Preferences of Users on Cross-Site OER Recommendations: Stay or Leave?

Ayşe Saliha Sunar
assunar@beu.edu.tr
Jozef Stefan Institute
Ljubljana, Slovenia
Bitlis Eren University
Bitlis, Turkey

Erik Novak
erik.novak@ijs.si
Jozef Stefan Institute
Jozef Stefan International
Postgraduate School
Ljubljana, Slovenia

Jasna Urbančič
jasna.urbancic@ijs.si
Jozef Stefan Institute
Ljubljana, Slovenia

Dunja Mladenić
dunja.mladenic@ijs.si
Jozef Stefan Institute
Jozef Stefan International
Postgraduate School
Ljubljana, Slovenia

ABSTRACT

In education we can find different open educational resource (OER) providers that are serving resources in different modalities, formats and languages. These providers can be the actual resource creators or re-distributors that redirect the user to the actual provider. In recent work, we developed a recommendation engine which provides content-based recommendations from multiple resource providers, enabling the users to navigate between the providers and their resources. In this paper, we investigate the users' choice on the recommended items focusing on the cross-site user learning activities. The results show that the users tend to stay on the same website and not choose the first item in the recommendation list.

CCS CONCEPTS

• **Information systems** → *Content ranking*.

KEYWORDS

open educational resources, recommendation system, cross-site recommendations, learning analytics, data visualization

1 INTRODUCTION

Open Educational Resources (OERs), as defined by the UNESCO,[1] are teaching, learning and research materials, digital or otherwise, that reside in the public domain or have been released under an open license that permits no-cost access, use, adaptation and redistribution by others with no or limited restrictions. As such, digital OERs have many advantages over traditional learning materials. Namely, OERs reduce the costs for the students, the dissemination of information is faster, and the resources can be accessed from everywhere. However, there are also some reservations, such as quality and reliability of the materials, and intellectual property rights ownership. Additionally, an OER user faces a very fragmented landscape of repositories containing OERs, which makes finding relevant OERs a difficult task for both students and teachers.

Therefore, we aim to connect the scattered OERs by enriching the material with additional semantic information, automatic and machine translation, as well as providing services for cross-site recommendations to make finding appropriate OERs easier. Currently, the repositories are highly specialised in terms of scientific domains, content type, level of education, and language. From the perspective of a student, the student may have to search OERs in different repositories for each class which is an undesirable situation. Such search is inefficient and time consuming, it also leads to sub-optimal search results and a negative user experience with OERs.

In this paper, we aim to provide some insight into the preferences of the users regarding cross-site recommendations. To discover users' preferences, we analyse user transition

---

[1]Definition adopted from https://en.unesco.org/themes/building-knowledge-societies/oer.

data from content-based cross-site recommender engine embedded into two OERs repositories. We focus on the cross-events, which happen when the user selects OER materials from a different repository from the list of recommendations. The findings are the first step in evaluating how users perceive such recommendations, how good the recommender is, and to further improve the recommendations.

2 RELATED WORK

Recommender engines produce the results which are identified as the most relevant to the user by using different methods such as content-based or collaborative filtering. Considering the personal differences such as interest, study goal, time restriction, and capabilities of owned devices, not all users would choose the first ranked item in recommendations. For example, Hajri et. al [4] created a recommendation engine to support MOOC learners with complementary OERs. Their study identifies that some learners did not prefer to use external resources to not disperse themselves. This result implies that user may not like to navigate cross-site even though they are provided with an item recommendation very related to their interest. Thus, understanding users' preferences on recommended items is very crucial for maintenance and improvement of recommender engines [2] so that personalised items could be provided to meet each user's preferences [1]. This paper is designed to understand users' cross-site navigation through recommendations on OER and their choice of the item to study.

3 CROSS-SITE RECOMMENDATIONS OF OPEN EDUCATIONAL RESOURCES

In order to provide cross-site and cross-language recommendations, we have developed a content-based recommender system which recommends resources based on the similarity of their content. We propose semantic representation of the resources based on Wikipedia concepts, i.e. Wikipedia pages which were identified and linked to particular parts of the material's content. Wikipedia concepts are then used to represent the resources of different modalities - providing a "concept" overview of the resource's content. This semantic representation allows comparing and calculating the similarity between resources of different languages, thus enabling recommendations containing resources in different languages. The recommender system design is described in [5].

Our recommendations are designed to provide the most similar resources based on the user query. The query can be either a) a link to another resource or b) a free-form text. For this paper, we focus on the first option, where the user provides a link to another resource. The recommendations are ordered by resource's similarity to the provided query, i.e. more similar resources appear higher on the list.

Figure 1: Snapshot of the recommendation window on videolectures.net resource (date: 25.07.2019).

The recommender system has been provided as a service with a public application programming interface (API).[2] In addition, we have also provided a plugin which allows a website to include the list of the OER recommendations given a query. Figure 1 shows an example of the recommender plugin output on the videolectures.net repository[3].

The data about users' transitions between the OER materials and recommendations is stored as csv files (we retrieved it on: 04.07.2019) Please note that we only have the transitions directed from www.VideoLectures.net (VL) and www.upv.es (Universitat Politècnica de València - UPV) due to the data sharing policies. We have implemented learning analytics techniques to analyse and visualise the users' preferences on recommendations by using the Python programming language.

4 ANALYSIS ON USERS' PREFERENCES

In the transitions dataset we have 233,221 transitions showing the users navigating from one page to another through the recommended items. Please note that we only have the transitions directed from VL and UPV due to the data sharing policies. The data used in our experiments was retrieved on July 4th, 2019.

The users are usually provided with around 20 items in the recommendation list. The data shows that the users tend to choose the item ranked 8.89. Usually, only 5 to 7 items

---

[2]Documentation is available at https://platform.x5gon.org/documentation
[3]http://videolectures.net/

Table 1: Frequency of Navigation amongst OER repositories

Directed from	Directed to	Frequency
VL	VL	176,594 (76%)
VL	UPV	14,212 (6%)
VL	UOS	553 (0.2%)
VL	Nantes	8 (0.0034%)
VL	MIT	8,854 (3.8%)
VL	Bologna	32,882 (14%)
UPV	VL	14 (0.006%)
UPV	UPV	92 (0.04%)
UPV	UOS	0 (0%)
UPV	Nantes	0 (0%)
UPV	MIT	0 (0%)
UPV	Bologna	12 (0.005%)

could fit into the recommendation window, which means that the users tend to scroll down in the recommendation window rather than click on the first recommended item.

According to the statistics, the users have chosen an item from the first page 88914 times (38%) and have scrolled down to chose an item 144,291 times (62%) in the case that 6 items shown at once in the recommendation window (see Fig.1).

Navigation amongst OER sites

Since the recommendations are cross-site, it is possible for the users to move from one OER repository to another. Because of the data sharing policy among the providers, we can track the transitions from VL and UPV to any partner providers. The sankey diagram in Figure 2 shows the navigation amongst the OER repositories. Table 1 shows the exact number how many times a user is directed from one repository to another.

Apart from the users' decision to choose an item from different domains, the number of items recommended by domain could have an effect on the users' choices. Figure 3 shows the percentage of recommended items by domains on VideoLectures.net and UPV, respectively. We can see that the providers mostly recommend an item from their own domain.

The results can be summarised as follows:

- When a user is on a material, most probably they choose the next material from the same domain, indicating that they prefer to stay on the same website.
- The users have mostly chosen the next item from the VL, Bologna, UPV and MIT respectively.
- All transitions to Nantes, UOS and MIT were directed from VL.

- There is at least one transition from the VL to each of the OER repository listed while there are no transitions between UPV and UOS, Nantes, and MIT. The reason could be that the content of the resources are not similar on these particular repositories so they are rarely shown to the users.

Chosen vs. Not Chosen

Even though there surely are reasons for a user not to choose the first ranked item in the recommendation list which cannot be traced in the data, we hypothesise that the users would choose one of the first shown items since they are more similar to the viewed item based on their contents. Therefore, observing the trends with the chosen item by investigating its similarity to the other given recommendations would give us insight into the users' behaviours.

A heat map is a convenient tool to visually show the trends. Figure 4 compares the features of the first item in the recommendation lists and the chosen item by the user.

The features are chosen for comparison are explained below.

Author. Author of the material

Language. Spoken language in the video or the provided language in the text

Provider. The website domain of the material provider e.g. videolectures.net

Type. Type of the material e.g. pdf or mp4

Wikipedia Concepts. The Wikipedia concepts linked to the resource's content. These were acquired through the Wikifier [3], a web service that finds and links text elements to Wikipedia concepts.

The items ranked first in the list and selected by the users are not included here. In the graph, the more purplish the more different, the more greenish the more similar.

According to the graph, the author information is the least similar feature between the selected item and the first item shown in the recommendation list. This implies that the users rarely chose the item authored by the same author of the first item in the list. It should be noted here that one of the reasons for this could be the lack of materials authored by the same author. It is also observed that language, provider and type are not necessarily the same.

The first ranked item is contentwise the most similar item to the currently viewed item (see Section 3). However, the items chosen by users are not always the most similar items. This result implies that the currently implemented content-based filtering method is not enough to meet the users' preferences, and other aspects must be considered.

5 CONCLUSION

The presented research is designed to investigate users' navigation between the different open educational resource
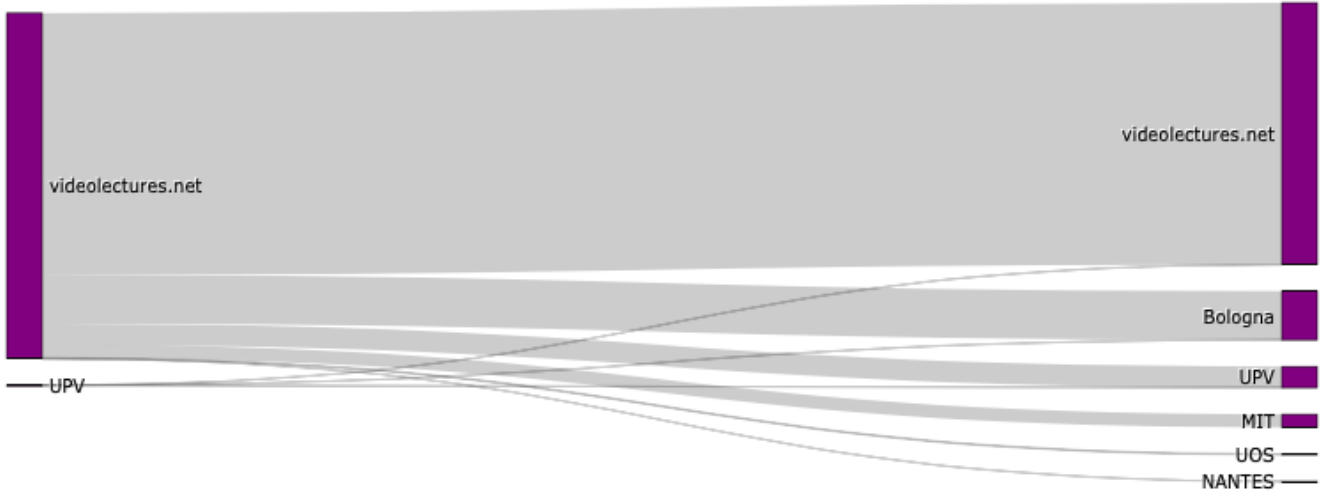
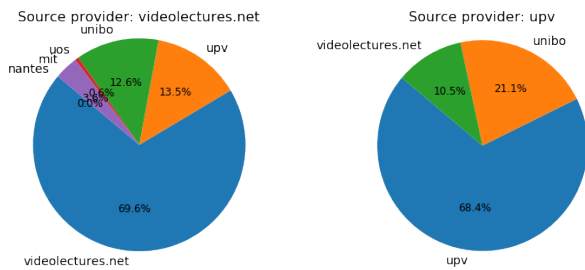**Figure 2: Navigation amongst the OER providers**

Figure 3: Number of materials recommended by domain

Figure 4: Comparison of the selected item and the first ranked item in the recommendation list

providers through an embedded recommendation engine on their platforms. We observe that the users mostly chose to stay within the same domain provider. Interestingly, we observed that the users did not choose the first couple of items that are ranked higher in the recommendation list, but they rather chose items ranked at around 8th place. This result shed light on users' preferences on cross-site OERs but also pave a way to further research to i) deeper behavioural analysis on user preferences and ii) improve the recommender engine which not only implement a content-based filtering but a method which is modified with personalised attributes.

6 ACKNOWLEDGMENTS

REFERENCES

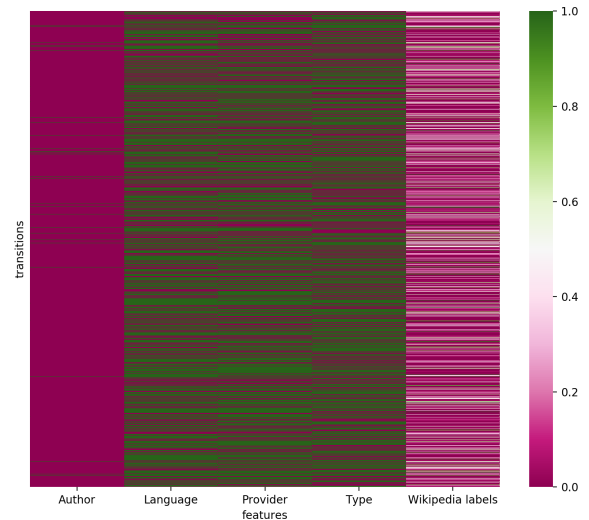[1] Eugene Agichtein, Eric Brill, and Susan Dumais. 2006. Improving web search ranking by incorporating user behavior information. In *In 29th Int'l ACM SIGIR Conf on Research and Development in Information Retrieval*. ACM, 19–26.

[2] Eugene Agichtein, Eric Brill, Susan Dumais, and Robert Ragno. [n.d.]. Learning user interaction models for predicting web search result preferences.

[3] Janez Brank, Gregor Leban, and Marko Grobelnik. 2017. Annotating documents with relevant Wikipedia concepts. *In SiKDD*.

[4] Hiba Hajri, Yolaine Bourda, and Fabrice Popineau. 2018. Personalized Recommendation of Open Educational Resources in MOOCs. In *International Conference on Computer Supported Education*. Springer, 166–190.

[5] Erik Novak, Jasna Urbančič, and Miha Jenko. 2018. Preparing multimodal data for natural language processing. *In SiKDD*.