. In this way, testing helps to increase the quality of any product, allowing errors to be corrected and not be relegated to the end of the process. Thus, quality standards and test protocols become key elements in the production process of any product.

In the context of Software Engineering, according to Pressman (2005), "Software Quality is the compliance with functional and performance requirements that have been explicitly stated, clearly documented development standards, and the implicit characteristics that are expected of all software developed by professionals." The same author indicates that in the building of a software, the sooner the error is discovered, the less impact it has on budget and time.

In the context of this study, the term "tests" or “testing” refers to the original context of Software Engineering, expanded to OER. In a Software Engineering scenario, the testing process aims to achieve two distinct objectives: to indicate that the system meets the specified requirements and also that the software does not behave in a different way than was specified (RIOS, 2013).

Matté (2011) points out that, in general, software testing should still provide for testing not only with the product, but also in the process. In order to a software reach the expected quality levels, it is fundamental that other element such as: system architecture, functional and non-functional requirements, development methods and tools, and others are to be tested. Thus, to find errors or failures it is necessary the application of some methodology and adoption of processes that can help the life of the software developers. The authors believe that is the same case for OER – which will be discussed in next section.

**Quality and testing on Learning Objects and OER**

The process of creating LO or OER sometimes presents many similar aspects to the software creation process – for instance, simulations and educational games requires many of the processes that “regular” software do. Although, according to the type of OER, its creation process is more consistent with media creation processes, as in the case of educational videos or animations. It is worth remembering that lessons learned in relation to Software Quality can be inspiring in the process of quality assessment of OER, if not straightforwardly applied.

Hylén (2012) comments that there are materials provided for free that are not considered of having “high quality”, including some created by institutions with local or international reputation – pointing out that reputation or recommendation could be not the only answer to assure quality in the educational context, as in Conesa (2012), for instance. New ways of ensuring and evaluating quality have to be developed, along with traditional quality assurance methods and mechanisms, mostly inspired by other elements of the openness movement.

The peer evaluation of OER learning materials can be combined with comments and user classifications on the web, as seen in Albuquerque and Silveira (2013), for example. In this sense, the adoption of OER by educational institutions could offer great value in terms of reputation quality. As an example, Rodés et al. (2012) discuss a pilot project for the adoption of open textbooks in the context of...
nine Latin American universities. Because the content created in all of the above initiatives is distributed under a Creative Commons Attribution (CC-BY) license, they can be easily shared with other educational sectors, which would be able to adapt them to their specific needs.

In this context, the discussion on quality aspects comes with a lot of emphasis on the openness aspects that surround the OER. As in Software Engineering, the development of OER involves a series of activities in which human failures may occur and errors may arise at some stages of the life cycle, from incomplete or imperfect requirements specification, to later stages of design, implementation and development. The fact that OER are susceptible to multiple versions and modifications makes this discussion even more broad and complex, which was not the case with LO, whose elements for quality are synthesized by Velásquez et al. (2005) as:

- Technological elements:
- Pedagogical Elements:
- Elements of Content.
- Aesthetic elements / interface

Evolving from the context of LO to OER, Kawachi (2014) brings a set of guidelines for product quality and the process of elaborating an OER called TIPS framework (Teaching and learning process; Information and material content; Presentation, product and format; System, technical and technology). This framework has four dimensions: teaching and learning processes; information and material content; presentation, product and format; and systems, techniques and technology.

Both the references of tests and quality for OA and OER mentioned start from the premise of meeting a set of pedagogical requirements, which in turn relate in a non-linear way with the classic concepts of functional and non-functional requirements of Software Engineering. Furthermore, considering that the terms "requirements" and "quality" do not apply consistently to the software industry, the same is true in the context of OER. What is known is that the end user expects both software and an OER to have a reasonable "quality" degree - which can be highly subjective. It is known, for example, that every student is frustrated when an OER stops functioning or presents some kind of failure or malfunction, whether caused by error or defect regarding to technical, pedagogical or content aspects.

**Framework Proposal - Mapping software testing to OER**

Considering the most common small OER types, Figure 01, based on Almendro and Silveira (2017) is here proposed, associating them with the types of software tests. This systematization is intended to guide which tests will be most adequate to guarantee the quality of OER. Figures 1 to 3 are graphical representations of an underlying adjacency matrix, which can be seen as chord diagrams. Figure 1 depicts the relationship of all test types from Software Engineering to the different types of small OER.
Figure 1. Test interconnections and OER types. Source: author

Figure 2. Interconnection of tests with OER Software type
As can be seen in Fig. 2 and Fig. 3, the types of OER that fit all tests are OER that fit the classical definition of software: educational software and games (put together with simulators and animations). In this sense, some considerations must be made regarding each of the tests coming from Software Engineering and its proposal of application in the OER context. Possibly, the main one is that not every OER fits into the “classical” software definition: according to Sommerville (2013), software is characterized as a computer program and all the documentation associated with it. Pressman (2005) defines software as a logical, non-physical element of a system that does not wear out, as it is commonly used in Software Engineering. With this definition in mind, Software Engineering tests should be taken, in some cases, just as reference elements for the preparation of tests appropriate to the context of OER.

It is a fact that OER developers run tests in their development process. But they are almost invariably performed informally, without the same rigor and testing procedures as exist in the software industry. One of the reasons is the cost of executing tests, which require a certain planning, or application of appropriate techniques, or the use of appropriate support tools that impact the time or the budget of an OER project.

Software testing, in general, is directed to a single feature of the software. For example, black-box testing targets the ultimate functionality of the software, stress tests analyse the performance of an application, and so on. In the case of OER, each of the tests adapted to this context should take into account at least three different dimensions to be tested in each OER, namely: the technical dimension, the didactic dimension and the content dimension. Fig. 4 shows a graphic schema of this concept.
Figure 4. Graphic representation of the technical dimension, the pedagogical dimension and content dimension.

This proposal involving three dimensions of performance for each test is a condensation of the organization proposed by Velásquez et al. (2005), which proposed four dimensions - the fourth dimension refers to the interaction aspects. Once the aspects of usability are considered as transversal to the didactic, content and technical dimensions, this dimension permeates all the others.

Considering these aspects, Almendro and Silveira (2017) present a brief classification where each type of test is directed to the verification of certain types and characteristics of OER, namely:

- **Unitary test** applies only to simulation, games, and software OER, since the main core of this type of test is to test a part of the programming involved in a software module.

- **Integration test** has its importance due to the joining of several parts of the same type of instrument, so there was justification for using the same in the OER of simulation / animation / games, software and questionnaires.

- **Operational test** covers all types of OER, as everyone must be ensuring that they are operational without encountering malfunctions.

- With respect to the **positive-negative test**, it is assumed that the resource will be executed with harmony for the intended purpose. Therefore, OER, audio, video, simulation / animation / games, software and questionnaires, do not enter images and texts for this type of test.

- All OER must pass the **regression test**, because if anything is rectified, it must be tested throughout the application again.

- The **black-box** test is applied to check the output results, i.e. given an input is assumed to have an already planned output, in this case the OER, which do not require this type of test are the Image and Text.

- Only the simulation / animation / game OER, software and questionnaires, must pass the **white-box test** since the objective is to test code. Because it may occur that parts of the code have never been tested. Moreover, that should go through all the paths of the code.

- The **functional test** is to validate the functionalities described in the documentation as soon as all the OER must pass through it.
• All OER types may pass the **usability test**, since the basic usability requirements must cover any digital artefacts, independent of their nature.

• The **performance test** should be applied also to all OER because it is the need to size the time and even the type of equipment that must be used to manipulate that particular type of OER.

• The main purpose of the **load test** is to verify the operation of the application with the use of a large number of concurrent users, which is why all OER must pass this type of test.

• The **user acceptance test** is also applicable to all OER, since potential users - teachers and students - should be able to assess their relevance to their specific teaching or learning needs.

• **Volume test** is equally universal, since for all OER it is necessary to verify the balanced amount of users using OER without performance problems.

• Testing the application in unexpected situations is what the **stress test** aims at; it should be used by all OER to ensure that different forms of use can be supported.

• Testing whether the application works properly in different hardware or software environments is what the **configuration test** aims at so that all OER must conform to this type of verification.

• When installing an OER it must have its functioning in symmetry in the execution therefore all OER must pass this **installation test**.

• The **Security test** of an OER is to give it dependability, and one must use the various roles, profiles, permissions, to test the different ways in which OER can be used.

It is noted that a considerable part of the tests can be applied to all OER. It is believed that without the application of these, OER will tend to be fragile when executed, put into production or even recreated from other OER. These are: Operational Testing, Regression Testing, Functional Testing, Interface Testing, Performance Testing, Load Testing, User Acceptance Testing, Volume Testing, Configuration Tests, Installation Tests, and Security Testing.

In what concerns the OER of simulations / animations / games and educational software, all types of tests cited in this study could be applied, because they are directly linked to the software production process and the tests referred to them come directly from Engineering of Software.

Videos, images and texts / hypertexts are OER that have low adherence to some types of test mentioned in the list, since these OER tend to be static and do not have sophisticated interactive elements, being closer to media products than software, in the strict sense of the term.

**Framework Architecture**

Figure 5 below shows diagrammatically the elements involved in the creation of a computational framework, called OERTrust, to support the categorization of quality criteria in OER tests. It should be emphasized that, although the framework is initially aimed at the context of tests and quality of small OER, its application can be extended to large OER.
Figure 5. Computational framework to support the categorization of quality criteria in small OER tests.

Figure 5 shows the architecture elements of the proposed Framework. Given any type of small OER, it should be classified according to its main characteristic: Software, Image, Audio, Questionnaires, Simulation, Animation, Games, Software, Theoretical Text / Hypertext or Video. The next step is to check which set of tests are to be applied on the OER. Given the set, the tests must be performed – clearly, not all tests are mandatory, but to each test it will be assigned different weight according to the type of OER. Each of these weights will be a coefficient of a Fuzzy Logic membership function (RUSSEL and NORVIG, 2003), which will impact the calculation of the OER $i_T$ index of trustability. If the OER already has a previously designed, it has to be recalculated, passing by the flow of the sub-set of tests that at another time had its guaranteed measurement, performing tests only on the changes made in the new versioning.
In summary, the activities of the OERTrust framework for validation of tests in OER are:

- Identify the type of OER to be validated whether large or small.
- Check that the OER already has the reliability index.
- Check which test group should be validated or revalidated.
- Define the weights of each test for each set of tests.
- Calculate the membership function
- Generate the reliability index $i_T$.

Such index $i_T$ is calculated according to the following mathematical modelling, explained in equations 1 and 2.

$$i_T^n = \sum_{a=1}^{m} w_a * R_a, \quad \text{for } n = 1$$

_Equation 1_

$$i_T^n = \Delta_n i_T^{n-1} + \sum_{a=1}^{m} w_a * R_a, \quad \text{for } n > 1$$

_Equation 2_

Where:

- $i_T^n = \text{Reliability index of new version } n$
- $i_T^{n-1} = \text{Reliability index of the previous version } n - 1$
- $\Delta_n = \text{differences between versions } n \text{ and } n - 1$
- $w_a = \text{test weight for the type of OER}$
- $R_a = \text{test result on OER}$
- $m = \text{number of tests of the set}$

It should be noted that in the case of OER that do not have any previously assigned $i_T$ coefficient, the calculation is done by eq. 1. Otherwise, it is done by eq. 2.

**Conclusions**

The main purpose of this work is to describe a framework to support OER quality assurance and validation processes through a set of Software Engineering-based tests, taking into account OER specific features, like versioning and remixing, and the nature of some OER, that must be considered more a media product than a piece of software. The framework is being designed under a three-dimensional approach, namely following the pedagogical, content and technical dimensions.

Future works include the implementation of a functional prototype of a controlled subset of this framework functionalities and its subsequent evaluation by instructional designers and professors. Another study could be based on the availability of the implementation of this framework in several virtual...
repositories of OER, according to the need of the repository. Such implementation also follows the openness principles so that other developers can adapt the framework activities, given some possible emergence of new types of OER, not related in this article. Another important factor in this implementation is the flexibility of weights to be assigned to the tests – used as parameters on Fuzzy Logic’s membership function - according to the type of OER being tested. Also it is being designed in an open, extensible way, the refinement of testing mechanisms involving instructional designers as well as faculty to help validate educational resource.

The pedagogical dimension is, until this moment, the most challenging part of the framework since, unlike repository-based LO, not every OER has proper documentation or metadata. Intelligent mechanisms of metadata automatic extraction and semantic-based inference – like in Souza et al. (2008) – or others approaches based on Machine Learning are needed to bridge this gap.

References
Albuquerque, A. G. P.; Silveira, I. F. (2013). Reuso de recursos educacionais abertos: uma proposta de mecanismo de recomendação o aplicado a textos abertos colaborativos [Reuse of open educational resources: a proposal of a recommendation mechanism applied to open collaborative texts]. In: Proceedings of World Congress on Systems Engineering and Information Technology.. p. 73-77.
Matté, M. A. (2011). Testes de software: uma abordagem da atividade de teste software em metodologias ágeis aplicando a técnica behavior driven development em um estudo experimental. [Software testing: an approach to software testing activity in


