WebEduRank: an Educational Ranking Principle of Web Resources for Teaching

Alessandro Marani

School of Information and Communication Technology,
Griffith University,
170 Kessels Road, Nathan, QLD, 4111 Australia
alessandro.marani@griffithuni.edu.au

Abstract. The seeking of teaching resources on the web is a very common practice among instructors. Some on-line systems aim to support instructors in this complex and delicate task. In fact, quality of teaching is without doubt an important factor for learning, and an important aspect of teaching is the delivery of quality learning material to students. Remarkable results have been achieved in the reuse and recommendation of teaching resources, producing the corner stone for Information Retrieval (IR) in education. Our research aims to connect those findings with web resources, overcoming the current limitation of analysing resources from educational repositories only. Firstly, this study addresses the problem of identifying those attributes of a teaching context that actually carry on a piece of information useful for rating web resources for teaching. Then, the project works on the design of WebEduRank, a ranking principle for rating web-resources according to the teaching requirements and needs of an instructor. So far, the research has already identified the attributes that are expected to positively inform the *WebEduRank*. Using these attributes, the design of the *WebEduRank*is now completed and the methodology for its validation is already devised and implemented.

Keywords: Educational ranking principle, teaching resources retrieval, teaching context

1 Introduction

Learning Object Repositories (LOR) are a very reliable source of Learning Objects (LO), where the recommendation approaches are facilitated by the metadata that those objects come with [3]. A number of issues in annotating such metadata, in both a manual and an automatic manner, are well known and some solutions have been proposed [2, 10]. However, there is still the limitation due to the low amount of resources hosted by such repositories, and low completeness of annotation of LOs metadata in LORs [10, 12].

Focusing on the applications of IR techniques for the retrieval of resources for teaching, it would be interesting to do not limit this search in resources hosted by LORs or other educational datasets, but to explore also the web. In practice, we are talking of going out on the web where we cannot rely on the fact that we have educational metadata like in LORs, and we do not even know if they actually are educational resources.

This research mainly aims to propose a method, called *WebEduRank*, for ranking or re-ranking web resources according to their suitability for a teaching context. The desired goal for the novel ranking principle is to provide a higher ranking of those web resources that present characteristics of interest for a teaching context.

2 Recommendation of Learning Objects

Learning Object is currently the most relevant approach for the description of educational characteristics of digital content, and it is adopted by most of the recommender systems of educational resources [3]. A Learning Object is a multimedia content provided with instructional information for the reuse of the content. This instructional information, called metadata, should be able to describe the educational material. However, Learning Object metadata has some semantic issues that makes information processing hard for computers [10]. These issues have to be carefully considered for the purposes of this research.

Learning Object has been widely used by recommender systems for achieving different recommendation goals. Some studies focused on recommending Learning Objects according to course structure and learning styles, among other aspects [13]. Hence, Learning Object metadata resulted useful in this field, but the current popular schema of Learning Object metadata are not enough for describing the educational traits of resources, especially depicting how they can be used to achieve some learning (or teaching) goals [1].

The metadata of Learning Objects can also facilitate the process of query expansion for a more effective retrieval of resources [6], but yet there is concern about the semantic ambiguity and completeness of annotated Learning Object metadata [11, 10, 12].

3 Attributes of teaching context for informing *WebEduRank*

This section aims to present the process for designing a teaching context towards the proposal of *WebEduRank*. The teaching context has the delicate role of representing the teaching requirements and experiences of instructors. This task is very important for proposing a ranking principle of web resources able to take into account the educational context of a profiled instructor. In order to achieve this goal, the teaching context should exhaustively represent the teaching requirements of an instructor that the web resources have to comply with.

From an analysis of a review study of instructors' knowledge involved in the process of teaching (called *teaching knowledge*), we have found that most of the contributions refer to some important categories of teaching knowledge such as *Content Knowledge*, *Pedagogical Knowledge* and *Pedagogical Content Knowledge* among others [15]. Also, the review of some significant studies for analysing the completeness rate and usefulness of some metadata in Learning Object Repositories can provide further attributes to be included in the *Instructor Profile*. [11] and [12] report some statistics on the completeness of Learning Object metadata in two different repositories. The first study analyses Learning Objects in GLOBE, an important repository, whilst the second one considers those Learning Objects used in a small repository of only agricultural materials called *Organic.Edunet*.

After the review of literature about the *teaching knowledge* and Learning Object, a first draft of the teaching context is presented. We suggest a teaching context TC of a course formed by *Prerequisite Knowledge* (PK), *Concept Name* (CN), *Course Title* (CT), *Education Level* (ED-LEV), *Difficulty* (DIFF), *Starting Knowledge* (SK), *Target Knowledge* (TK).

3.1 Evaluation of the attributes for informing the WebEduRank

This evaluation is only a first refining of the teaching context, so that the *WebEduRank* can be focused on only those aspects that are likely to be useful for our purposes. Then, the final evaluation of the *WebEduRank* is the final evidence whether or not such aspects have been useful and exhaustive. The evaluation has been carried out by surveying instructors about the usefulness of the attributes of the proposed teaching context for ranking web resources for teaching.

The research question that piloted the formulation and structure of the questionnaire is the following:

RQ_1: What information of the proposed Teaching Context is perceived useful by instructors for informing the WebEduRank for the retrieval of web resources for teaching?

For addressing the research question, the questionnaire simply focused on presenting the attributes of the *Teaching Context* to the participants and asking them the perceived usefulness and comprehensiveness of such information for the process of ranking teaching resources. So far, a total of 33 responses have been collected; most of the participants have been recruited by e-mail. Such number of participants already allows us to conduct some statistical analysis with a good degree of significance.

Structure of the questionnaire We approached the research question with a questionnaire of 11 items divided into three constructs as follow:

- General questions (3).
- Attitude in using the Internet for seeking teaching material (6).
- Perceived usefulness of the attributes of the teaching context (2).

The first construct provides some general information about the sample such as age and teaching experience. We assume that instructors with different age, teaching experience and place where they teach consider a teaching context similarly. Hence, this first construct is only for describing the sample and for further analysis of the results.

Instead, the second construct shows the current attitude, and so practical experience, of instructors in using current on-line systems for seeking teaching resources. This aspect should also tell us if the sample massively use the Internet for such purposes, or it is just a beginner user in this regard. We believe that answers provided by instructors that massively use the Internet for teaching can provide us with more reliable and, more importantly, experienced feedback. In this construct, there are six 5-point Likert scale questions.

Finally, the last construct is related to the perceived usefulness of the attributes in the proposed teaching context for seeking teaching resources on the web. This evaluation is based on what information instructors usually consider when they look for resources on the web; information used in different phases of the search, from formulating the query to browsing the results. Two items belong to this construct. The first item is a multiple-choice question where instructors are asked to select those attributes that they think useful when searching teaching resources on the web for a specific concept. Participants have to choose at least one option, where also the 'none of the above' option is available. There is an open field for additional considerations by participants. The second question is a matrix of 5-points Likert scales, where each row represents one attribute of the proposed *Teaching Context*; participants are required to assign a relevance score to each attribute. Also for this question, there is an open field for additional comments from the participants.

Data analysis For analysing the results of the questionnaire, statistics for measuring the perceived usefulness of respondents and repeatability of the results is used. The questions involved in this task are 5-points Likert scales, so we can use statistical hypotheses test methods for rejecting or not the null hypothesis that an attribute in the *Teaching Context* is believed useless by the population of instructors, not only the sample of this study. The null hypothesis can be stated as follows:

$$H_0 : \mu_{relevance} \leqslant 3.0 \tag{1}$$

If H_0 is rejected, it means that the research hypothesis, namely the true mean relevance for the population is greater than 3.0, is retained. A value greater than 3.0 in a 5-points Likert scale is a positive response to the relevance of the attribute. The null hypothesis is tested for all the attributes in the *Teaching Context*.

The application of t for this study is two-fold: i) retain or reject H_0 for the population, and ii) compute a confidence interval for the population mean of the relevance of all the attributes in the *Teaching Context*.

Relevance of the attributes of the *Teaching Context* This section presents the perceived usefulness of the attributes of the *Teaching Context* according to the surveyed instructors (N = 33). We use t distributions for testing H_0 for all the attributes of the *Teaching Context*. For each attribute, it is reported whether or not H_0 (it means the attribute is not relevant) is retained or rejected.

Two questions of the questionnaire are focused on the relevance of the attributes of the *Teaching Context*; only the most important one for the purposes of the *WebEdu-Rank* is presented here.

By means of 5-point Likert scales, participants evaluated the relevance of all the attributes of the *Teaching Context*, in order to capture to what extent the attributes of the proposed *Teaching Context* actually represent important aspects of a teaching context, independently of the specific retrieval task. For each attribute, all the participants are

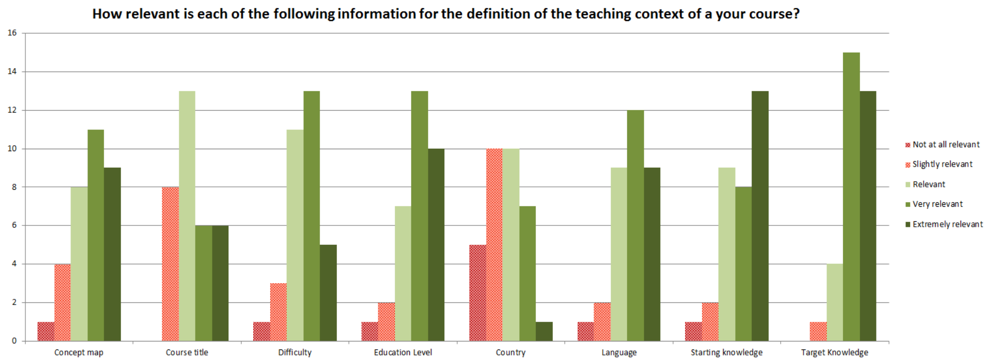

Fig. 1. Relevance of the attributes of the proposed *Teaching Context* for an effective and exhaustive description of a real teaching context.

required to express the relevance. Looking at the results achieved so far, other than the attributes *country* and *course title*, all the attributes got a very positive outcome.

In order to reject H_0 for the degree of freedom equal to 32 (N-1), t must be greater than 2.738 at .01 significance level. All the attributes but *course title* and *country* rejected H_0, presenting a t-value significantly over 2.738 and p-values smaller than .05.

course title is a borderline case where H_0 is retained but the p-value is just over .05 and the lower bound of the confidence interval mean for the population is 2.995, slightly lower than the desired value of 3.0. Totally different is the situation for the attribute *country*, where the lower bound of the population mean is 2.348, remarkably lower than 3.0. For this reason, only the attribute *country* is removed.

4 The *WebEduRank*

The *WebEduRank* aims to rate web pages according to some aspects of the teaching of an instructor. The rating should reflect the suitability of a web resource for teaching a certain concept in a particular teaching context or situation. A first note is that the proposed principle is not much focused on the ranking of the resources according to a topic. Google is already able to do it very well and with a remarkable performance. The main problem that the *WebEduRank* is trying to solve is the ordering of a set of web resources according to their suitability for teaching in a certain context. At this stage of the research, we consider as input a set of web resources that is assumed to be about the topic of a query, for example the results of a Google interrogation. The novel contribution of the *WebEduRank* would be the re-ranking of the web resources about a topic according to the teaching context of an instructor. This task is not simple at all, mainly because of the problematic structure of a web resource and the detection of educational attributes by text-analysis only. The structure of web resources, sometimes even unstructured, is a very well-known problem, challenging many activities around the text-analysis of web pages including information retrieval. However, this issue is better addressed by studies related to semantic web and web crawling, which are not research areas of this project.

Instead, the detection of educational attributes by text-analysis is a main problem to be addressed in this research. In order to achieve this goal, the *WebEduRank* has to be able to i) capture educational traits of web resources, and then ii) use such traits for ranking them accordingly to the teaching context of the instructor. These two problems are addressed in this research, where the traits to be extracted from web resources are those included in the *Instructor Profile* presented by this thesis.

4.1 Attributes of the *Instructor Profile* for informing the *WebEduRank*

According to the finding in Section 3, during the search of teaching resources on the web, an instructor knows that he is looking for resources about a concept (CN) for a particular course (CT), when some prerequisite knowledge (PK) is assumed to be already learnt by the students. Implicitly, the resources have to comply with the education level (ED-LEV) of the class with a certain level of difficulty (DIFF). The course may also be a bit different in terms of concepts taught in it because of the starting knowledge (SK) of a class. Finally, the instructor knows what is the target knowledge (TK) of the course. Some of these elements, such as *PK*, *SK* and *TK* can be derived by analysing the concept map of the teaching context, where the *SK* can be represented by source concepts (concepts with no prerequisites) in the map, and the *TK* can consist of the leaves of the concept map.

The proposed *WebEduRank* should replicate the aforementioned behaviour of instructors when ranking web resources for teaching. The *Teaching Context* here presented can provide the *WebEduRank* with such information, now the problem is how to use these attributes for analysing web resources accordingly to the process previously described. The purpose is to use these data for ranking the resources, not for formulating or expanding the user query. It would be interesting to deeply explore also these aspects, but they are out of the reach of this Ph.D. project. Only a shallow investigation has been conducted within this research, and the results are presented in [7]. The attributes for informing the *WebEduRank* is a first important finding of this research towards the design of the *WebEduRank* itslef.

The next challenge is in which parts of a web page the principle should look for that information. We can expect that some information of the teaching context is more likely to be found in the links of a web page (e.g. prerequisite knowledge or concepts) than the body or the title. Usually, for ranking web pages the body is the part of the page that is mostly analysed beside to some metadata if available. However, within the body of a web page we can find different tags that may contain different kind of information. For example, links (expressed by the html tag *a*) can have as text the name of other related concepts. So, we want to distinguish the text that comes from the following four parts of a web page: *title*, *body*, *links* and *highlights*. We believe that these parts of a web page can contain different attributes of the *Teaching Context*, reflecting a more informative information than just finding the information in the whole body-text. The identification of those four parts of the web pages is based on associating some HTML tags that are usually used for expressing those parts of a web page. Table 1 shows the HTML tags that are extracted from a web page and whose text is used for composing the four texts of the web page that the *WebEduRank* will analyse.

Part of web page	HTML tags
Title	title
Body	body
Links	a
Highlights	strong, h3, h2, h1, b

Table 1. For each of the four parts of a web page analysed by the *WebEduRank*, it is reported the HTML tags that are used for composing their texts.

4.2 The Expectancy Appearance Matrix

The last step for the design of the *WebEduRank* is defining the way of how the attributes of the *Teaching Context* are matched or searched into the four texts representing the web page. As anticipated earlier, we expect that the attributes can appear differently, we can say with a certain likelihood, into the four components of the web page defined in this research. In this stage, we have 7 attributes of the *Teaching Context* considered over the rating process of *WebEduRank*. We propose the *Expectancy Appearance Matrix* (EAM) for representing the expectation that an attribute appears into a component of the web page. Formally, EAM is a 4x7 matrix, where the rows are the four components of the web pages analysed by the *WebEduRank*, and the columns are the attributes of the *Teaching Context*. The element of $a_{ij} \in EAM$ is the likelihood that the j-th attribute is found in the i-th part of the web page.The events of finding or not an attribute in a certain part of the web page are independent of each other, because an attribute can appear in a section of the web page without any correlation with the appearance in other sections. In this regard, the definition of the elements of the EAM raises a new challenge, allowing the tuning of the *WebEduRank* for discovering the most appropriate values for each element of the matrix. The analysis of the dataset DAJEE [4] can provide some insights about the values to be assigned to the elements of the EAM.

4.3 Formulation of the *WebEduRank*

The idea behind the *WebEduRank* is to find a match between the text of a web resource and the text of the attributes of the *Teaching Context*, being the values of the *Teaching Context* are texts for describing the different aspects of the teaching. The matching of the texts is performed dividing the web page into four sections providing different information about the content of the web page. The *WebEduRank* basically analyses the frequency of the terms of the values of the attributes of the *Teaching Context* with a weighting system based on the values in the EAM. As introduced in Section 4.2, the columns EAM indicate the expectancy that an attribute of the *Teaching Context* of an instructor appears in the four sections of the web page. Hence, the frequency of the terms of an attribute j-th into a section i-th of the web page is weighted by the expectancy value a_{ij} reported by the EAM. In practice, each column of the EAM is a vector stating the importance of considering an attribute of the *Teaching Context* along the four sections of a web page.

5 Evaluation of the *WebEduRank*

This section presents the experiment for validating the *WebEduRank*. Part of this experiment consists of collecting teaching contexts from instructors. Then, instructors are asked to provide the rating of a set of web resources according to their suitability for teaching a concept in the context defined by them. This preliminary phase has to be conducted because the teaching context used for informing the *WebEduRank* is a novel approach, so there are no datasets that we can use for our purposes. For this reason, we need instructors for defining a collection of teaching contexts and web resources so that we can run a number of experiments for tuning and validating our new proposal against current practice and baseline methods. In addition, we do not have the capability of Google in indexing most of the web pages on the web, so the data collected are about web pages retrieved by Google after a query submitted by instructors. In this way, the *WebEduRank* proposes a ranking of the same resources presented by Google, overcoming the issue of gathering web pages to be rated by instructors in a particular teaching context.

The goal of this evaluation is two-fold: tuning the *WebEduRank* with different settings of the EAM and proving its improvement in the ranking of web resources for teaching compared to some baselines. As said in the previous section, the EAM is a critical aspect of the *WebEduRank*, so it needs to be carefully tuned for achieving the best performance that the *WebEduRank* can actually reach. For this reason, we can have a number of different settings of the ranking principle to be evaluated against themselves first, and then against other approaches. Therefore, we can formulate the following research question:

RQ_2: Given a set of web resources about a topic, can the WebEduRank offer an educational ranking of the web resources better than current practices?

5.1 Methodology

For addressing the research question, the experiments need a set of data on which the benchmark systems and our proposal can be run. Usually, web search results produced by an IR system are compared to what the users of those systems actually believe useful and relevant for their query. In this regard, it has been pointed out a small but yet important difference between *usefulness* and *relevance* of an object for the web search [9]. In our case, we have to make sure that external assessors rate the *usefulness* of web resources retrieved by Google more than their relevance to the query. The best way for avoiding this issue is providing external assessors with the highest level of knowledge and awareness of the purpose of the web search. Hence, it is recommended to i) clearly describe the contextual information about the web search, ii) use very informative query, not too short or ambiguous, iii) state clearly the purpose and need of the user [9].

In this study we need to undertake the data collection about the usefulness of some web resources for teaching of a concept in a certain teaching context; this is the purpose (or user need) in our case. For this reason, we ask instructors to i) define a teaching context of their interest, which includes a concept map, ii) formulate a query for retrieving web resources for a concept of the concept map using Google, and iii) rate the retrieved resources according to their usefulness for teaching the concept in the given teaching context. In practice, instructors are the external assessors of the web pages presented by Google, where the same assessors define the teaching context and formulate the query. We believe that this protocol for the data collection cover the three points highlighted by [9] for actually getting the assessment of *usefulness*, not just *relevance*, of web pages.

Once the data collection phase is completed, we can finally run a number of experiments for validating our proposal. Traditionally, when evaluating an IR system the rating of resources provided by the external assessors are considered the most reliable and truthful ratings. A number of metrics are usually involved for measuring how close to or far from the assessors' ratings is the ranking or order of the resources produced by the novel IR system.

5.2 Structure of the experiments

A first experiment is used for comparing the performance of the different tuning of the *WebEduRank* and some baselines. The goal of this experiment is to have evidence of the improved ranking produced by the novel approach against some simple baseline methods. The baseline methods consist of two scoring approaches: i) the documents are scored based only on the terms in the query provided by the instructors during the data collection phase, ii) plain *WebEduRank* where TF-IDF score of the documents is applied using a query composed by the values of the attributes of the *Teaching Context*. More details about the benchmark systems and baseline are reported in the next subsection. For this first experiment, prediction-accuracy metrics are more suitable for electing the best approach. In particular, Root Mean Squared Error (RMSE) and Mean Average Precision (MAP) [14] are used. The first metric measures how close to actual users' ratings are the ratings predicted by the subject approaches as follows:

$$RMSE = \sqrt{\frac{1}{|items|} \sum_{i \in items} (\widehat{r_i} - r_i)^2} \tag{2}$$

where *items* is the set of web resources rated by instructors during the phase of data collection, r_i is the rating for the resource i-th given by the user and $\widehat{r_i}$ is the rating for the i-th object predicted by the system. The latter metric simply computes the mean of the average precision of the results presented by the subject system. In this regard, a resource is considered relevant when it has been given a rate of 3 out of 5 by the external assessors. During the collection data phase, the first ten results presented by Google have been rated by the assessors, so we decide to average the top-3 precision measure of the result sets. The results are ordered by descent rating as produced by the subject system. It is expected that a configuration of the *WebEduRank* performs better than the baselines according to these two metrics. The best tuning of the *WebEduRank* will be tested against the benchmark systems in the next experiment.

During the second experiment, the *WebEduRank* is tested against Google which is pointed out to be the most popular system for retrieving teaching resources [8]. The *WebEduRank* should benefit of the information provided by the *Teaching Context* for predicting more accurately the usefulness of a web resource for teaching. In this experiment, we use position-based and prediction-accuracy metrics for comparing the two approaches. About the prediction-accuracy, we can use only the MAP metric because from Google we have only the ranking of the resources, not the ratings. Among the position-based metrics we choose the Discounted Cumulative Gain (DCG) [5] and the Average Precision Correlation Coefficient (τ_{AP}) [16]. Both metrics work analysing the position of the presented resources by the subject systems against their actual usefulness.

The DCG metric tells us to what extent the produced ranking reflects the relevance of the resources, by discounting the gain for users of recommending relevant resources at lower positions. The formula of DCG given a list of items J is the following [5]:

$$DCG = \frac{CG[j]}{max(1, \log_2 j)}$$

where $CG[j]$ expresses the cumulative gain that the user has from being recommended the item at position j within the list of items J. The gain values are cumulative, so, for each J, it is built a Gain vector (G) of the 10 presented items (ten is the number of resources rated for each query). The element $G[j]$ is the user-relevance of the item returned in position j. Then, the Cumulated Gain (CG) is simply the sum of the gains from position 1 to j, recursively defined as follows:

$$CG[j] = \begin{cases} G[1], if\ j = 1 \\ CG[j-1] + G[j], otherwise \end{cases} \tag{3}$$

Similarly, τ_{AP} works on the positions of the items but with a different approach. It measures the differences in the ordering of resources as provided by the user and the one produced by the subject system. This particular metric is more sensible of differences happening in the highest positions of the list than the ones in the bottom of the list. Hence, τ_{AP} is more strict when resources at the top positions of the user ranking appear differently or appear at the lowest positions in the order generated by the subject method. The formula of this metric is [16]:

$$\tau_{AP} = \frac{2}{J-1} \sum_{i=2}^{J} \left(\frac{C(i)}{i-1}\right) - 1 \tag{4}$$

where $C(i)$ is the number of correctly ranked items above position i.

5.3 Benchmark system and baselines

The only benchmark system used in this study is Google. We do not need to spend many words for justifying this decision, being Google the most reliable search engine also for teaching. However, before testing our proposal against Google, we have to show some improvements that WebEduRank can offer compared to some baseline approaches. Also, this comparison allows us to test the *WebEduRank* with different tuning of the EAM. In this study, two baseline approaches are used. The first one does not use any information from the *Teaching Context* and it is used with the aim of motivating the usage of the *Teaching Context* for informing the *WebEduRank*. The second approach is pretty similar to the method of the *WebEduRank*, but without using the EAM. Both methods use the same score function for predicting the rating of the items, based on the TF-IDF score of a document given the query. The only difference is in the formulation of the query; the first approach uses the same query used by instructors during the collection data phase (no usage of *Teaching Context*), while the second approach builds the query using the attributes of the *Teaching Context* by following the structure of the fifth query presented in Table 2 of [7]. Both approaches do not divide the web page into different parts, but only the body-text is analysed. Given the query text q and the body-text $text$, the rating is computed as follows:

$$score(q,text) = \sum_{t \in q} \sqrt{frequency(t \ in \ text) \cdot idf(t)^2} \tag{5}$$

where $idf(t)$ is defined as follows:

$$idf(t) = 1 + \log \frac{numDocs}{docFreq(t)+1}$$

This score method of documents given a query is taken from the *TFIDFSimilarity* class of Apache Lucene[1], after removing the normalisation and boosting factors.

6 Potential Advances

A successful proposal of a ranking principle of web resources for educational users (i.e. students and instructors) would be a very important improvement in the field of recommender systems in Technology Advanced Learning (TEL). As discussed along this contribution, the predominant limitation of current solutions in TEL against Google or other generic search engines is the number of resources and data that these systems can handle. The proposal of this doctoral work will put a first important step towards the consideration of web resources into the recommendation process of teaching resources.

As part of the proposal of this research, the definition of the teaching context, which is expected to positively inform the proposed *WebEduRank*, may finally concentrate the proposal of new recommendation approaches on those aspects that are actually important for such task. Finally, the impact of building such teaching context should be evaluated. After a first analysis of the data about MOOCs from Coursera offered by the DAJEE dataset [4], it is anticipated that most of the elements of the proposed *Teaching Context*, among other data, can extracted out of the information available in DAJEE (and so in MOOCs).

[1] https://lucene.apache.org/core/4_0_0/core/org/apache/lucene/search/similarities/TFIDFSimilarity.html

References

1. Bozo, J., Alarcón, R., Iribarra, S.: Recommending learning objects according to a teachers' contex model. In: Sustaining TEL: From innovation to learning and practice, pp. 470–475. Springer (2010)
2. Dietze, S., Yu, H.Q., Giordano, D., Kaldoudi, E., Dovrolis, N., Taibi, D.: Linked education: interlinking educational resources and the web of data. In: Proceedings of the 27th Annual ACM Symposium on Applied Computing. pp. 366–371. ACM (2012)
3. Drachsler, H., Verbert, K., Santos, O.C., Manouselis, N.: Panorama of recommender systems to support learning. In: Recommender systems handbook, pp. 421–451. Springer (2015)
4. Estivill-Castro, V., Limongelli, C., Lombardi, M., Marani, A.: Dajee: A dataset of joint educational entities for information retrieval in technology enhanced learning. In: Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 681–684. SIGIR '16, ACM (2016)
5. Järvelin, K., Kekäläinen, J.: Cumulated gain-based evaluation of ir techniques. ACM Transactions on Information Systems (TOIS) 20(4), 422–446 (2002)
6. Limongelli, C., Lombardi, M., Marani, A., Sciarrone, F., Temperini, M.: A recommendation module to help teachers build courses through the moodle learning management system. New Review of Hypermedia and Multimedia 22(1–2), 58–82 (2015)
7. Lombardi, M., Marani, A.: A comparative framework to evaluate recommender systems in technology enhanced learning: a case study. In: Advances in Artificial Intelligence and Its Applications, pp. 155–170. Springer (2015)
8. Maloney, S., Moss, A., Keating, J., Kotsanas, G., Morgan, P.: Sharing teaching and learning resources: perceptions of a university's faculty members. Medical education 47(8), 811–819 (2013)
9. Mao, J., Liu, Y., Zhou, K., Nie, J.Y., Song, J., Zhang, M., Ma, S., Sun, J., Luo, H.: When does relevance mean usefulness and user satisfaction in web search? In: Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 463–472. SIGIR '16, ACM (2016)
10. Mylonakis, M., Arapi, P., Pappas, N., Moumoutzis, N., Christodoulakis, S.: Metadata management and sharing in multimedia open learning environment (mole). In: Metadata and Semantic Research, pp. 275–286. Springer (2011)
11. Ochoa, X., Klerkx, J., Vandeputte, B., Duval, E.: On the use of learning object metadata: The globe experience. In: Towards ubiquitous learning, pp. 271–284. Springer (2011)
12. Palavitsinis, N., Manouselis, N., Sanchez-Alonso, S.: Metadata quality in learning object repositories: a case study. The Electronic Library 32(1), 62–82 (2014)
13. Rodríguez, P.A., Tabares, V., Mendez, N.D.D., Carranza, D.A.O., Vicari, R.M.: Broa: An agent-based model to recommend relevant learning objects from repository federations adapted to learner profile. IJIMAI 2(1), 6–11 (2013)
14. Shani, G., Gunawardana, A.: Evaluating recommendation systems. In: Recommender systems handbook, pp. 257–297. Springer (2011)
15. Voogt, J., Fisser, P., Pareja Roblin, N., Tondeur, J., van Braak, J.: Technological pedagogical content knowledge–a review of the literature. Journal of Computer Assisted Learning 29(2), 109–121 (2013)
16. Yilmaz, E., Aslam, J.A., Robertson, S.: A new rank correlation coefficient for information retrieval. In: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval. pp. 587–594. ACM (2008)