Evidence of Cross-boundary Use and Reuse of Digital Educational Resources

doi:10.3991/ijxx.vxnx.xxx (Please do not delete this line)

Riina Vuorikari\textsuperscript{1,2} and Rob Koper\textsuperscript{1}

\textsuperscript{1} OUNL, Heerlen, The Netherlands
\textsuperscript{2} European Schoolnet, Brussels, Belgium

Abstract— In this study an investigation using log-files of teachers' Collections of educational resources in two different platform was conducted. The goal was to find empirical evidence from the field that teachers use and reuse learning resources that are in a language other than their mother tongue and originate from different countries than they do, for this, the term cross-boundary use of learning resources is used. In both contexts behavioral evidence was found that cross-boundary use and reuse takes place, and it was shown that it correlates with the general use and reuse trends. Moreover, it was found that cross-boundary reuse, when compared to 20\% of general reuse, was notably less (37\% to 55\% of it). The motivation to study cross-boundary use and reuse is to set a baseline for future studies, and to understand how it can be supported and enhanced in the future.

Index Terms— behavioral evidence, cross-boundary use, learning resources, multilinguality, reuse

I. INTRODUCTION

Since the late 1990’s, digital repositories for learning purposes have gained ground. Such repositories store “any digital resources that can be reused to support learning” [1] and/or their respective metadata. Repositories have been set up on regional, national and international levels to offer content for teachers and learners from K-12 to tertiary and vocational education [2,3]. Sharing, using and reusing digital educational resources are the main drivers of the learning object economy [4]. Participants of the economy are educational institutions, digital libraries and learning object repositories (LOR) and their policy makers, managers, content providers, educators and learners, each with their own needs, requirements and agendas.

Fig. 1 depicts learning resource landscape: the vertical axis is used to present how the authoring of the content takes place. On the one end, there are any third party sources of educational resources such as textbook publishers, museums, broadcasting companies, etc. On the other end of the axis, the content comes from teachers who create it, i.e. the “author users”. The horizontal axis introduces the context: “institutional context” refers to content repositories that are managed by educational institutions and authorities (e.g. Ministry of Education), whereas the other end of the axis represents “community driven context” where publishing typically takes advantage of Web 2.0 tools.

This paper focuses on the use and reuse of educational resources when it happens across language and country borders, e.g., between users and communities that do not share the same mother tongue and/or the same country. This is called cross-boundary use of educational resources and it means that the user and the content come from different countries, and/or that the content is in a language other than the user’s mother tongue.

![Image](Please do not delete this line)

Figure 1. The learning resources landscape representing the diversity of available resources. EdReNe [5] members, for example, run learning repositories by European educational authorities

The evidence finding focuses on teachers in K-12 education in a European multilingual context where 497 million people live with diverse linguistic backgrounds. Multilinguality can be defined as a situation where several languages are spoken within a certain geographical area, as well as the ability of a person to master multiple languages. Multilinguality has an important role in the European Union (EU); there are 23 official EU languages, three alphabets, and some 60 other languages are used commonly [6]. 56\% of EU citizens say that they are able to hold a conversation in one language apart from their mother tongue, and 28\% in at least two languages. English remains the most widely spoken foreign language throughout Europe (38\%), second and third place are French (14\%) and German (14\%), whereas 6\% have foreign language expertise in Spanish and Russian, respectively. Over two-thirds say that they learned foreign languages at school [7].

This paper starts by identifying how much cross-boundary use and reuse currently take place in order to have a proper baseline for future studies. Moreover, the interest is to better understand what variables affect on the eventual probability of discovery, and the use and reuse of cross-boundary resources. First, the related work is introduced. Then, the research method, data collection and
analyses procedures are described, after which the results are outlined. Following that, the outcomes and their implications for future studies are discussed from the perspective to enhance the discovery and eventual use and reuse of cross-boundary educational content.

II. RELATED WORK

A number of studies has focused on the use and reuse of digital educational resources in different settings from blended learning in a classroom to fully functional distance education at the university level [e.g., 8,9,10]. In some cases learning resources are used to complement social interaction in learning contexts, whereas in other settings they can be used to imitate or replace social interactions, as can be observed in Technology Enhanced Learning in general [11]. The adoption and use of digital resources vary between different educational settings in a way such that each have their specific requirements; unlike the university sector, standardized local or national curricula are common in schools and colleges which can affect how educators view the reuse [12]. Additionally, informal learning communities have their specific requirements, too [13].

The use and reuse of learning resources can be interpreted as the result of the success of a chain of consecutive events where each event needs to succeed for the use or reuse to take place [14,15]. In learning object repositories (LOR), these are related to the lifecycle of Learning Objects, the main steps of which are obtaining, labeling, offering, selecting, using and retaining [16]. Depending on the context the steps differ, however, the important ones are the discovery of the resource, evaluating its usefulness for the given context, accessing it, and adapting and integrating it into a new context.

Reference [17] defines three levels of reuse of learning resources: the creator of the resource reuses it (first level), the user reuses a resource created by someone else within the same community (second level) and the third level of reuse when the resource reused is created by someone else from outside of the user’s community. Reference [15] conducted a quantitative analysis of learning object reuse and observes that the reuse is around 20% across resources of different granularity of content. The following criteria of reuse is applied: “considered reuse if a component is present in more than one slide, if a module is used in more than one collection, or if a course is mandatory in more than one curriculum” [15, p.66]. Reference [18] also studied reuse of learning resources in a repository, where a similar finding was reported.

Reference [19] studied the availability of content on the Internet and how this content is accessed in users’ first and second languages. The results indicate that non-English languages are under-represented on the Web and that this is partly due to content-creation, link-setting and users link-following behavior. Thus, making educational content available in the users’ mother tongue has been the goal of institutionalized learning resource repositories that have been set up by national or regional educational authorities in Europe [5]. Main efforts also include the labeling of educational resources for indexing and search purposes by using standardized metadata, e.g., [20] and Application Profiles with multilingual vocabularies and thesauri [21]. Despite these efforts, in a context of crossing national, cultural and language boundaries, locating suitable content has proven challenging [22], as the gap between the end-user and expert vocabularies remains wide [8, p. 53].

Turning the emphasis away from technical issues, reference [10] studied the reuse behaviors of open educational resources and found that language translations represent only 1% of cases (p.110). Seen from this light, Littlejohn’s question [23, p.5] “Is global sharing of resources a possibility?” seems relevant. The challenges’ list is long: problems not only within disciplines that “differ in their languages, in their methods of enquiry and in their social and cultural organization” [24]”, but also at a transnational level, where “cultural and language differences add a further complexity to the idea of resource sharing”, e.g. the concern of the fit to the local curriculum [8]; diverse models of teaching and related cultural expectations, as well as types of tasks for which learning resources are used [25].

It has been suggested that studying sharing and the reuse cannot only include the dimensions of the repository and individuals, but other dimensions influence on users’ decisions. These can depend on issues such as the subject discipline, who contributes to the repository, its reward and incentive schemes as well as pedagogical approaches [26]. Towards this direction, Community key dimensions framework has been introduced which includes dimensions such as “community purpose”, “dialogue”, and “composition” [25]. Emphasising the importance of the community around the reuse of learning resources, reference [18] reported that when there was at least one person in common with both the module (i.e. learning resource) and the collection, the modules were included in collections 3.67 times more often. Similar preferential behaviour is found in other areas, for example reference [27] reports on American fund managers investing more money in firms run by people who were known via shared education networks.

III. METHOD

For this study, two learning resource platforms were selected. Table 1 presents the datasets for Calibrate [28] and LeMill [29].

<table>
<thead>
<tr>
<th>TABLE I. DESCRIPTION OF DIFFERENT SETS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibrate</td>
</tr>
<tr>
<td>Users in this study</td>
</tr>
<tr>
<td>Users in the system</td>
</tr>
<tr>
<td>Users in the study</td>
</tr>
<tr>
<td>Resources in system</td>
</tr>
<tr>
<td>General use</td>
</tr>
<tr>
<td>General, 2nd level reuse</td>
</tr>
<tr>
<td>Coverage</td>
</tr>
<tr>
<td>Average size of a Collection</td>
</tr>
<tr>
<td>Date of dataset</td>
</tr>
</tbody>
</table>

The Calibrate portal (currently known as Learning Resource Exchange for Schools) represents institutional context, it federates content from a number of European
Schoolnet partners [30], whereas in LeMill, teachers create resources in a community-driven context. The systems and their tagging tools have been described in [31].

In this study, bookmarks and users’ Collections of learning resources are considered as a proxy for the use and reuse of resources. Reference [32] has identified implicit and explicit Interest indicators in the context of resource discovery on the Internet, they can be described as a way to create a digital handle to a resource so that it can be later retrieved again (e.g., bookmark, tag, addition to a Collection). Such digital handle was used to gather data for this study. Namely, log-files on teachers’ bookmarks and Collections of educational resources were gathered from both platforms. They consisted of the following data: the user (user id, country of origin, languages spoken); the resource (resource id, title, URL and some other LOM metadata); and the Interest indicator (e.g. an id that indicated the content is part of a Collection or an id of the bookmark with tags).

When a learning resource has an Interest indicator at least once, “General use” is considered, even if no further evidence is gathered about its use in teaching or learning. Similar measures were taken to study cross-boundary use. A classical example is represented in Fig. 2, the resource with an Interest indicator originates from a different country than the user and is in different language than that of the user’s mother tongue.

![Figure 2. Cross-boundary resource](image)

While processing the data other cross-boundary cases were found, namely that the resource with an Interest indicator is in a user’s mother tongue, but the user and resource come from different countries, e.g., Austria and Germany (Fig. 3).

![Figure 3. Cross-country, but not a cross-language resource](image)

Lastly, the user and the resource come from the same country, but the content is in a different language, for example in English (Fig. 4), which was often the case for the content in LeMill. The prominence of English language in the reuse setting is discussed at a later point.

![Figure 4. Cross-language, but not a cross-country resource](image)

The reuse is considered taking place on different levels using the vocabulary from [17]. The first level reuse by the creator of the resource is not considered in this study. The second level of reuse takes place within the repository, users of which are considered as a meta-community. When the resource is integrated in a new context with other components of the repository, and when this occurred more than once, this is considered “General reuse” or second level of reuse.

As the focus is on languages and national boundaries, they are also used to define communities. When the resource reused is created by someone else from outside of the user’s language or country-related community, it is considered the third level of reuse. For example, the following two cases can be identified for language related cross-boundary reuse:

a. A Finnish teacher with competences in English and French is not only part of the meta-community of the repository, but also part of the language-based sub-communities like Finnish, English and French. The reuse takes place on this level when she reuses resources in English or French.

b. A Finnish teacher who reuses a resource that is in a language outside of her language competencies (e.g., Hungarian), but still within the same disciplinary or pedagogical culture, is also considered the third level of reuse.

For each data set, a number of measures, which are presented in Table 2, are established. The general use is counted by how many times the resource appears in Collections and then compared with the total number of resources in that given set. The reuse is counted using the number of times that the resource appeared in more than one Collection, and then compared with the total number of resources in that given set. This gives us a comparable figure to the reuse of about 20% as in [15]. The same measures are used for cross-boundary usage.

<table>
<thead>
<tr>
<th>Name</th>
<th>Metric</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>General use</td>
<td>Number of resource integrated in a new context at least once</td>
<td>Number of resources in Collections at least once/Number of all resources in the set</td>
</tr>
<tr>
<td>General cross-boundary use</td>
<td>Same as above, but using cross-boundary measures</td>
<td>Same as above, but using cross-boundary measures</td>
</tr>
<tr>
<td>General reuse (2nd level)</td>
<td>Number of resource integrated in a new context more than once</td>
<td>Number of resource in more than one Collection/Number of all resources in the set</td>
</tr>
<tr>
<td>Cross-boundary reuse (3rd level)</td>
<td>Number of cross-boundary cases integrated in a new context more than once</td>
<td>Same as “General reuse”, but only for cross-boundary usage</td>
</tr>
<tr>
<td>Coverage</td>
<td>Degree to which the used/reused resources cover the entire set of items within a system</td>
<td>Distinct number resources used/reused - Total of distinct resources in the system</td>
</tr>
</tbody>
</table>

In addition to cross-boundary use, another metric was used. In the recommender system literature, Coverage measures the degree to which the recommendations cover the entire set of items [33]. In this context, Coverage measures the degree to which the used or reused resources cover the entire set of items within a system. It is calculated using the distinct number of resources (i.e. individual resources) and subtracting it from the total of distinct resources in the system.
IV. RESULTS ON USE AND REUSE ACROSS BOUNDARIES

A. LeMill

In LeMill, only about 10% of all the users have used the Collection-feature. In this study, Collection-feature is used as a proxy for defining the use and reuse of resources. The general use in LeMill is 82% and the general reuse is 21% (Table 3). As for the cross-boundary usage: 29% of resources in the entire LeMill are used across boundaries, whereas the figure for the cross-boundary reuse is 12%. The latter amounts to about 5% of distinct resources in LeMill.

When the cross-boundary usage is studied, it is found that most of it (64%) takes place across language borders, i.e., resources have been produced in the same country as the user, but are in another language (Fig. 4). For example, 51% of cross-language content was used in English, but produced by non-native English speaker in the same country. The other major languages for cross-language use are Russian (21%) and Estonian (12%). The cross-boundary use both across languages and countries was 35% (Fig. 2).

In LeMill, it was possible to look how users used and reused resources of different types in English and in other available languages. Out of all resources added to Collections at least once (n=1649), multimedia material was the most used resource type, out of which 55% was used in English. The reuse in LeMill is not distributed evenly across different content types: the multimedia material is by far the more reused type. When it comes to cross-boundary use and reuse, 35% of the time resources added in Collections were cross-boundary resources. Again, multimedia material was by far more used (17%) and reused across boundaries (8%), and in this case the material was mostly in English.

B. Calibrate

In Calibrate, 30% of all users have used the Collections-feature that is used as a proxy for use and reuse (Table 4). 14% of general use is found across all 4,000 learning resources (*) and 7,000 assets (i.e. resources with lower level granularity such as individual photos). The general reuse figure is 7%. The cross-boundary use and reuse follows very closely the general use and reuse. It can thus be observed that the use and reuse in Calibrate takes place mostly across language and country borders.

A manual verification of URLs to infer the file format was performed to better understand the type and granularity of the resource. This selection method gave returns of 60% of the used resources. 13% of the URLs indicated file formats such as images, videos and flash files, which usually cannot be disaggregated to smaller granularity. 87% were file format like .htm, .php, zip-files, .pdf,.exe from which it can be inferred that these are more likely aggregated learning resources. As for the reused resources, it was found that 95% of analyzed URLs represented the latter category. Based on this data, an assumption was made that most reused resources exclude small granularity resources. Thus, the general reuse for Calibrate can be calculated to be 18.6%(*). This result, again, very clearly indicates that resources of different granularity were reused differently.

C. General use and reuse vs. cross-boundary use and reuse

The use and reuse have now been studied among the general level (2nd level), and to what extent does it take place across language and country boundaries (3rd level). As LeMill and Calibrate datasets and user-base are similar, the correlation between the second level and third level of use and reuse of resources was calculated by using the measures in Table 2. The Pearson R correlation between the use and reuse on the second level and the use and reuse on the 3rd level was performed. A strong and significant positive correlation r = .928 was observed (p<0.01).

V. DISCUSSION AND FUTURE WORK

Two different sources of digital educational resources coming from different contexts and authoring backgrounds were studied; Calibrate with institutional content and LeMill with teacher-generated content. Despite the social, cultural and technical barriers that were reviewed in the Related work-Section, evidence was found that the use and reuse of educational resources take place across language and country boundaries (i.e. 3rd level of reuse).

In LeMill, where the content is created by teachers, it was found that the users create and share material both within the language communities and across them, indicating that the purpose of the platform fits and supports the typical activities that the meta-community carries out in order to achieve its goals. When the cross-boundary use is studied, it was found that most of it takes place across languages, e.g. content is produced in English by a non-native speaker in the same country where the use takes place (Fig. 4). A similar situation is becoming more common especially in Europe, where educational institutions and individuals who aim at attracting a wider cross-boundary audience for their content, make resources available in widely spoken foreign languages. The Calibrate portal, on the other hand, operates in an institutional context where the focus is not on the creation of material, but on the discovery and eventual use and reuse. The observed use of resources was generally less than in LeMill, however, the resources were used as much within countries (2nd level) as across-boundaries (3rd level).
The statement that in common settings, the amount of learning resources reuse is around 20% across collections of different granularity [15] was tested. The second level of reuse in LeMill (22%) follows the general trend, however, the reuse is not evenly distributed across the different types. In Calibrate, the second level of reuse when looked across all the items was well below the baseline (7%). Excluding the small granularity content in Calibrate (the data from Table 4), the reuse at 19% was observed. This finding also hints that not all the content with different granularities is equally reused. The figures for the cross-boundary (i.e. 3rd level) reuse are 12% in LeMill and 7% in Calibrate. The cross-boundary reuse is 55% to 37% of the general reuse (respectively).

Additionally, to better understand the extent to which the resources are used, the metric of Coverage was used. The Coverage in LeMill is 69%, whereas in Calibrate this is low, 9% of all items. Regarding the cross-boundary coverage, the figures are 21% and 9% respectively. The “inconvenient truth” reveals that the coverage is rather low and better ways to support both 2nd and 3rd level reuse should be created.

Literatures in the Related work-Section suggests that the use and reuse of resources cannot be solely studied by looking at the figures from behavioral data, as those are not an indicator of attitudes and preferences. Reference [19] cautions against drawing too simplistic conclusions based on behavior alone: in the absence of links and/or content in their native languages, users will acquiesce to English-language content, however, their preference will persist. In user-group sessions (documented in [34]) many of the barriers discussed in the Related work-section have re-emerged, however, many positive usages of cross-boundary resources have been observed. Supporting such behavior, the sessions highlighted that sharing cross-boundary resources alone is not sufficient, but sharing usage scenarios and stories along with the resource is important for teachers.

Using Community key dimensions framework [25] and supporting elements such as “dialogue”, i.e. different modes of participation and communication, for example through tagging, evaluations of resources and usage scenarios can thus offer ways to engage in more use and reuse on both levels, a field that offers interesting future studies. Additionally, future implementations and studies focusing on how different types of sub-communities (“composition”) can support the use and reuse on different levels should offer interesting insights also for the cross-boundary discovery of resources and their eventual use. Such sub-communities could be based on disciplines, spoken languages, user-behavior (downloads, bookmarks, tags and ratings), to mention but a few. Such Community key dimensions could also help establishing not only more non-English content on the Web, but also enhance creating new ways for users’ link-following behavior.

Moreover, establishing better metrics for cross-boundary use and reuse of learning resources is needed to allow a better quantitative and periodical measuring of the cross-boundary actions. This is a logical extension of already existing Learnmetrics [15], Web metrics for digital libraries [35] and metrics evaluating tagging behavior in social bookmarking systems [36]. Having established a baseline for cross-boundary use and reuse in this study can be considered as a contribution towards such metrics, as well as the work in [37] where metrics were proposed for tagging in a multilingual context.

Lastly, as a limitation of this study it can be pointed out that using Interest indicators as a proxy for the use and reuse of learning resources can be misleading, as there is no further evidence on their use in teaching related activities. The real figure of cross-boundary reuse can be bigger (or smaller).

VI. CONCLUSIONS

An investigation using log-files and social bookmarks on teachers’ Collections of educational resources as a proxy for the use and reuse of digital educational resources was conducted. The goal was to find out how much cross-boundary use and reuse currently take place by comparing the origin and languages of the user to the origin and languages of the learning resources. The motivation for setting a baseline for such use and reuse is to better understand how it could be supported and enhanced in the future.

In both sources behavioral evidence was found that cross-boundary use and reuse takes place and that it correlates with the general use and reuse trends. The cross-boundary reuse within a platform (third level of reuse), when compared to 20% of general reuse, was notably less (37% to 55% of it). Following the idea that improving even one of the steps in the reuse chain would improve the probability of reuse and therefore, the amount of reuse within the platform, it can be suggested that better ways to support and enhance cross-boundary use should be the focus of future studies.

ACKNOWLEDGMENT

The authors would like to thank the people involved in LeMill and Calibrate for sharing the data for these analyses.

REFERENCES


AUTHORS

R. H. Vuorikari is a PhD candidate in CELSTEC, OUNL and works in European Schoolnet in Brussels, Belgium (e-mail: riina.vuorikari@eun.org).

E. J. R. Koper is Professor and Dean in CELSTEC, OUNL in Heerlen, the Neterlands. (e-mail: rob.koper@ounl.nl).

Manuscript received 09 October 2005. (Write the date on which you submitted your paper for review.) This work was supported in part by the U.S. Department of Commerce under Grant BSI23456 (sponsor and financial support acknowledgment goes here). Published as submitted by the author(s).