

Article

Examining the Usefulness of Quality Scores for Generating Learning Object Recommendations in Repositories of Open Educational Resources

Aldo Gordillo ^{1,*} , Daniel López-Fernández ¹  and Katrien Verbert ²

¹ Departamento de Sistemas Informáticos, Universidad Politécnica de Madrid, 28031 Madrid, Spain; daniel.lopez@upm.es

² Department of Computer Science, KU Leuven, B-3001 Leuven, Belgium; katrien.verbert@kuleuven.be

* Correspondence: a.gordillo@upm.es

Received: 9 June 2020; Accepted: 2 July 2020; Published: 4 July 2020



Abstract: Open educational resources (OER) can contribute to democratize education by providing effective learning experiences with lower costs. Nevertheless, the massive amount of resources currently available in OER repositories makes it difficult for teachers and learners to find relevant and high-quality content, which is hindering OER use and adoption. Recommender systems that use data related to the pedagogical quality of the OER can help to overcome this problem. However, studies analyzing the usefulness of these data for generating OER recommendations are very limited and inconclusive. This article examines the usefulness of using pedagogical quality scores for generating OER recommendations in OER repositories by means of a user study that compares the following four different recommendation approaches: a traditional content-based recommendation technique, a quality-based non-personalized recommendation technique, a hybrid approach that combines the two previous techniques, and random recommendations. This user study involved 53 participants and 400 OER whose quality was evaluated by reviewers using the Learning Object Review Instrument (LORI). The main finding of this study is that pedagogical quality scores can enhance traditional content-based OER recommender systems by allowing them to recommend OER with more quality without detriment to relevance.

Keywords: OER; open educational resources; learning objects; recommender systems; quality; digital repositories; TEL

1. Introduction

Open educational resources (OER) have a huge and irrefutable potential to provide universal and equal access to education at all levels, contributing to education's democratization [1]. This potential has been noticed by the European Union, who recognized that stimulating supply and demand for high-quality OER is essential for modernizing education and that these resources contribute to broadening access to education, as well as to alleviating costs for educational institutions and students, especially among disadvantaged groups [2]. The first definition of the OER term was proposed at a UNESCO conference in 2002 which defined OER as “the open provision of educational resources, enabled by information and communication technologies, for consultation, use, and adaptation by a community of users for non-commercial purposes” [3]. A definition more often used nowadays states that OER are “digitized materials offered freely and openly for educators, students, and self-learners to use and reuse for teaching, learning, and research” [4]. The educational benefits of OER most frequently recognized include the encouragement of improvement and localization of content, the drastic reduction of costs without detriment to instructional effectiveness, and the capacity to offer equal access to knowledge for all.

The current OER panorama reveals that there is an enormous and growing amount of digital learning resources available for learners and teachers through OER repositories. Repositories such as Wikimedia Commons [5], Europeana [6], OER Commons [7], MERLOT [8], ODS [9], or the LRE [10] are offering thousands or even millions of OER. For example, Wikimedia Commons offers more than 61 million freely usable media files, whereas Europeana makes available nearly 59 million digitized materials from European museums, galleries, and multimedia archives.

This massive amount of resources available in OER repositories makes it difficult for users to find relevant and high-quality content when searching using these systems. A recent systematic literature review [11] indicated that discoverability of quality open learning resources was one of the major barriers hindering OER adoption and impact. According to the results of the OER Data Report 2013–2015 [12], finding resources of sufficiently high quality and related to a specific subject area are challenges often experienced by both educators and learners when searching for OER. Other studies have also pointed out difficulties in locating relevant and quality educational resources. Najjar et al. [13] investigated usability problems of search tools for repositories of educational resources and concluded that these tools were not easy to use. John et al. [14] explored the search services offered by eight of the most popular OER repositories and pointed out gaps, as well as possible solutions. Another study [15] showed that teachers could require more than an hour to find a suitable learning resource, and that this search process was perceived in many occasions to be time consuming and occasionally frustrating.

An aspect that deserves special attention in OER repositories is the pedagogical quality of the learning resources, as many of these repositories follow a production model in which OER are created by communities of volunteers and lack an effective quality control mechanism. These are the main reasons why quality assurance is still often pointed out as an open issue for OER [16–19]. There are two main reasons why quality assurance is especially important for OER repositories. On the one hand, teachers need to have some guarantee of quality before including OER in the syllabus. Two surveys [12,20] have provided evidence that lack of quality control has been a problem for teachers when using OER authored by others and that user ratings were important for teachers when searching for OER. On the other hand, if no quality control mechanism is used in an OER repository, learners in self-directed or student-centered educational settings in which they are expected to choose their own materials could be at risk of being misinformed by wrong, inaccurate, or outdated content, as well as wasting time with poor or inappropriate instructional designs [21]. In order to assure the pedagogical quality of their learning resources, many OER repositories have developed their own quality control mechanisms. Evidence of this fact has been provided by a survey of 59 popular repositories of learning resources [22], which found that 27 (46%) of them followed a quality control policy, and 23 (39%) had a resource assessment or rating policy. Moreover, with the aim of addressing the clear need of OER repositories to assess the quality of their learning resources in a systematic way, plenty of evaluation models have been elaborated such as the Learning Object Review Instrument (LORI) [23], UNE 71362 [24], WBLT-S and WBLT-T [25,26], COdA [27], LOEM [28], and MECOA [29]. Furthermore, some software systems have been developed in order to facilitate the quality evaluation of digital learning resources according to these evaluation models [29–33].

In order to facilitate users to find OER of their interest and of enough quality, OER repositories can apply different measures such as developing easy-to-use search tools [13], adopting ranking metrics [34–37], and using recommender systems [38–40]. Recommender systems are software tools that provide suggestions for items likely to be of use to a user [41,42]. In the context of technology enhanced learning (TEL), these systems are mainly used to help teachers and learners find suitable digital learning resources and, to a lesser extent, to suggest sequences of resources, didactic activities, full courses, and peer learners [39,40]. Therefore, in OER repositories, the items suggested by recommender systems are generally learning resources, educational activities, or courses relevant to the users' interests. In this regard, it should be noted that, in the literature, the term "learning objects" is often used to refer to digital learning resources, concretely to refer to any "reusable digital resources tagged with metadata that are self-contained and that can be used for education" [43]. With respect to

the use of recommender systems by OER repositories, it should be noted that, although the automatic recommendation of digital learning resources has been identified as a feature that any OER repository should have in order to meet the current needs of the educational community [44], only a very small percentage have implemented a recommender system to provide this feature [45].

Different types of recommender systems exist according to the recommendation technique used for generating the recommendations, including content-based, collaborative filtering, demographic, knowledge-based, social (or community-based), and hybrid recommender systems. The major types are: content-based recommender systems, which recommend items that are similar to the ones that the user found relevant or liked in the past; collaborative filtering, where item recommendations are generated by the system using only information about ratings or usage from other users; and hybrid recommender systems, which are those that combine two or more techniques in order to generate the recommendations. The context, understood as “any information that can be used to characterize the situation of a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves” [46], is an important source of information from which recommender systems can take advantage. Recommender systems with a recommendation process that makes use of context are called “context-aware recommender systems” [47,48]. These systems are able to suggest items to users by using contextual information such as activity, time of day, or location. In TEL environments, it is also quite common that recommender systems consider learning resources that are relevant to the user [48], for instance, with the goal of recommending resources similar to those that the user is currently consuming, viewing, or searching. Other information that could be useful for TEL recommender systems is information related to the pedagogical quality of the learning resources [49].

There is a wide range of studies related to the use of recommender systems for assisting users with the search and discovery of OER based on their preferences and needs. Clough et al. [50] discussed the implementation of non-personalized suggestions of digital learning resources from Europeana using both content-based and collaborative recommendation techniques. In a more recent work, Otegi et al. [51] examined the use of personalized PageRank to recommend Europeana resources. Another study [52], also related to Europeana, described and evaluated a hybrid recommender system capable of generating recommendations of digital learning resources retrieved from Europeana. Gordillo, Barra, and Quemada [53] implemented two different recommender systems for two OER repositories based on the same hybrid recommendation model, which combined three recommendation techniques (content-based, demographic, and context-aware) and the use of quality and popularity scores. The authors of [54] introduced a recommender system for suggesting learning objects to teachers by exploiting their ICT competence profiles elicited from their relevance feedback data and evaluated this system using datasets retrieved from three OER repositories. Lemos et al. [55] evaluated a traditional collaborative filtering technique and a new recommendation technique based on clustering and collaborative filtering in a dataset extracted from the MERLOT repository. In [56], a semantic recommender system capable of assisting learners in searching learning objects was proposed and discussed. Another recommender system to assist learners in searching for learning objects in repositories of educational resources was presented by [57]. More recently, an implementation in this same system of a learning object group recommendation approach was described [58]. In another related work, Ruiz-Iniesta et al. [59] described the application of a knowledge-based OER recommendation strategy to enhance the users’ search experience in OER repositories. Cechinel et al. [60] evaluated OER recommendations generated according to distinct collaborative filtering recommendation algorithms using OER from the MERLOT repository. A theoretical model for generating proactive context-aware recommendations in an OER repository by using community-based and collaborative filtering techniques was proposed by Gallego et al. [61,62]. In a later work, Gallego et al. [63] presented another model that could be used to generate proactive context-aware OER recommendations in learning object authoring tools.

Despite the number of research works that have been carried out on the use of recommender systems to find and discover relevant OER, and the key role that OER pedagogical quality plays in OER repositories, research analyzing the use of information related to pedagogical quality for generating OER recommendations is very scarce. Among all related works described, only two ([53,57]) have reported an evaluation of a learning object recommender system that took into account information of this kind, and neither of them investigated the specific added value of incorporating this information into the recommendation process. Another issue, which has been raised in the literature on TEL recommender systems, has been the high number of studies that lacked evaluations involving human users in spite of the need to evaluate these systems from the user perspective [38]. Regarding this matter, it is also important to note that most evaluations of these systems, so far, have focused solely on algorithm accuracy [40], despite the fact that other factors beyond this one need to be considered when evaluating TEL recommender systems since high recommender accuracy does not always correlate with user satisfaction or perceived usefulness [64]. In the particular case of recommender systems for OER repositories, it is essential to take into account the pedagogical quality of the recommended OER in order to evaluate the actual value of the recommendations. In this regard, only two studies ([53,55]) have been found that reported an evaluation of OER recommendations considering pedagogical quality. Nevertheless, neither of them analyzed the influence of using measures of OER quality on this variable.

This article examines the usefulness of using pedagogical quality scores for generating learning object recommendations in OER repositories through a user study involving 53 participants and 400 OER. This user study compares the following four different recommendation approaches: a traditional content-based recommendation technique; a quality-based non-personalized recommendation technique; a hybrid approach that combines the two former techniques; and a random recommendation method, which was used as a baseline for comparison purposes. These four recommendation approaches were compared in terms of both relevance and pedagogical quality of the recommended OER in order to answer the following two research questions:

1. Are learning object quality scores useful to generate OER recommendations?
2. Can existing OER recommender systems be enhanced by using learning object quality scores?

The results of this article provide strong evidence that learning object pedagogical quality scores, although they have limited utility by themselves, can be successfully used to enhance traditional content-based OER recommender systems by enabling the recommendation of OER with higher quality without detriment to relevance.

The remainder of the article is organized as follows: the user study is explained in Section 2, including details about the sample, the recommendation approaches, the dataset, and the quality scores of the OER that compose it; Section 3 presents the results of the user study, which are then discussed in Section 4; finally, Section 5 concludes the article by summarizing the main findings and giving an outlook of future work.

2. Materials and Methods

2.1. Sample

A total of 53 potential users of OER repositories participated in the user study. Of these participants, 42 (79%) were teachers and 11 (21%) were learners, and of these participants, 33 (62%) were males and 20 (38%) were females. All participants were between 21 and 70 years old, the mean age being 39 with a standard deviation of 13.4. When asked about their experience using OER repositories, 10 (18.9%) participants declared to have no experience, 14 (26.4%) declared to have low experience, 15 (28.3%) declared medium experience, and the remaining 14 (26.4%) declared high or very high experience.

2.2. User Study

This user study was conducted online, and therefore participants could complete it by themselves using their own computers through a web browser. The only guidance received by the participants

was a text document with the necessary instructions to participate in the user study. All participants performed the following steps:

1. Participants accessed an OER repository and create a user account. Concretely, participants accessed ViSH [43], an OER repository publicly available at <http://vishub.org>, which is enriched with additional features such as authoring tools, an audience response system, recommendations, and a social network.
2. Participants accessed an online evaluation tool provided through the ViSH web portal. This tool was developed specifically for this user study.
3. Once participants had accessed the online evaluation tool, they were asked to provide some basic information, including demographic data (age, gender, language, and occupation) and their previous experience using OER repositories.
4. Participants chose a topic of their interest (e.g., engineering, biology, or history) among different options. It should be clarified that this topic was used in the user study only to select an OER in the next step (it was not taken into account by any of the recommendation approaches).
5. In this step, participants were told to put themselves in the situation that they are searching for learning resources in an OER repository (such as ViSH) and that, finally, they find one in which they are interested and, after clicking on a link, they navigate to a new page where they can see it. An OER of ViSH tagged with the topic chosen by the participants in the previous step was randomly selected and shown to the participants in order to act as the particular OER that each participant found in their fictitious search. Below this OER, a list of recommendations with 20 randomly sorted OER were presented to each participant. Finally, participants were required to rate, given the described situation, the relevance of each of the recommended OER using a five-point Likert scale from 5 (very relevant to me) to 1 (not at all relevant to me).

The lists of OER recommendations shown to the users in the last step of the user study were built by concatenating four distinct lists of five OER, each of them generated by a system using a different recommendation approach. Each of these four recommendation approaches is described in detail in the Section 2.4. Therefore, in order to build each list shown to the participants, a total of 20 OER recommendations were generated: five by a traditional content-based recommender system, five according to a quality-based non-personalized approach, five by a hybrid recommender system that combines content-based and quality-based recommendation techniques, and five chosen at random. It should be highlighted that after concatenating the different lists, all OER were shuffled and hence it was not possible for participants to know which OER was recommended by which system. In fact, participants did not even know that the recommendations had been generated by different systems. Furthermore, by using this research design, position bias was prevented. In the case of overlap (i.e., when the same OER was recommended by two or more systems), duplicates were removed, and therefore no OER were shown to a participant more than once. Therefore, in these cases, it should be taken into account that, although 20 OER recommendations were generated, the OER list shown to the participants was actually comprised of less than 20 recommendations.

2.3. Dataset

In order to perform a reliable comparison, all the systems used to generate recommendations in this user study used the same dataset. This dataset was comprised of 400 OER authored by individuals and published on the ViSH OER repository [43]. All of these OER were learning objects, since they were self-contained digital learning resources, they were tagged with metadata according to the IEEE LOM [65] standard, and they were also compliant with the SCORM [66] standard.

Besides metadata specified according to the IEEE LOM standard, each OER of the dataset had an associated quality score, which represented a measure of its pedagogical quality. This quality score was generated based on evaluations conducted by experienced reviewers using the LORI (Learning Object Review Instrument) [23] evaluation model, which defined the following nine items: content

quality, learning goal alignment, feedback and adaptation, motivation, presentation design, interaction usability, accessibility, reusability, and standards compliance. In order to get a reliable quality measure, each OER was assessed by at least two different reviewers using LORI. All reviewers conducted the OER quality evaluations by means of the LOEP [33] tool. LORI is likely the most widely used instrument for evaluating the quality of digital learning resources, and it has been evaluated by some studies which have concluded that it can be reliably used to assess learning object quality [21,67]. When a learning object is assessed with LORI, each reviewer rates each LORI item using a scale from 1 to 5. In this study, overall quality scores were calculated for each OER of the dataset as the mean of the scores of all LORI items and averaging reviewers scores. Nevertheless, it should be taken into account that the ViSH repository transforms these scores on a 1–5 scale to scores on a 0–10 scale, which are the ones that are finally available for the recommender systems.

2.4. Recommendation Approaches

2.4.1. Traditional Content-Based Recommendation (CB)

Content-based recommender systems learn to recommend items that are similar to the ones the user found relevant or liked in the past. In these systems, recommendations are generated based on item features, user actions or user ratings. In OER repositories, content-based recommender systems can be mainly used to recommend OER similar to the ones the user has visited, downloaded, shared, bookmarked, or rated positively, as well as OER similar to the ones the user is currently viewing. In this context, content-based recommender systems can generate the recommendations by using OER metadata, which usually include data such as their title, language, resource type, and keywords. All this information can be defined with IEEE LOM [65], the most used metadata standard for learning objects.

A traditional content-based recommender system was implemented for this user study. This recommender system generates OER recommendations based on the similarity between the OER that is being viewed by the user and those OER that can be recommended (hereafter referred to as candidate OER) taking into account three metadata fields: title, language and keywords. The overall similarity score for each candidate OER on a 0–1 scale was calculated by this recommender system according to the following equation:

$$S_{CB}(O_x, O_y) = 0.3 \times \text{Sim}_T(O_{x_t}, O_{y_t}) + 0.4 \times \text{Sim}_L(O_{x_l}, O_{y_l}) + 0.3 \times \text{Sim}_K(O_{x_k}, O_{y_k}) \quad (1)$$

where O_x is the OER the user is currently viewing and O_y is the candidate OER; O_{x_t} and O_{y_t} are, respectively, the titles of O_x and O_y ; O_{x_l} and O_{y_l} are, respectively, the languages of O_x and O_y ; O_{x_k} and O_{y_k} are, respectively, the keywords of O_x and O_y ; $\text{Sim}_T(O_{x_t}, O_{y_t})$ is the similarity score between the textual fields O_{x_t} and O_{y_t} on a 0–1 scale calculated by using the TF-IDF function [68] and the cosine similarity metric according to Equation (2); $\text{Sim}_L(O_{x_l}, O_{y_l})$ is 1 if O_x and O_y have the same language and 0 otherwise; and $\text{Sim}_K(O_{x_k}, O_{y_k})$ is the similarity score between the lists of keywords O_{x_k} and O_{y_k} on a 0–1 scale as calculated by Equation (3).

$$\text{Sim}_T(T_x, T_y) = \begin{cases} 0, & \|N\| = 0 \\ \frac{\sum_{i=1}^N \text{TFIDF}(w_i, T_x) \times \text{TFIDF}(w_i, T_y)}{\sqrt{\sum_{i=1}^N \text{TFIDF}(w_i, T_x)^2} \times \sqrt{\sum_{i=1}^N \text{TFIDF}(w_i, T_y)^2}}, & \|N\| > 0 \end{cases} \quad (2)$$

where N is the quantity of distinct words in the texts T_x and T_y , and w_i is the i -th of these words.

$$\text{Sim}_K(L_x, L_y) = \begin{cases} 0, & |N_{L_x} + N_{L_y}| = 0 \\ \frac{2 \times N_C}{N_{L_x} + N_{L_y}}, & |N_{L_x} + N_{L_y}| > 0 \end{cases} \quad (3)$$

where N_{L_x} and N_{L_y} are, respectively, the number of keywords of the lists L_x and L_y , and N_C is the quantity of common keywords between L_x and L_y .

In order to generate a list of N recommended OER, this recommender system, first, calculates a similarity score for each OER in the dataset (i.e., for each candidate OER) using Equation (1), then, sorts all candidate OER according to their similarity score, and finally, generates the recommendation as a sorted list of the N OER with the highest similarity score. This recommendation approach was included in the user study for obtaining a baseline that could later be used for comparison, in order to examine the specific added value of incorporating pedagogical quality scores into a common content-based recommender system.

2.4.2. Quality-Based Non-Personalized Recommendation (Q)

In this approach, OER recommendations are generated exclusively based on the pedagogical quality of the OER measured with the LORI evaluation model, as explained in Section 2.3. Therefore, these recommendations are non-personalized since neither the profile of the users nor their actions are taken into account for their generation. Thus, the OER that users are viewing at a certain moment are also not taken into account. A system capable of suggesting OER to the users according to this quality-based non-personalized recommendation approach was implemented for the study. This system generates lists of N recommended OER by sorting by quality scores all OER in the dataset and randomly selecting N OER among the ones in the top quartile (i.e., among the 25% of OER with highest quality in the dataset). One option for this approach could have been to directly suggest the top N OER (i.e., always recommend the OER with more quality in the dataset). However, we decided to introduce some randomness in the recommendation process in order to evaluate a more realistic approach.

There were two main reasons for including this non-personalized recommendation approach in the user study. On the one hand, this approach gathers evidence on the usefulness of this kind of recommendations, which can be useful to determine the benefits for OER repositories of providing users with suggestions such as “the top ten selections of OER” or “the five best OER of the year”. Although it is recognized that non-personalized recommendations can be useful in some scenarios, these types of recommendations are not often addressed by recommender systems research [41]. In fact, no study has ever examined the usefulness of non-personalized OER recommendations based exclusively on pedagogical quality indicators. On the other hand, this approach provides another baseline to be used for comparison purposes.

2.4.3. Hybrid Approach: Content-Based and Quality-Based Recommendation (CB + Q)

A hybrid recommender system was developed with the aim of examining the usefulness of learning object quality scores for generating OER recommendations and for enhancing existing content-based OER recommender systems. This system combines the traditional content-based recommendation technique, described in Section 2.4.1, and the quality-based non-personalized recommendation technique, described in Section 2.4.2. The hybrid recommender system calculates the overall score on a 0–1 scale for each candidate OER as follows:

$$S_{CB+Q}(O_x, O_y) = 0.8 \times S_{CB}(O_x, O_y) + 0.2 \times S_Q(O_y) \quad (4)$$

where O_x is the OER the user is currently viewing and O_y is the candidate OER, $S_{CB}(O_x, O_y)$ is a score calculated according to Equation (1), and $S_Q(O_y)$ is the quality score of O_y on a 0–1 scale.

The recommendation process followed by the hybrid recommender system is the same as the one followed by the traditional content-based recommender system, except for the way in which scores are calculated. Therefore, the hybrid recommender system, first, calculates a score for each candidate OER using Equation (4), then, sorts all candidate OER according to the calculated scores, and finally generates the recommendation as a sorted list of the N OER with the highest calculated scores.

This recommendation approach allowed to analyze the specific added value of incorporating learning object quality scores into the OER recommendation process by comparing its performance against the one obtained by the traditional content-based approach. This comparison also made it possible to determine if OER quality scores are useful for enhancing existing recommender systems for OER repositories.

2.4.4. Random Recommendation (R)

The last approach for providing OER recommendations involved in this user study simply consisted of choosing, at random, OER from the dataset. In order to enable this approach, a system that generated recommendations in the form of lists of N OER randomly retrieved from the dataset was implemented. The random recommendations were included just to get a baseline for comparison purposes. Thereby, the usefulness of the OER recommendations generated according to the other three recommendation approaches can be compared with a reference.

2.5. Usefulness Evaluation of the OER Recommendations

Results related to two different factors were obtained in the user study for each of the recommendation approaches examined. On the one hand, the relevance of each OER recommendation was rated by the participants using a 5-point Likert scale. These ratings indicated, for each recommendation approach, how relevant the recommended OER were for the participants. On the other hand, the quality score (i.e., the pedagogical quality according to the results obtained from the LORI evaluations conducted by the reviewers) of each of the recommended OER was recorded by the evaluation tool used by the participants in the user study. These data indicated how good, in terms of pedagogical quality, the recommended OER were. Thereby, measures of relevance and pedagogical quality were obtained for each OER recommendation and for each recommendation approach. These two factors are the most important ones in order to evaluate the usefulness of OER recommendations. A proof of this fact can be found in the OER Data Report 2013–2015 [12], whose results showed that the relevance and user ratings were important factors for both teachers and learners when searching for OER.

In order to obtain an overall measure of the relevance of the OER recommendations generated by the different recommendation approaches based on the results of the user study, the normalized R-Score metric was used. This metric, which is a variation of the R-Score metric [69], calculates the utility score on a 0–1 scale of a list of N_L recommendations presented to a user, U , according to the following equation:

$$R_U = \frac{\sum_{i=1}^{N_L} \frac{\text{MAX}(S_i - d, 0)}{2^{\frac{i-1}{\alpha-1}}}}{\sum_{i=1}^{N_L} \frac{S_{\text{MAX}} - d}{2^{\frac{i-1}{\alpha-1}}}} \quad (5)$$

where S_i is the rating assigned by U to the item in the i -th position of the list, $S_{\text{MAX}} \geq S_i$ is the maximum numerical rating that a user can assign to an item, $d < S_{\text{MAX}}$ is the value of an adaptable threshold that determines the minimum numerical rating that a user should assign to an item in order for that item to be considered useful, and $\alpha > 1$ is a half-life parameter that regulates the exponential decrease of the recommendations utility according to their position on the list.

In order to measure the utility of a set of lists of recommendations presented to a set of N_U users, the normalized R-Score metric provides an overall score calculated as follows:

$$R = \frac{1}{N_U} \times \sum_{i=1}^{N_U} R_{U_i} \quad (6)$$

where U_i is the i -th user of the set of N_U users and R_{U_i} is the utility score of the list of recommendations presented to U_i calculated according to Equation (5).

The normalized R-Score metric, like the R-Score metric, presupposes that the utility of a list of recommendations is determined by the sum of the utilities of each individual recommendation. It also presupposes that the utility of an individual recommendation decreases exponentially as its position on the recommendation list decreases. The maximum utility score of a set of lists of recommendations calculated by means of the normalized R-Score metric is 1 and is reached when all users assign the maximum rating S_{MAX} to all the items recommended. On the contrary, the minimum utility score of a set of lists of recommendations calculated through the normalized R-Score metric is 0 and is obtained when all users rate all the recommended items with ratings that are equal or lower than the threshold value d . In this user study, the parameter S_{MAX} was set to 5 since participants rated the relevance of the OER using a 5-point scale, the threshold value d was set to 1, and the half-life parameter α was set to 3.

In addition to calculating overall measures of relevance, the normalized R-Score metric was used, in this user study, for obtaining an overall measure of pedagogical quality of the OER recommended for each of the different recommendation approaches. These overall quality measures were calculated in the same way as the overall relevance measures, but with the following two changes in Equation (5): S_i was the quality score of the OER in the i -th position of the list (calculated according to the LORI evaluations) instead of the participant rating, and S_{MAX} was set to 10 since it is the maximum value the quality score of an OER could be. As for calculating the relevance, the threshold value d was set to 1 and the half-life parameter α was set to 3. When calculating overall quality measures with the normalized R-Score metric, on the one hand, the maximum utility score of a set of lists of recommendations (i.e., a value of 1) is reached when all recommended OER have a quality score of 10, which only occurs if all reviewers rate all those OER with the maximum score in all LORI criteria. On the other hand, the minimum utility score of a set of lists of recommendations (i.e., a value of zero) is obtained when all recommended OER have a quality score lower or equal than 1, the value of the parameter d , which in this case acts as a quality threshold instead of as a participant rating threshold.

In summary, the following data were obtained from the results of the user study by using the normalized R-Score metric:

- A relevance score for each recommendation approach and participant of the user study on a 0–1 scale indicating how relevant for the participant were the recommendations generated by the approach.
- An overall relevance score for each recommendation approach on a 0–1 scale indicating, on average, how relevant were the recommendations generated by the approach for all participants.
- A quality score for each recommendation approach and participant of the user study on a 0–1 scale indicating how good in terms of pedagogical quality were the OER recommended to the participant according to the approach.
- An overall quality score for each recommendation approach on a 0–1 scale indicating, on average, how good in terms of pedagogical quality were the OER recommended by the approach to all the participants of the user study.

Thereby, the usefulness of the OER recommendations generated by each of the four recommendation approaches was evaluated in terms of both relevance and pedagogical quality. This evaluation compared these recommendation approaches based on the relevance of their recommendations and the quality of the recommended OER.

3. Results

Table 1 shows the mean (M) and standard deviation (SD) of the overall relevance and quality scores of the OER recommendations calculated by means of the normalized R-Score metric based on the results of the user study for each of the four recommendation approaches analyzed: the hybrid approach combining content-based and quality-based recommendation techniques (labeled as CB + Q in the tables), the traditional content-based recommendation technique (labeled as CB), the quality-based

non-personalized recommendation technique (labeled as Q), and the random recommendations (labeled as R).

Table 1. Relevance and quality scores of the open educational resources (OER) recommendations.

Factor	CB + Q		CB		Q		R		Kruskal-Wallis H Test	
	M	SD	M	SD	M	SD	M	SD	H	p-Value
Relevance	0.64	0.25	0.60	0.29	0.25	0.21	0.17	0.21	91.6	<0.001
Quality	0.77	0.10	0.60	0.17	0.78	0.04	0.53	0.10	115.1	<0.001

The Kruskal-Wallis and Mann-Whitney U tests were used to determine statistical significance. Firstly, a Kruskal-Wallis H test was used to determine if there were statistically significant differences between the overall relevance and quality scores of the different recommendation approaches. The results of this test can be seen in Table 1. Secondly, a series of Mann-Whitney U tests were conducted to compare the overall relevance and quality scores between each pair of recommendation approaches (i.e., CB + Q vs. CB, CB + Q vs. Q, CB + Q vs. R, CB vs. Q, CB vs. R, and Q vs. R). Table 2 shows the results of these tests.

Table 2. Results of Mann-Whitney U tests comparing relevance and quality scores.

Approaches	Relevance			Quality		
	U	Effect Size (r)	p-Value	U	Effect Size (r)	p-Value
CB + Q vs. CB	1318	0.02	0.586	414	0.60	<0.001
CB + Q vs. Q	351	0.64	<0.001	1148	0.12	0.104
CB + Q vs. R	246	0.70	<0.001	122	0.78	<0.001
CB vs. Q	487	0.55	<0.001	379	0.62	<0.001
CB vs. R	329	0.65	<0.001	1053	0.19	0.026
Q vs. R	1054	0.19	0.026	31	>1.0	<0.001

As shown in Table 1, the OER recommendations generated by the hybrid approach were the most relevant for the participants (M = 0.64, SD = 0.25), followed very closely by those generated by the traditional content-based recommendation technique (M = 0.60, SD = 0.29). In fact, this difference was not found to be statistically significant. The relevance of the OER recommendations generated according to the other recommendation approaches were found to be much less relevant for the participants, with those generated based on OER pedagogical quality being more relevant (M = 0.25, SD = 0.21) than the random recommendations (M = 0.17, SD = 0.21). With the exception of the hybrid and content-based recommender approaches, the differences between the overall relevance scores for the other pairs of recommendation approaches were found to be statistically significant. Large effect sizes (r > 0.5) were found in all comparisons except between Q and R, where the effect size was small to medium, and between CB + Q and CB, where the effect size was insignificant and, as mentioned before, the difference between overall relevance scores was non-statistically significant.

Regarding overall quality scores, the recommendation approaches that recommended OER with highest quality were the quality-based approach (M = 0.78, SD = 0.04) and the hybrid approach (M = 0.77, SD = 0.10). The difference between the quality scores of these two approaches was small and non-statistically significant. Both of these approaches outperformed, in terms of overall quality scores, the traditional content-based recommendation technique (M = 0.60, SD = 0.17), as well as the random recommendations (M = 0.53, SD = 0.10). In all these cases (CB + Q vs. CB, CB + Q vs. R, Q vs. CB, and Q vs. R), the difference between overall quality scores was found to be statistically significant with a large effect size. Finally, the traditional content-based recommendation technique outperformed randomly generated recommendations in terms of overall quality scores, being the difference between these scores statistically significant with a small to medium effect size.

The results obtained from the user study show that the recommendation approach that generated the more useful OER recommendations in terms of relevance and quality was the hybrid recommendation approach that combines content-based and quality-based techniques. In terms of relevance, the better recommendation approaches were the hybrid approach and the traditional content-based approach, whereas, in terms of quality, the better approaches were the quality-based approach and the hybrid one. If both factors (relevance and quality) are taken into account for measuring the usefulness of the OER recommendations, the better approach was the hybrid one, because it was capable of generating the recommendations with highest relevance (achieving relevance scores similar to those achieved by the traditional content-based recommender system) and, at the same time, recommending OER with high pedagogical quality, achieving quality scores similar to those achieved by the quality-based approach, which was the one that performed best in this factor. Therefore, the hybrid recommendation approach was capable of recommending OER as relevant as the ones recommended by the traditional content-based approach but with higher quality, as well as recommending OER of similar quality as the ones recommended by the quality-based approach but more relevant to the participants.

Regarding the quality-based approach, the results of this user study indicated that it was useful to generate high-quality OER. In fact, as expected, this approach was the one that achieved recommending OER with the most quality. However, in view of the obtained results, this approach clearly failed to suggest OER relevant to the users' interests. In this regard, it should be pointed out that, although the results achieved by the quality-based OER recommendations in terms of relevance were far from those achieved by the personalized recommendations, these quality-based OER recommendations clearly outperformed random recommendations, in terms of pedagogical quality, and also in terms of relevance.

4. Discussion

The results of the conducted user study show that learning object pedagogical quality scores can be used to improve traditional content-based OER recommender systems, allowing these systems to recommend OER with more quality and equally relevant for the users. Therefore, this study proves that OER repositories can make use of OER quality measures in order to help their users to find OER of sufficiently high quality and related to a specific topic, which is a challenge often faced by both teachers and learners when searching for OER [12]. The results of this research show that the hybrid recommendation approach statistically significantly outperformed the traditional content-based recommendation approach in terms of quality, the quality-based recommendation approach in terms of relevance, and the random recommendations in terms of both quality and relevance. Therefore, these results are consistent with those of [53], which also compared the relevance (by means of a user study) and quality (by means of A/B testing) of random OER recommendations and OER recommendations generated by a hybrid recommender system using quality scores.

This study has also examined the usefulness of non-personalized OER recommendations generated by using solely pedagogical quality scores. On the basis of the obtained results, it can be concluded that these types of recommendations, although they could be useful to provide users of OER repositories with features such as "the 2020 top OER selections", they are not very useful for helping teachers and learners to find educational resources of their specific area of interest. In summary, measures of OER pedagogical quality have limited utility by themselves to suggest OER to users, but they can be successfully used for enhancing existing OER recommender systems in order to recommend OER with higher pedagogical quality.

Most of the results obtained in the user study were expected. Firstly, the hybrid recommender system was expected to outperform the quality-based approach in terms of relevance because this latter approach is not able to generate personalized recommendations, and hence it recommended OER of high quality but that were not found relevant by the users. Secondly, the fact that the hybrid recommender system outperformed the traditional content-based recommender system in terms of

quality was also expected, because this latter system recommended the most similar OER to the one the participant was viewing, without taking into account data related to the pedagogical quality of these OER. An interesting and unanticipated finding of this study was that the hybrid recommender system was able to recommend OER as relevant as the ones suggested by the content-based approach and of similar quality to those recommended by the quality-based approach. This means that this hybrid recommender system was able to improve the pedagogical quality of the OER it recommended without having a cost in terms of relevance. A reasonable hypothesis would have been that this system was capable of enhancing OER quality but with some detriment to relevance. However, the results obtained in this study indicate that this enhancement occurred at no cost. This fact confirms that hybrid OER recommender systems, by combining multiple techniques together, are able to overcome the shortcomings of some of these techniques by using the strengths of the other ones. Lastly, regarding the comparison of the random recommendations against the other approaches, these recommendations were, as expected, the least useful ones in terms of both relevance and quality.

By taking into account that a high recommender accuracy does not always correlate with user satisfaction or perceived usefulness [64], in order to evaluate the real value of OER recommendations it is necessary to not only evaluate their relevance or accuracy, but also the pedagogical quality of the recommended OER. In spite of this fact, most evaluations of TEL recommender systems reported in the literature have focused only on the accuracy of the recommendation algorithms [40]. In this study, the usefulness of the OER recommendations was evaluated in terms of both relevance and quality. Therefore, the recommendations considered to be more useful were those that recommended high-quality OER that were found relevant by the user. If only relevance had been considered for measuring the usefulness of the generated recommendations, recommendations of OER with low pedagogical quality (e.g., for having inaccurate content or an inappropriate level of detail) would have been considered useful in those cases in which participants, after a quick review, considered that the OER could suit their interests. Obviously, these recommendations should not be considered to be useful (or at least not as useful as high-quality OER recommendations) because there is little benefit from providing users with OER with poor quality, and this can even be counterproductive in some cases.

Quality assurance is of major importance for OER repositories since teachers need some guarantee of quality before incorporating OER into their lesson plans and low-quality OER can put learners at risk of being misinformed or wasting time. Therefore, there is little doubt that pedagogical quality is a key feature of OER that these repositories should consider and evaluate. Notwithstanding, this feature is usually ignored by OER recommender systems, even for those recommender systems using content-based techniques, for which it should be easy to incorporate OER quality measures into the recommendation process since they are able to generate recommendations based on item features and ratings. This work proves that OER repositories can benefit from evaluating the pedagogical quality of their OER by incorporating the obtained quality measures into a recommender system. In spite of these benefits, it is clear that there are important barriers preventing the use of quality-based recommendation approaches in OER repositories, with the most critical barrier being the difficulty to effectively collect pedagogical quality scores for all OER. In those OER repositories, such as MERLOT or ViSH, that systematically evaluate the quality of the published OER following a community-based mechanism, evaluations are conducted by volunteer reviewers or users, and hence no quality scores are generated for a significant percentage of OER. Although some measures have been proposed to face this barrier, such as assigning a temporary score to non-evaluated OER [53], automatically evaluating quality using intrinsic features [70], or estimating the OER pedagogical quality based on user interactions [71], this issue remains as an open challenge for OER repositories that requires further research. In the dataset used for this user study, all OER were evaluated by at least two experienced reviewers and, based on these evaluations, quality scores were calculated for all of them. Therefore, in other scenarios where no reliable pedagogical quality data are available for all OER in the dataset, the performance of the quality-based recommendation approaches is expected to be worse. Nonetheless,

in the worst scenario where no data related to pedagogical quality would be available at all, the hybrid approach analyzed in this study would still be able to generate OER recommendations as useful as the ones generated by the traditional content-based approach.

5. Conclusions

This article examines the usefulness of using pedagogical quality scores for generating OER recommendations by means of a user study involving 53 participants and 400 OER, which compared, in terms of both relevance and pedagogical quality, four different recommendation approaches: a hybrid approach combining content-based and quality-based recommendation techniques, a traditional content-based recommendation technique, a recommendation approach that generates OER recommendations based exclusively on the pedagogical quality of the OER, and random recommendations. The results reported in this article show that learning object quality scores can be successfully used to enhance traditional content-based OER recommender systems by enabling these systems to recommend OER that are of higher quality and equally relevant for the users. These results provide evidence of the usefulness to OER recommender systems of including information related to the OER pedagogical quality in the recommendation process. Although previous studies [53,57] had evaluated learning object recommender systems that make use of this type of information, none of them specifically examined its added value. Furthermore, this study also contributes to the TEL recommender systems research by examining the usefulness of non-personalized recommendations of OER that are generated based solely on pedagogical quality scores, which is a valuable contribution since, to the best of our knowledge, no work had addressed this issue before. In conclusion, the results of this study show that the appropriate use of learning object quality scores can contribute to enhance the discoverability of high-quality OER, one of the major barriers hampering the use and uptake of OER worldwide [11].

In addition to providing better OER recommendations, OER repositories can benefit from evaluating OER quality by using the obtained quality scores for enhancing the search and discovery of OER through other types of systems such as search tools or catalog-like applications. For example, similar to the way in which recommender systems can incorporate quality scores to recommend OER that are both relevant and of high quality, search tools can use ranking metrics that combine relevance scores with quality scores in order to sort results not only by relevance, but also taking into account the OER pedagogical quality. Thereby, between those OER that are equally relevant for the user search query, the ones with higher quality can be shown first to the users. An example of this type of ranking metrics is described in [36]. OER repositories can also use quality scores by themselves to provide quality-based sorting of OER search results. In this regard, it should be mentioned that OER repositories, such as MERLOT, use this approach, and that there is evidence that quality metrics based on learning object evaluation models, such as LORI, can be effectively used for this purpose, as well as for filtering low-quality OER [67], which can be useful, for instance, for allowing catalog-like applications to specify a minimum quality for the offered OER. Future studies should examine the usefulness of OER quality scores for improving existing search and discovery tools in the context of real-world OER repositories.

In order to confirm the results obtained in this study, we plan to conduct a new online experiment where OER recommendations generated according to the analyzed recommendation approaches will be presented to users of an OER repository under normal conditions over a long period of time. Another interesting line of future work would be to examine the usefulness of quality scores for enhancing other types of OER recommender systems in addition to those using content-based techniques, such as collaborative filtering, knowledge-based, or demographic recommender systems. Finally, we recommend research on sustainable solutions for OER repositories that have the capacity to provide effective quality assurance for a large number of digital learning resources.

Author Contributions: Conceptualization. A.G., D.L.-F., and K.V.; methodology. A.G., D.L.-F., and K.V.; software. A.G.; investigation. A.G., D.L.-F., and K.V.; writing—original draft preparation. A.G.; writing—review and editing. A.G., D.L.-F., and K.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. *Open Educational Resources: Breaking the Lockbox on Education (White Paper)*; William and Flora Hewlett Foundation: Menlo Park, CA, USA, 2013.
2. *Opening Up Education: Innovative Teaching and Learning for all through new Technologies and Open Educational Resources*; European Commission: Brussels, Belgium, 2013.
3. *Forum on the Impact of Open Courseware for Higher Education in Developing Countries*; UNESCO: Paris, France, 2002.
4. *Giving Knowledge for Free: The Emergence of Open Educational Resources*; OECD: Paris, France, 2007.
5. Wikimedia Commons. Available online: <http://commons.wikimedia.org> (accessed on 9 June 2020).
6. Europeana. Available online: <http://www.europeana.eu> (accessed on 9 June 2020).
7. OER Commons. Available online: <http://www.oercommons.org> (accessed on 9 June 2020).
8. MERLOT (Multimedia Educational Resource for Learning and Online Teaching). Available online: <http://www.merlot.org> (accessed on 9 June 2020).
9. Open Discovery Space Portal (ODS Portal). Available online: <http://www.opendiscoveryspace.eu> (accessed on 9 June 2020).
10. Learning Resource Exchange (LRE). Available online: <http://lreforschools.eun.org> (accessed on 9 June 2020).
11. Luo, T.; Hostetler, K.; Freeman, C.; Stefaniak, J. The power of open: Benefits, barriers, and strategies for integration of open educational resources. *Open Learn. J. Open Distance e-Learn.* **2019**, *35*, 140–158. [CrossRef]
12. De los Arcos, B.; Cirigottis, G.E.; Egglestone, N.; Farrow, R.; McAndrew, P.; Perryman, L.-A.; Pitt, R.; Weller, M. *OER Data Report 2013–2015: Building Understanding of Open Education*; OER Research Hub: Milton Keynes, UK, 2015.
13. Najjar, J.; Klerkx, J.; Vuorikari, R.; Duval, E. Finding Appropriate Learning Objects: An Empirical Evaluation. In *Research and Advanced Technology for Digital Libraries: Proceedings of the 9th European Conference on Digital Libraries (ECDL 2005)*; Rauber, A., Christodoulakis, S., Tjoa, A.M., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; pp. 323–335.
14. John, B.; Thavavel, V.; Jayaraj, J.; Muthukumar, A.; Jeevanandam, P.K. Comparative Analysis of Current Methods in Searching Open Education Content Repositories. *Online J. Sci. Technol.* **2016**, *6*, 21–29.
15. Kay, R.; Knaack, L.; Petrarca, D. Exploring Teachers Perceptions of Web-Based Learning Tools. *Interdiscip. J. E-Learn. Learn. Objects* **2009**, *5*, 27–50. [CrossRef]
16. Wiley, D.; Bliss, T.J.; McEwen, M. Open Educational Resources: A Review of the Literature. In *Handbook of Research on Educational Communications and Technology*; Spector, J.M., Merrill, M.D., Elen, J., Bishop, M.J., Eds.; Springer: New York, NY, USA, 2014; pp. 781–789.
17. Clements, K.; Pawlowski, J.; Manouselis, N. Open educational resources repositories literature review—Towards a comprehensive quality approaches framework. *Comput. Human Behav.* **2015**, *51*, 1098–1106. [CrossRef]
18. Larsen, K.; Vincent-Lancrin, S. The impact of ICT on tertiary education: Advances and promises. In *Advancing Knowledge and the Knowledge Economy*; Kahin, B., Foray, D., Eds.; MIT Press: Cambridge, MA, USA, 2006; pp. 151–168.
19. Atkins, D.E.; Brown, J.S.; Hammond, A.L. *A Review of the Open Educational Resources (OER) Movement: Achievements, Challenges, and New Opportunities*; Technical report to the William and Flora Hewlett Foundation: Menlo Park, CA, USA, 2007.
20. Clements, K.; Pawlowski, J. User-oriented quality for OER: Understanding teachers' views on re-use, quality, and trust. *J. Comput. Assist. Learn.* **2012**, *28*, 4–14. [CrossRef]
21. Vargo, J.; Nesbit, J.C.; Belfer, K.; Archambault, A. Learning Object Evaluation: Computer-Mediated Collaboration And Inter-Rater Reliability. *Int. J. Comput. Appl.* **2003**, *25*, 198–205. [CrossRef]

22. Tzikopoulos, A.; Manouselis, N.; Vuorikari, R. An Overview of Learning Object Repositories. In *Learning Objects for Instruction: Design and Evaluation*; Taylor Northrup, P., Ed.; IGI Global: Hershey, PA, USA, 2007; pp. 29–55.
23. Leacock, T.L.; Nesbit, J.C. A Framework for Evaluating the Quality of Multimedia Learning Resources. *Educ. Technol. Soc.* **2007**, *10*, 44–59.
24. Fernández-Pampillón, A. UNE 71362: Calidad de los materiales educativos digitales. *AENOR* **2017**, *329*, 44–47.
25. Kay, R. Evaluating learning, design, and engagement in web-based learning tools (WBLTs): The WBLT Evaluation Scale. *Comput. Human Behav.* **2011**, *27*, 1849–1856. [[CrossRef](#)]
26. Kay, R. Examining the Effectiveness of Web-Based Learning Tools in Middle and Secondary School Science Classrooms. *Interdiscip. J. E-Learning Learn. Objects* **2011**, *7*, 359–374. [[CrossRef](#)]
27. Domínguez, E.; de Armas, I.; Fernández-Pampillón, A. The COdA scoring rubric. In *Technology-Enhanced Language Learning for Specialized Domains: Practical Applications and Mobility*; Routledge: Abingdon, UK, 2016; pp. 86–94.
28. Kay, R.; Knaack, L. A multi-component model for assessing learning objects: The learning object evaluation metric (LOEM). *Australas. J. Educ. Technol.* **2008**, *24*, 574–591. [[CrossRef](#)]
29. Eguigure, Y.; Zapata, A.; Menendez, V.; Prieto, M. Quality evaluation model for learning objects from pedagogical perspective. A case of study. *Iberoam. J. Appl. Comput.* **2011**, *1*, 16–28.
30. Nesbit, J.C.; Leacock, T.L. Web-Based Tools for Collaborative Evaluation of Learning Resources. *J. Syst. Cybern. Informatics* **2006**, *3*, 102–112.
31. Luke, R.; Mallory, D.; Pinet, R.; Segiun, A. Evaluating Learning Objects with an Online Version of the Learning Object Review Instrument: Results of a Design Study. In *Proceedings of the World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, Montreal, QC, Canada, 15 October 2007.
32. Sarasa, A.; Fernández-Pampillón, A.; Rueda, A.; Riani, C. A web tool for assessment of the quality of digital educational materials. In *Proceedings of the 2016 International Symposium on Computers in Education (SIIE 2016)*, Salamanca, Spain, 13–15 September 2016.
33. Gordillo, A.; Barra, E.; Quemada, J. A flexible open source web platform to facilitate learning object evaluation. In *Proceedings of the 44th Frontiers in Education Conference (FIE 2014)*, Madrid, Spain, 22–25 October 2014; pp. 3080–3087.
34. Ochoa, X.; Duval, E. Relevance Ranking Metrics for Learning Objects. *IEEE Trans. Learn. Technol.* **2008**, *1*, 34–48. [[CrossRef](#)]
35. Yen, N.Y.; Shih, T.K.; Chao, L.R.; Jin, Q. Ranking Metrics and Search Guidance for Learning Object Repository. *IEEE Trans. Learn. Technol.* **2010**, *3*, 250–264. [[CrossRef](#)]
36. Sanz-Rodríguez, J.; Dodero, J.M.; Sánchez-Alonso, S. Ranking Learning Objects through Integration of Different Quality Indicators. *IEEE Trans. Learn. Technol.* **2010**, *3*, 358–363. [[CrossRef](#)]
37. Ochoa, X.; Duval, E. Use of Contextualized Attention Metadata for Ranking and Recommending Learning Objects. In *Proceedings of the 1st International Workshop on Contextualized Attention Metadata: Collecting, Managing and Exploiting of Rich Usage Information*; Association for Computing Machinery: New York, NY, USA, 2006.
38. Manouselis, N.; Drachsler, H.; Vuorikari, R.; Hummel, H.; Koper, R. Recommender Systems in Technology Enhanced Learning. In *Recommender Systems Handbook*; Ricci, F., Rokach, L., Shapira, B., Kantor, P.B., Eds.; Springer: Boston, MA, USA, 2011; pp. 387–415.
39. Drachsler, H.; Verbert, K.; Santos, O.C.; Manouselis, N. Panorama of Recommender Systems to Support Learning. In *Recommender Systems Handbook*; Ricci, F., Rokach, L., Shapira, B., Eds.; Springer: Boston, MA, USA, 2015; pp. 421–451.
40. Shakya, S.; Prasad, P.W.C.; Alsadoon, A.; Maag, A. A systematic review: Machine learning based recommendation systems for e-learning. *Educ. Inf. Technol.* **2019**.
41. Ricci, F.; Rokach, L.; Shapira, B. Introduction to Recommender Systems Handbook. In *Recommender Systems Handbook*; Ricci, F., Rokach, L., Shapira, B., Kantor, P.B., Eds.; Springer: Boston, MA, USA, 2011; pp. 1–35.
42. Bobadilla, J.; Ortega, F.; Hernando, A.; Gutiérrez, A. Recommender systems survey. *Knowl. Based Syst.* **2013**, *46*, 109–132. [[CrossRef](#)]

43. Gordillo, A. Contribution to the Authoring, Distribution, Evaluation and Integration of Learning Objects. Ph.D. Thesis, Universidad Politécnica de Madrid, Madrid, Spain, 2017.
44. Sampson, D.G.; Zervas, P. Learning object repositories as knowledge management systems. *Knowl. Manag. E-Learn.* **2013**, *5*, 117–136.
45. Zervas, P.; Alifragkis, C.; Sampson, D.G. A quantitative analysis of learning object repositories as knowledge management systems. *Knowl. Manag. E-Learn.* **2014**, *6*, 156–170.
46. Dey, A.K. Understanding and Using Context. *Pers. Ubiquitous Comput.* **2001**, *5*, 4–7. [[CrossRef](#)]
47. Adomavicius, G.; Tuzhilin, A. Context-Aware Recommender Systems. In *Recommender Systems Handbook*; Ricci, F., Rokach, L., Shapira, B., Kantor, P.B., Eds.; Springer: Boston, MA, USA, 2011; pp. 217–253.
48. Verbert, K.; Manouselis, N.; Ochoa, X.; Wolpers, M.; Drachsler, H.; Bosnic, I.; Duval, E. Context-aware recommender systems for learning: A survey and future challenges. *IEEE Trans. Learn. Technol.* **2012**, *5*, 318–335. [[CrossRef](#)]
49. Rodés-Paragarino, V.; Gewerc-Barujel, A.; Llamas-Nistal, M. Use of Repositories of Digital Educational Resources. State of the Art Review. *IEEE Rev. Iberoam. Technol. Aprendiz.* **2016**, *11*, 73–78. [[CrossRef](#)]
50. Clough, P.; Otegi, A.; Agirre, E.; Hall, M. Implementing Recommendations in the PATHS System. In *Theory and Practice of Digital Libraries—TPDL 2013 Selected Workshops*; Springer International Publishing: Cham, Switzerland, 2014; pp. 169–173.
51. Otegi, A.; Agirre, E.; Clough, P. Personalised PageRank for Making Recommendations in Digital Cultural Heritage Collections. In Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2014), London, UK, 8–12 September 2014; pp. 49–52.
52. Gordillo, A.; Charleer, S.; Verbert, K. An open source hybrid learning object recommender system based on Europeana. In Proceedings of the 11th International Conference of Education, Research and Innovation (ICERI 2018), Seville, Spain, 12–14 November 2018; pp. 8862–8872.
53. Gordillo, A.; Barra, E.; Quemada, J. A hybrid recommendation model for learning object repositories. *IEEE Lat. Am. Trans.* **2017**, *15*, 462–473. [[CrossRef](#)]
54. Sergis, S.; Sampson, D.G. Learning Object Recommendations for Teachers Based On Elicited ICT Competence Profiles. *IEEE Trans. Learn. Technol.* **2016**, *9*, 67–80. [[CrossRef](#)]
55. Lemos, H.; Cechinel, C.; Giustozzi, F.; Casali, A.; Deco, C. Users opinions about Learning Object Recommendations: A case study. In Proceedings of the 2016 International Symposium on Computers in Education (SIIE), Salamanca, Spain, 13–15 September 2016.
56. Fraihat, S.; Shambour, Q. A Framework of Semantic Recommender System for e-Learning. *J. Softw.* **2015**, *10*, 317–330. [[CrossRef](#)]
57. Zapata, A.; Menéndez, V.H.; Prieto, M.E.; Romero, C. A framework for recommendation in learning object repositories: An example of application in civil engineering. *Adv. Eng. Softw.* **2013**, *56*, 1–14. [[CrossRef](#)]
58. Zapata, A.; Menéndez, V.H.; Prieto, M.E.; Romero, C. Evaluation and selection of group recommendation strategies for collaborative searching of learning objects. *Int. J. Hum. Comput. Stud.* **2015**, *76*, 22–39. [[CrossRef](#)]
59. Ruiz-Iniesta, A.; Jiménez-Díaz, G.; Gómez-Albarrán, M. A semantically enriched context-aware OER recommendation strategy and its application to a computer science OER repository. *IEEE Trans. Educ.* **2014**, *57*, 255–260. [[CrossRef](#)]
60. Cechinel, C.; Sicilia, M.Á.; Sánchez-Alonso, S.; García-Barriocanal, E. Evaluating collaborative filtering recommendations inside large learning object repositories. *Inf. Process. Manag.* **2013**, *49*, 34–50. [[CrossRef](#)]
61. Gallego, D.; Barra, E.; Aguirre, S.; Huecas, G. A Model for Generating Proactive Context-Aware Recommendations in e-Learning Systems. In Proceedings of the 2012 Frontiers in Education Conference (FIE 2012), Seattle, WA, USA, 3–6 October 2012.
62. Gallego, D.; Barra, E.; Rodríguez, P.; Huecas, G. Incorporating Proactivity to Context-Aware Recommender Systems for E-Learning. In Proceedings of the 2013 World Congress on Computer and Information Technology (WCCIT 2013), Sousse, Tunisia, 22–24 June 2013.
63. Gallego, D.; Barra, E.; Gordillo, A.; Huecas, G. Enhanced recommendations for e-learning authoring tools based on a proactive context-aware recommender. In Proceedings of the 43rd Frontiers in Education Conference (FIE 2013), Oklahoma City, OK, USA, 23–26 October 2013; pp. 1393–1395.

64. McNee, S.M.; Riedl, J.; Konstan, J.A. Accurate is not always good: How Accuracy Metrics have hurt Recommender Systems. In Proceedings of the Extended Abstracts of the 2006 ACM Conference on Human Factors in Computing Systems (CHI 2006), Montreal, QC, Canada, 22–27 April 2006.
65. 1484.12.1-2002—IEEE Standard for Learning Object Metadata; The Institute of Electrical and Electronics Engineers: New York, NY, USA, 2002.
66. ADL (Advanced Distributed Learning) Initiative SCORM, 4th ed.; 2004. Available online: <http://adlnet.gov/projects/scorm-2004-4th-edition> (accessed on 9 June 2020).
67. Gordillo, A.; Barra, E.; Quemada, J. Towards a learning object pedagogical quality metric based on the LORI evaluation model. In Proceedings of the 44th Frontiers in Education Conference (FIE 2014), Madrid, Spain, 22–25 October 2014; pp. 3088–3095.
68. Robertson, S. Understanding inverse document frequency: On theoretical arguments for IDF. *J. Doc.* **2004**, *60*, 503–520. [[CrossRef](#)]
69. Shani, G.; Gunawardana, A. Evaluating Recommendation Systems. In *Recommender Systems Handbook*; Ricci, F., Rokach, L., Shapira, B., Kantor, P.B., Eds.; Springer: Boston, MA, USA, 2011; pp. 257–297.
70. Cechinel, C.; da Silva Camargo, S.; Sicilia, M.Á.; Sánchez-Alonso, S. Mining Models for Automated Quality Assessment of Learning Objects. *J. Univers. Comput. Sci.* **2016**, *22*, 94–113.
71. Gordillo, A.; Barra, E.; Quemada, J. Estimación de calidad de objetos de aprendizaje en repositorios de recursos educativos abiertos basada en las interacciones de los estudiantes. *Educ. XXI* **2018**, *21*, 285–302. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).