Exploratory Analysis of the Main Characteristics of Tags and Tagging of Educational Resources in a Multi-lingual Context

Riina Vuorikari
European Schoolnet
Riina.Vuorikari@eun.org

Xavier Ochoa
Escuela Superior Politecnica del Litoral
xavier@cti.espol.edu.ec

Abstract

Although social, collaborative classification through tagging has been the focus of recent research, the effect of multi-linguality is often overlooked. This work presents an exploratory study of the production and use of tags in multiple languages in a context of European Learning Resources Exchange. We describe a tagging tool used by teachers from 6 countries and study the main characteristics of tags and how users tag when multiple languages are presented. We find early indication that tags and bookmarks could be used to facilitate the discovery of educational resources across country and language borders. "Hiding all but the right tags" becomes crucial for the success of a multi-lingual collaborative tagging system.

1. Introduction

The use of social, collaborative classification systems has grown dramatically in recent years. An example of this is a multitude of sites that provide some type of social annotation of digital artefacts and social navigation system (Flickr, del.icio.us, CiteULike, Last.fm, among others). Social tagging, i.e. allowing individuals to apply free text keywords to digital objects, potentially offers advantages in terms of personal knowledge management, serendipitous access to objects through tags, and enhanced possibilities to share content with emerging social networks among other users. In the core of the tagging system, there are the implicit and/or explicit relationships between resources through the users that tag them; similarly, users are connected by the resources they bookmark and tag (Marlow *et al.* 2006).

Several studies have been undertaken to better understand the behaviour and evolution of social tagging systems. Early research conducted by Mathes

([2004](#)) coined the term "folksonomy" to be used for the emerging socially generated vocabulary that he compared with more formal ontologies. Golder and Huberman ([2006](#)) first looked at user patterns of collaborative tagging systems. Recent studies also focus on understanding the network properties ([Gatutto et al. 2007](#)).

A prevailing aspect among current studies concerning tagging is that they assume that tags are represented in a common language ([Hammond et al. 2005](#)) understandable by all the members of the user community. Guy and Tonkin ([2006](#)) suggest that this is not always the case; they found that the bulk of tags in their study was valid English. However, tags from other languages were present in small numbers. They acknowledge that gauging the source language of tags is challenging due to technical issues as well as linguistics (e.g. many words exist in multiple languages with differing meaning or grammatical structures). The most difficult aspect outlined in that study was "malformed" tags, which put them beyond the grasp of a multi-lingual spell-checker. Lately, multi-lingual tags have started to emerge on popular social tagging systems as their user base grows at the international level. Roughly, two different ways to process multiple languages can be observed: by users and by "system".

Examples of how users deal with multiple languages include [Flickr](#) or [del.icio.us](#) where users share the same system and use multiple languages to tag. Tags are added in different languages (e.g. "achat", "shopping"), and, even on some occasions, a tag identifying the source language has been added (e.g. lang:fi). This is very marginal, though, we found less than 18 000 such tags applied in del.icio.us (accessed in July 2008), which has more than 10 000 000 tags. There is no system level support that allows users to see tags, say, only in French or Finnish. In [LibraryThing](#), which recently has launched different language versions of the service, experienced users can also combine tags under one tag. On some occasions, tags in different languages have been grouped together. As for the community of [Flickr](#), its tag base has become a source for cross-language retrieval studies by [iCLEF](#). On the other hand, approaches like [Yahoo!'s MyWeb](#) offer tags and tagclouds in different languages in localised parts of the portal (e.g. .fr, .es, ...), which indicates that there is some system level support for multiple languages. An outcome of this is that users from different countries and language groups are kept separated.

Our work, still at its early stage, attempts to shed light on a community of users who use a common tagging system across country and language borders, but does not share a common language. One of our main questions is to study whether a tagging system, where users tag in multiple languages, still functions as one system, or is it split into separate communities of users based on their languages? This exploration takes place in the context of two

European Community funded projects, Calibrate and Melt, both focusing on sharing and re-using digital learning resources in primary and secondary education.

We start by studying the phenomenon of tagging in multiple languages in Section 3. We first look at it from the system point of view; we observe the general tagging activity and distribution of post, what is the tag growth and reuse in our system. We then turn to look at the users and their tagging behaviour in the multi-lingual context: in what languages do users tag, what are the characteristics of tags, and introduce the idea of "travel well" tags. We also present a user study on how users perceive multi-lingual tags. In Section 4, we attempt to answer our research questions and contribute to the design requirements of a multi-lingual tagging system that helps bridge across languages and country barriers. Lastly, we outline the future work and present a conclusion in Section 7.

2. Research rationale and methodology

In this section we introduce the research terminology, outline our research goals and early hypotheses. We also explain our research methodology, give an overview of our tagging system, and finally also describe our main dataset.

2.1 Research terminology

Marlow *et al*. (2006) present a conceptual model for social tagging system; tags are represented as typed edges connecting users and resources. We study such a tagging system where users from different pilot countries (Austria, Estonia, Hungary, Lithuania, Poland and Slovenia) assign tags to resources that they find from a federation of learning repositories. We are interested in the implicit relationships between resources through the users that tag them. Moreover, we are interested in the connection that users form through the resources they bookmark and tag.

The basic unit of study in this paper thus consists of a (user, resource, {tags}) triple, which Gatutto *et al*. (2007) also described as a "post". Farooq *et al*. (2007) call it a "tag application". Hereafter in this paper we refer to our unit of study as a *post* (Table 1).

Table 1. Unit of study as presented in this paper

When discussing tags in our system, we use the terminology from Farooq *et al*. (2007): global tags (previously used by all users of the system), personal tags (previously used by the user) and paper-specific tags (previously used by

all users of the system for the target paper), which we hereafter refer to as resource-specific tags. Moreover, we use the tag categorisation factual, subjective and personal tags from Sen *et al*. (2006), which is also based on the categories of Golder and Huberman (2006).

2.2 Research goals

The primary goal of our analysis is to explore our dataset to better understand the phenomenon of tagging in the context of multiple languages. We have two main research questions that we wish to advance:

1) What happens when users tag in multiple languages instead of one common language?, and
2) Can we find evidence that tagging and bookmarking through implicit connection between users, resources and tags, could be used to facilitate the cross border use of learning resources?

We start by decomposing the first question: Does the presence of multiple languages have any implication on the global growth of tags in a tagging system and how tags are reused (Section 3.1.2). Second, we need to understand better how users behave in a tagging system where multiple languages are present (Section 3.2.1), i.e. in what languages do users tag and how do they reuse tags? Additionally, we seek to understand how users perceive multi-lingual tags (Section 3.2.3).

As for our second research question, we want to discover early indicators as to whether the implicit connection between users, resources and tags could be used to facilitate cross-border discovery and use of educational resources in the context of multi-lingual and multi-cultural federation of educational repositories. By cross-border we mean users who discover resources that come from different countries than they do, which also can be in different language from their mother tongue. Our context of research is European education, especially that prior to the tertiary level, which is inherently multi-lingual and multi-cultural. Offering educational resources and services in native languages is deemed important, but equally important is the exposure to other languages (COM, 2007). One way to promote this kind of multi-linguality is to make learning resources available across national and linguistic borders.

This complicates semantic interoperability, i.e. how well the content and its metadata can be understood by other systems and users. Controlled vocabularies, such as the multi-lingual LRE Thesaurus (2002), can be used to overcome some semantic interoperability hurdles. However, the gap between the terms created by experts, like in the LRE Thesaurus, and practitioners in the field is problematic (McCormick *et al.* 2004). For that reason, our current

approach looks into the co-existence of taxonomies and end user generated tags.

Our second main research question relates to the value of multi-lingual tagging system, and can be decomposed into the following two parts: on the one hand, we want understand what kind of information multi-lingual tags can yield about the resources and their possible use in different contexts (Section 3.2.2). On the other hand, we are interested in the value of tags for resource discovery and as a navigational tool to enhance the discovery of new resources across country and language borders (discussed in Section 4).

Finally, we wish to contribute to better understanding of system requirements for a tagging tool that supports multiple languages (Section 5). Design heuristics for social bookmarking tools are well covered in Farooq *et al*. (2007), and cross-language retrieval is discussed elsewhere (iCLEF, 2008). Our work focuses on the intersection between these two.

2.3 Research methodology

To attain our research goals, we start with this descriptive, qualitative analysis that uses our server-side logging data, which was gathered in a multi-lingual context from November 2006 to October 2007. We use this analysis as a requirements survey to better understand the user needs and requirements. On the other hand, it also helps provide more information as to which issues to focus on in the future in order to better shape our hypotheses for subsequent correlational, quasi-experimental and experimental studies.

To analyse the tags and tagging behaviour, we manually apply a number of metrics that have been used in previous studies, notably those from Farooq *et al*. (2007) and Sen *et al*. (2006). We offer observations based on log-file analyses on user tagging behaviour.

Finally, our methodology also includes a user study with 13 participants which we summarise in Section 4. Details of this study are discussed elsewhere in Vuorikari *et al*. (2007). Our aim is to gain a better understanding of how users react when they are confronted with tags in multiple languages, especially in those languages that they did not speak or have knowledge of. The results of this user study are useful to guide design decisions in the development of retrieval tools for learning resources in a multi-lingual environment.

2.4 System set-up and dataset

Since November 2006, a group of pilot teachers in Austria, Estonia, Hungary, Lithuania, Poland and Slovenia had access to a portal which was made

available within the Calibrate project. One of the main goals of the project was to facilitate the reuse of learning resources among primary and secondary schools in Europe and beyond. The Calibrate portal is connected to a federation of learning resource repositories (Colin and Massart 2006). Approximately 4000 learning resources and nearly 7000 learning assets (e.g. images, sound) were provided by the Ministries of Education in Austria, Estonia, Hungary, Lithuania, Poland and Slovenia for pilot school teachers to use.

The pilot teachers were asked to use the Calibrate portal from November 2006 to October 2007 to search for useful educational resources among those made available by the participating Ministries of Education. The pilot group was asked to use the available search modes such as browsing resources by topic category, as well as simple and advanced search options. They were asked to produce lesson plans in which they describe the learning resources and how they used them in their teaching.

One of the tools to facilitate this work is called the Favourites. It allows teachers to create personal collections of resources and assign tags to them in any desired language(s). The Favourites-tool creates a unique handle to a resource that is available through the Calibrate portal, so that the user can easily retrieve it again.

Figure 1. Viewable-tagging interface. The user has found a resource "Comparison in action" and adds tags. She is shown all her personal tags, and additionally one resource-related tag from other users who tagged in English

The Calibrate portal was made available in all the languages of the pilot (language choices seen on top right corner of Figure 1) and the tagging interface was always in the language that the user had selected. Figure 1 shows the Favourites-tool and its tagging interface in English. The user is about to add tags to a resource named "Comparison in action". The personal tags of the user are displayed below the text field for tags with a number in parenthesis that indicates how many times it has already been applied. The user can choose a tag by clicking on it or by typing in a new one into the empty text box. When the user now adds a new tag while using the English interface, the tag will automatically be assigned "English" as metadata regarding its language. Tags are to be separated with the use of comma, otherwise they appear as compound terms.

The tagging interface additionally supports viewable tagging whenever resource-specific tags are available in the language of the interface. In this

case (Figure 1) the user is shown the tag "adjectives (1)" in English because the interface language is in English. No tags in any other languages are exposed, even if they exist. Additionally, users could add comments to the resource that they tag. These comments can be made public or kept private, but they are out of the scope of this study.

At the beginning of the pilot the system had no tags attached to resources, thus users were left to invent their own tags. No incentives were given to users to add tags, other than the fact that the tags would help the user to retrieve these resources later.

Table 2 presents part of the data that the Calibrate system logs regarding the users information, resources and tags. This data was used for these analyses. Vuorikari and Van Assche (2007) introduce additional information about the multi-lingual enrichment environment.

Table 2. Metadata regarding the unit of study, like LOM based on the LRE Application profile

Finally, to conclude on our system set-up, it is worth noting that the Favourites bookmarking and tagging tool used in this pilot differs from some other well-known services on the Internet in terms of offering very little social features or support. The bookmarks were not shared among users (this was planned for future development), and users were not able to take advantage of navigational cues such as how many other users have bookmarked resources, tagclouds, etc.

2.4.1 Dataset

A total of 478 users were registered to the Calibrate portal during the time of the pilot. However, only 142 of them had made at least one post (there was no obligation to use the tagging tool). Our dataset is comprised of the users who made at least one post, which represents 30% of pilot participants. It is out of the scope of this study to find out why the remaining 70% were not interested in the tagging tool. This study does not include any data on the use of tags for resources discovery.

Table 3. Description of the dataset

The data for this analysis is from a period of twelve months, November 1 2006 to October 31 2007 (Table 3). However, a number of posts (16) before the initial start were recorded, and we kept them as part of the dataset. Our

dataset is comprised of 1022 posts, covering 682 individual learning resources. There were 1301 individual, distinct tags, however, users had deleted some, resulting in 832 individual multi-lingual tags in the system. We also analysed the deleted tags to gain more insight into the tagging behaviour.

2.4.2 Validity

We analyse the tags and the tagging behaviour of a pilot group of teachers who participate in the Calibrate project . The implication of the data being gathered from a closed pilot, with a rather small sample size, is that the outcomes of this analysis cannot be generalised in a straightforward way to any web-based tagging system. The results will be valuable, however, to define better system design criteria for a tagging tool that should support the use of multiple languages (Section 5).

3. Outcomes of the analysis

In this section we present the main results of our analyses. We first look at it from the system point of view and in the second part the view is shifted on the tagging behaviour in a multi-lingual context: in what languages do users tag and what are the characteristics of tags. We also introduce a summary of a user study on how users perceive multi-lingual tags.

3.1 Observations on the tagging system

To analyse the tags in the tagging system and to better understand the phenomenon, we analyse the general tagging activity, look at the tag growth and the tag reuse both on the global and personal level.

3.1.1 General tagging activity and distribution of posts

The general tagging activity over time is presented in Figure 2. The low number of posts in the summer months can be explained by the holiday period, and more intense activity in February and October by users performing their pilot activities as explained in 2.4.

Figure 2. Number of posts by month

Figure 3 represents the distribution of posts per user (grey points). The graph is presented in logarithmic or log-log scale. As in some other systems (e.g. CiteULike), we find that most posts were generated by a small group of "super users": the top users had 54 and 53 posts respectively. On the average each user had 7.2 posts (median 3 posts per user). The wide distribution

(dotted line) can be better illustrated by an inverse power law (an exponent of -0.78) with an exponential cut-off (with a rate of 0.062). This distribution suggests that highly productive users are very rare; nonetheless they provide most of the tags in the system.

Figure 3. Distribution of posts in the Calibrate system

The average number of posts per resource is 1.5 posts (median 1). Again, there are a small number of resources with many users (the maximum is 9 posts), whereas about 73% of resources had only one user who had added tags to them. This is lower than for example in CiteULike, as reported by Farooq *et al*. (2007).

Finally, the correlation between the number of individual resources and the number of tags that users had applied to them is 0.863, somewhat lower that that in CiteULike (0.944). Farooq *et al* (2007) explain that in their case the strong linear relationship between the number of resources bookmarked and the number of tags for each user can explain that the system is still maturing and has not yet reached its relatively stable stage. We can speculate that this is also the case in Calibrate.

When we look at the coverage of learning resources that have tags applied to them in the Calibrate portal, we find that only 6.2% of all resources available through the federation have tags applied to them.

3.1.2 Tag growth and reuse

The Growth metric by Farooq *et al*. (2007) measures how the tags are evolving over time, at what rate the new ones are created and whether there are signs of the vocabulary stabilising. Creation of new tags in our system has closely followed the number of posts that the users have entered in the system (Figure 4, pink line).

Figure 4. Growth in absolute numbers per month and reuse of tags

Reuse relates to how tags are shared among users; whether tags converge over time or if users only reuse their own personal tags over and over again. We used this metric for both global tags in the system and personal tags. Moreover, in the future we are also interested in using it for calculating the reuse of resource-specific tags. We calculated the tag reuse using the following formula by Sen et al. (2006) for which their baseline was 1.10 users per tag.

Tag reuse= ∑ (# of distinct users for each tag)/ # of tags.

The reuse of tags on the global level was very low, 1.22 users/tag. It was also rather low in CiteULike (1.59users/tag). We further followed the metrics from the CiteULike analysis and calculated the number of occurrences of tag reuse for each tag (number of posts per tag minus 1). Our average (3.2) was even lower than that of CiteULike (3.9).

Table 4 lists the twenty most reused tags in the pilot. We give the tag name, its language, number of times it was reused and the number of users. Additionally, we name the category of tags, which will be explained in 3.2.2.

Some of the tag reuse indicates common pilot activities (e.g., Table 4, Hungarian tags 1, 2, 3, 5). These were tags used to make a personal collection of good learning resources in foreign languages by a group of about ten teachers. Additionally, there are indications of rather unintentional sharing of a few tags among a few users (e.g. Table 4, tags 7 or 9).

Reuse on a personal level (i.e. applying previously created tags to posts) followed the same trend as the global reuse. 58% of the users did not reuse their personal tags; their posts only contained distinct tags. This was often times related to the low number of personal tags. In some cases, the users had created many distinct tags and never reused them. We are interested in finding out more about different patterns in personal tagging behaviour.

Table 4. Most used tags, the language, number of applications, tag class and number of users

Interpretation of the results on growth and reuse: The growth of posts in the system is sporadic, which may be explained by school holidays and teachers' active periods during the pilot. The fact that the number of new tags follows closely the number of posts (pink and blue lines in Figure 4) indicates that users are creating their personal tags as they create new posts, which most likely means that they have not yet developed a steady personal tag base. Others have observed that the growth entirely diminish over time (Marlow *et al.* 2006). We will further observe whether these trends will also appear in our system as it matures.

When it comes to the tag reuse in our system, we can look for reasons for it to be very low (1.22 users/tag). Similar to the interpretation from Farooq *et al.* (2007), we can partly opt for the influence of the tagging interface where global tags were absent, and in our case where resource-specific tags were only shown in the same language as the interface. The so-called "cold start

problem" may also contribute to the low reuse of tags; only 6.2% available learning resources have tags applied to them. When no social cues were made available, e.g. "5 users have added this to Favourites", it is rather random that a user tags a resource that was previously already tagged.

Lastly, we can speculate that the low level of personal reuse of tags was partly due to the fact that user was not familiar with tagging and was not able to see its benefits. In Table 4 we can see some examples of tags that were reused personally in order to create a collection of resources related to literature, chemistry and geometry (tags 6, 12, 15). This indicates that some teachers see the value of tagging for creating personal collection. We can assume that once others see this type of example through a tagcloud, for example, they would follow. Thus, we are interested in seeing to what level "social functionalities" such as a tagcloud affects both personal and global reuse of tags.

3.1.3 Problems with tags per post

Guy and Tonkin (2006) list a number of "sloppy" tags. We manually analysed a sample of posts (n=477) to see whether similar problems appear in our tagging system. We found redundancy within tags due to different spellings, use of quotes, capitalisation (the known problems of "sloppy tags"), but also due to the tag encoding, which required the user to enter a comma in order to separate tags from one another.

Out of our sample, 55% of the tags which appeared to be single tags according to our system actually were comprised of more terms. Orange bars (with pattern) in Fig. 5 show that 28% of the sample posts included a single tag, most posts (60%) include two tags and 12% of posts include three or more tags. These multi-term tags were not often compound words or literal concatenations of words (e.g. "thisisaspecialtag") as found by Guy and Tonkin (2006), but rather two separate terms, or in some cases, even sentence-like structures were found.

Figure 5. Number of tags per post: Orange bars (with pattern) that most posts (60%) in our sample (n=477) had 2 tags applied to them, whereas the system logs (red bars) erroneously show that most posts had only one tag (80%)

Interpretation of the results: We observed that our users' tagging behaviour divided. Even if they were told to use a comma, more than half ignored it. Our sample analysis showed quite a big discrepancy between what tags really are (Fig. 6 orange bars) and what our system records (red bars). The fact that many terms were bundled has an impact on their reuse, both on personal and global level. Thus, opting for a del.icio.us-like decision to treat each separate

term as a single tag could contribute to more reusable tags in the system. On the other hand, this type of decision needs user guidance in order to avoid a big diversity of "compound separators" (e.g. symbols like "_" or "+").

3.2 User's tagging behaviour in the multi-lingual context

We first look at how users tag in multiple language and then apply a tag classification to better understand the characteristics, such as non-obviousness and "travel well", of our tags in an educational setting. We also introduce a user study on how users perceive multi-lingual tags.

3.2.1 Tagging in multiple languages

We analysed all the unique tags that were recorded in the system. We included even the ones that users had deleted (199) from the posts to better understand the tagging behaviour. There were a total of 1031 tags in the system. Each tag has a unique ID. Additionally the system adds the language of the tag, timestamp and the ID of the learning resources that the tag is applied to. The language of the tag is inferred from the user interface language used while tagging. In this analysis we refer to this as inferred language. The interface was made available in the languages of the pilot and in English.

We studied the choice of the tagging interface language (Table 5). We can observe that pilot participants mostly chose to use the interface in their mother tongue (77%), and the rest of the time they mostly used the English interface.

Table 5. Tagging behaviour by language groups

We also undertook a manual language verification of our tags comparing the inferred language to the real language of the tag (Fig. 6). Most tags were in English (32%), although none of the pilot users are native English speakers. Other tag languages were Hungarian (20%), Polish (15%) and Czech (11%), which also were the major languages in the pilot (the other languages being German, Estonian, Lithuanian, Dutch, Slovenian and French). In Figure 6 the orange bars represent the real language of the tag that we verified manually, whereas the red is the inferred language from the user interface. This manual language verification revealed an error rate of about 30% in our simple approach to identifying the language of tags.

Figure 6. Real tag language (orange with pattern) and inferred language in red

Interpretation of the results on tagging in multiple languages: We found that users explore the tagging system in different languages. On average, every fourth tag was entered while using the tagging interface in another language than the user's mother tongue. More studies in this area would allow us to better understand personal tagging preferences: does everyone change languages while tagging, or only some of the users?

We can speculate that how users tag and in which languages they tag has ramifications on the viewable-tagging, we need further studies on what languages to display in order to promote multi-linguality and cross-border use of resources. Most likely this will have implications also on the convergence of tags over time within a language and languages in a multi-lingual and cultural context.

Inferring the language of the tag from the tagging interface left us with the error rate of about 30%. This is about the same as what Guy and Tonkin (2006) obtained while checking against a multi-lingual software dictionary. This discrepancy of language identification has ramifications on the usability of the Calibrate portal, reuse of tags, and how they can be used as navigational support. For example, the tag "Internet" was found four times in the system, twice with different capitalisation, once in Hungarian and once in English. Similar double entries of the same word with different language identification contributed to the fact that almost every 7th tag was redundant in the system.

3.2.2 Tag classification and "Travel well" tags

Apart from statistical properties of tags, we are also interested in the semantics of tags. In two different periods we manually categorised a sample of 819 of reused tags according to the classification from Sen *et al*. (2006), which is also based on the categories of Golder and Huberman (2006). They are Factual tags (Golder: item topics, kinds of item, category refinements); Subjective tags (Golder: item qualities) and Personal tags (Golder: item ownership, self-reference, tasks organisation). We have indicated these categories for our most used tags in Table 4.

Table 6. Categories of tags

Table 6 presents the tag categories of our sample. 74% of the tags applied are of factual type, such as describing the topic of the resource, its file type, the language or country the resources is related to. The second main category, some 25%, is subjective tags. These tags are used to describe the qualities of the resources or how the person felt about them. Apart from common pilot activities, there were very few subjective tags.

During our semantic tag analyses we also discovered a number of tags that stood apart hinting to us of some emerging trends. These tags were hard to group with one language as the spelling was identical in many languages (e.g. "chemie" has the same spelling in German, Dutch and Czech). Moreover, there were tags that presented a general term, a name, a place, or a country/area (e.g. EU, Euroopa, Evropa, Europa, europe) that is easily understood in other languages even if the spelling is slightly different. Other similar groups were people's names (e.g. Pythagoras, da Vinci) and commonly known acronyms (e.g. AIDS, USA). We call these tags "travel well" tags as users from different countries could easily understand them even without translation.

Some of these "travel well" tags were among the most reused tags in the system, examples of which can be seen in Table 4. The term "Matematika" (Table 4, no 7), for example, has the same spelling both in Czech and Hungarian. On the other hand, "test" (Table 4, no 9, we verified this tag was not to "test" the system), is used in many languages to indicate material suitable for exams or evaluation.

Interpretation of the results on semantic analysis: We had two interests in our semantic analysis of tags. On the one hand, we are interested in getting tags that add value to the system, and on the other hand, we wanted to better understand their usefulness for discovering resources across country and language border.

Others have also looked at the value of tags for an information system. Farooq *et al*. (2007) studied their system and introduced the Tag Non-obviousness metric. This metric could be used to detect tags that do not add much intellectual value to the tagging system as a whole. An example is a tag that repeats a term in the resource title. Such tag, when part of personal tags, can be useful as a personal descriptor and for retrieval, however, for the global use in the tagging system, it adds little new information.

In our case the LRE Application profile metadata already contains information such as the title of the resource, its language, etc. (indicated with * in Table 6). Thus, this type of information gathered from tags is redundant from the system point of view and adds little intellectual value to the tagging system as a whole.

On the other hand, tags from different categories can also add value in terms of helping users in their tasks. Sen *et al*. (2006) have looked at how different categories of tags were found useful for different tasks. For example, in MovieLens factual tags were good for finding movies and learning more about them, whereas subjective tags were good for making a decision on

which movie to watch. Similarly, we will continue observing our tag categorises to see if any similarities emerge.

As to our second goal with tags, using them as a navigational support to discover resources across borders, we think that "travel well" tags, due to the intrinsic properties that make them easily understood by many people, could act as a bridge between language groups to connect like-minded people across country and linguistic borders. In our future studies we will focus on the navigational aspects of "travel well" tags.

We also assume that "travel well" tags, which seem to be present mostly in the factual category, could be useful especially for less used languages in the system. We plan to display tagclouds in separate languages, and "travel well" tags could prove useful for less used languages. Also, when a user's language preferences is not known, or when no other resource-specific tags are available in the user's language, "travel well" tags can be used.

This analysis helped us to tune our system towards "travel well" tags and make sure that our new system requirements take advantage of these tags, either through an automated process or by asking users to identify them. The peril of this approach is that there are also words that look similar but have different meaning in different languages. There exist, for example, many faux amis (false friends) between English and French.

3.2.3 How users perceive tags in multiple languages

So far our Calibrate system has used tags only for personal management of learning resources, to "keep found things found" and managed. We plan to use tags as part of the resource metadata and in a tagcloud. Thus users reactions to tags in multiple languages became focus of our study. Especially, taking into account the issue discussed regarding language verification of tags (3.1.3) we were interested in how users react and cope with tags in languages that they are not familiar. In Vuorikari et al. (2007) we have reported this user study in detail.

In this study users indicated which thesaurus keywords and user-generated tags they found useful. Among the two most useful terms for each resource, we find that thesaurus terms were somewhat more popular (60%) than tags (40%). Another interesting outcome is that users occasionally found tags useful even if they were in languages that they did not have skills in. Most of these tags were what we described above as "travel well" tags. Figure 7 shows five bars that display the language of useful keywords to users. The orange is in a language that the user says he has skills in, and the red bars (with pattern) represent keywords in the languages that users did not know.

Figure 7. Percentage of keywords per LO in known languages (orange) and unknown language (red) that users found descriptive

Lastly, from our user study we can say that the issue of multi-lingual tags evokes sentiments and also splits users. Half of the users found them useful, whereas the other half found them confusing. One user even claimed to hate seeing keywords in languages that he/she does not understand. Participants in the last group also described that seeing tags in multiple language was rather irritating, especially when they were in languages that they did not recognise. It was also mentioned that multi-lingual tags make it harder and slower to pick the useful terms out of all the tags.

Interpretation of the results: The user study, which focused on users' attitudes towards multi/lingual tags, shows that tags in multiple languages divide users: some like them and others don't. Moreover, it gave us the indication that users may also find tags useful even if they are in languages that they do not claim to have competencies. This hinted to the direction of importance of "travel well" tags.

4. Discussion of the results in the light of our two main questions

In our discussion of the results and their interpretation, we attempt to find indicative answers to our two questions:
1) What happens when users tag in multiple languages?, and
2) Can we find any indication towards the use of tags and bookmarks to facilitate the cross border use of learning resources?

Due to the small sample size and the pilot nature of our tagging system, it is impossible to conclude whether tagging in many languages has a real impact on tag growth. We can see that in our system, the growth was rather similar to another tagging system in a similar context (Farooq *et al*. 2007) and that the users create new tags, either in their mother tongue, but also in English, in a manner similar to what happens in a mostly monolingual system. When it comes to reuse of tags, we also found indications of similar behaviour. However, we identified two main issues that hindered our analyses. First and foremost, the correctness of tag encoding and its related metadata needs to be addressed. Moreover, we discovered indications that our tag reuse most likely suffered from the design of the tagging interface, i.e. how multi-lingual tags were supported in viewable tagging.

Besides tag growth and reuse, we have been able to see that users discover resources in different languages and tag them using multiple languages

(Section 3.2). We found that some clear patterns emerge in how users tag in a multi-lingual context: they mainly tag in their mother tongue and in English (Table 4). More importantly, we found that despite tagging in different languages, there are tags that seem to be somewhat widely spread despite language borders. We call these "travel well" tags as they seem to be more easily understood without translation.

Our second question concerns whether tags and bookmarks could be used to facilitate the cross border discovery and use of learning resources? With cross-border use we mean users who use resources that come from different countries than they do, and can also be in different language from their mother tongue. As mentioned before, the cross-border discovery of resources can be challenging for users even if the searchable metadata is made available in multiple languages.

With a multi-lingual tagging system we have worked with the hypothesis that multi-lingual tags can yield new information regarding the resource itself and its usage. Tags could, for example, indicate the suitability of a given learning resource in a new lingual and cultural context. The semantic analyses preformed for this study help us see that users mostly apply tags that are factual (3.2.2). Even if we found that some of these tags were redundant with the information that we already have in the metadata (e.g. they repeat the title or the language of resource), it appears that users find tags in multiple languages somewhat descriptive and useful (3.2.3). This gives us an incentive to conduct future studies on their usefulness as a navigational tool.

Moreover, we discovered "travel well" tags. We assume that they could, due to the properties that make them easily understood by many people, act as a bridge across language and national borders, thus helping to create communities and clusters of like-minded users around tags and resources. During these analyses we found indications in this direction, e.g. shared use of some tags, as presented in Table 4, and small groups of users that formed around a number of tags.

Similar to the work of identifying informally powerful tags (Farooq *et al*.), we need to work on understanding what such tags are in our system (e.g. travel well, factual and subjective) and need to investigate whether those tags really foster creation of cross-language and cross-border communities.

Lastly, to demonstrate the across the national boundaries usage of digital resource, we used a visualisation tool to visualise all the bookmarked resources. Figure 8 represents bookmarked resources by users from different countries, each big round represent a node of users from a pilot country (some of which are orange). The nodes are connected by edges to the resources that users have bookmarked. In Figure 8 the resource in the middle, Match-

Teacher Educational Software, is highlighted in orange with edges connecting to users in five different countries (Poland, Estoia, Hungary, Lithuania and Chez Republic). This illustrates across the borders usage of the resource in question. Similarly, a number of small clusters are visible between the country nodes. These represent resources that bridge across national boundaries. In another paper of this Special Issue (Klerkx and Duval, 2008) another visualisation tool is described in details.

Figure 8. Visualisation of bookmarked resources that cross national borders

5. Contribution to design requirements of a multi-lingual tagging tool

This early study contributes to the understanding of tags and tagging behaviour in multiple languages. It can serve as a requirements survey for a multi-lingual tagging and navigation tool that needs to support multiple languages and discovery of resources across languages and country borders. In the spirit of "how to hide all but the right tags for each user", our analyses allowed us to further identify issues to work on.

These descriptive analyses show the importance of a correctly fine-tuned system that supports tagging in multiple languages; first of all, correct identification of the tag language is crucial, which will also allow the correct metadata on tag language. Moreover, it will enable calculating metrics similar to those presented in this paper possible without need of human intervention.

For fine-tuning a suitable language identification mechanism there is a need to investigate approaches using both existing software solutions and the ones that could take advantage of users' tagging behaviour. Although if our approach yields almost as good results as using multi-lingual dictionary software (e.g. Guy and Tonkin 2006), ours was only able to cover the languages in which the user interface was created. This is clearly insufficient in the future. Possible ways forwards could investigate, for example, tags against a properly managed multi-lingual list (e.g. WordNet) or creating lists of previously entered and validated tags. Also, testing new tags against characters specific to each language (language recognition chart in Wikipedia) could offer interesting results. Moreover, similar methods could be used for identifying "travel well" tags. Once the tag language has been correctly identified, its metadata can be added to the system correctly.

This study also showed the importance of the tagging interface and how it can passively affect on the tag reuse through the resource-specific or global tags that the user sees while tagging. Multi-linguality of tags adds an additional layer of complexity to the design of the tagging interface; overwhelming the user with tags in languages that they do not have competencies in can do a disservice for a multi-lingual system. This needs to be carefully considered also for the creation of a multi-lingual tagcloud.

6. Further work

Our analyses make it clear that using established metrics for monitoring tags and bookmarking activities allows comparing one's system to other existing systems and thus benchmark against them. We have realised that in the future there is a need to create more varied metrics that allow us to keep track of our multi-lingual tagging activities in a similar manner to Ochoa and Duval (2006). Apart from systematic and automated computation of the metrics introduced here, we are keen to create metrics to better track cross-border interactions, e.g. tags and bookmarks from users who come from a different country than that of the resource. Such metrics could be used to calculate the cross-border interactions of a given resource and tag. This could help identify resources that previous users from varied lingual backgrounds have found attractive within a large-scale collection of multi-lingual resources.

We are also keen to find more behavioural evidence on the usefulness of multi-lingual tags for users as for the resource discovery. We envisage metrics that can show how often a tag has been used to discover the resource, as opposed to using more conventional methods such as thesaurus terms or keyword based searches. In this area we are interested in enlarging the Contextual Attention Metadata framework to also support social information retrieval methods (Najjar *et a*l. 2006).

Moreover, now that new, effective technical architectures are in place to enable better discovery of educational resources across learning repositories on the international level, we are also interested in sharing tags with other learning resource repositories. Currently, there is a number of educational repositories that allow end user tagging (e.g. LeMill, OERCommons, KlasCement). Many of these repositories already share metadata regarding resources (through LRE network, Ariadne, Globe). Currently, however, currently tags are not shared and not used for navigational aid across repositories. Our small initial study on tags in Calibrate, LeMill and OERCommons show that there are many overlapping tags and interests by users in all systems (Vuorikari and Poldoja, 2008). Therefore, offering a way to navigate between systems by using tags could provide interesting avenues for end-users to cross system borders.

The issue of multi-lingual resources and tags is intriguing and offers interesting possibilities not only for end-users, but also for learning resources repository managers and administrators. We are interested in using our future metrics on multi-linguality to identify the information that the repository can gather from bookmarking and tagging activity to flag out learning resources that "travel well". Similar to the concept of "travel well" tags, these are resources that cross language and country borders easily. To identify potentially interesting "travel well" resources, we plan to use our cross-border metrics to better filter out or rank these resources. Future studies on validating this idea will be carried out.

A potential direction for future work will also need to consider recommender systems. A hybrid recommender system could consider a bookmark or tag as a vote for the resource. Additionally, other metadata (e.g. LOM) could be used to support content based filtering. Thirdly, information that the repository gathers through Contextualised Attention Metadata could also be taken advantage of (Najjar et al. 2006).

7. Conclusions

In this paper, we have presented some early and initial analyses of a multi-lingual tagging system. We analysed the general characteristics of our system, its tag growth and reuse, as well as categorisation of tags. We investigated how users tagged in a multi-lingual context. We discussed the findings in conjunction with design requirements to enhance our system. Lastly, we outlined our future work in this field.

We conclude that tags in a multi-cultural and multi-lingual context offer potential advantages to the collaborative tagging system and its multi-lingual user communities (e.g. Europe, on the international level). However, there are challenges and research questions that need further attention. As it becomes clear that some tags are useful for some users, the design challenge becomes "hiding all but the right tags".

9. References

- Cattuto, C., Schmitz, C., Baldassarri, B., Servedio, V.D.P., Loreto, V., Hotho, A., Grahl, M., Stumme, G. (2007) "Network Properties of Folksonomies". AI Communications 20, 245. Special Issue on Network Analysis in Natural Sciences and Engineering, 245-262. http://pil.phys.uniroma1.it/~ciro/files/cattuto_aicom2007.pdf
- Colin, J.-N. and Massart, D. (2006) "LIMBS: Open source, open standards, and open content to foster learning resource exchanges". In Kinshuk, R. Koper, P. Kommers, P. Kirschner, D. Sampson, and W. Didderen, editors, Proc. of The Sixth IEEE International Conference on Advanced Learning Technologies, ICALT'06, 682-686. http://doi.ieeecomputersociety.org/10.1109/ICALT.2006.222
- COM (2007) "Promoting language learning and linguistic diversity". Report on the implementation of the Action Plan Commission Working Paper, COM(2007) 554 final/2 http://ec.europa.eu/education/policies/lang/policy/report_en.html
- Farooq, U., Kannampallil, T. G., Song, Y., Ganoe, C. H., Carroll, J. M., and Giles, L. (2007) "Evaluating tagging behavior in social bookmarking systems: metrics and design heuristics". In Proceedings of the 2007 international ACM Conference on Supporting Group Work (Sanibel Island, Florida, USA, November 04 - 07, 2007), 351-360. http://doi.acm.org/10.1145/1316624.131667712
- Golder, S.A. and Huberman, B.A. (2006) "Usage patterns of collaborative tagging systems". Journal of Information Science 32(2), pp. 198—208. http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf
- Guy, M. and Tonkin, E. (2006) "Tidying up tags". D-Lib Magazine, Vol.12, month 1. http://www.dlib.org/dlib/january06/guy/08guy.html
- Hammond, T., Hannay, T., Lund, B., Scott, J. (2005) "Social bookmarking tools (i)". D-Lib Magazine, Vol.11, No.4. http://www.dlib.org/dlib/april05/lund/04lund.html
- IEEE LOM (2002). Standard for Learning Object Metadata, IEEE Learning Technology Standards Committee, IEEE 1484.12.1-2002.
- Klerkx, J., Duval, E. (2008) "Visualising Social Bookmarkstags"
- Lux, M., Granitzer, M., Kern, R. (2007) "Aspects of Broad Folksonomies". DEXA Workshops 2007, 283-287. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?tp=&arnumber=4312902&isnumber=4312839

- LRE Thesaurus (2002). ETB-project, European Schoolnet. http://insight.eun.org/ww/en/pub/insight/interoperability/learning_resource_exchange/metadata.htm
- LRE Application Profile (2007). http://insight.eun.org/intern/shared/data/insight/lre/AppProfilev3p0.pdf
- Marlow, C., Naaman, M., Boyd, D., Davis, M. (2006) "Position paper, tagging, taxonomy, flickr, article, toread". Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland. http://www.danah.org/papers/WWW2006.pdf
- Mathes, A. (2004) "Folksonomies-cooperative classification and communication through shared metadata". Computer Mediated Communication, Graduate School of Library and Information Science, University of Illinois Urbana-Champaign. http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html
- McCormick, R., Scrimshaw, P, Li, N., Carmel, C. (2004) Celebrate Evaluation, Deliverable 7.2, European Schoolnet. http://celebrate.eun.org/eun.org2/eun/Include_to_content/celebrate/file/Deliverable7_2EvaluationReport02Dec04.pdf
- Najjar, J., Wolpers, M., Duval, E. (2006) "Towards Effective Usage-Based Learning Applications: Track and Learn from User Experience(s) ", IEEE International Conference on Advanced Learning Technologies (ICALT 2006), 1022-1024. http://portal.acm.org/citation.cfm?id=1156068.1156352&coll=GUIDE&dl=GUIDE&CFID=5856084&CFTOKEN=72510094
- Ochoa, X., Duval, E. (2006) "Towards automatic evaluation of learning object metadata quality". In: Advances in Conceptual Modeling - Theory and Practice, ER 2006 Workshops BP-UML, CoMoGIS, COSS, ECDM, OIS, QoIS, SemWAT. Lecture Notes in Computer Science, 372–381. http://ariadne.cti.espol.edu.ec/M4M/files/TowardsAutomaticQuality.pdf
- Sen, S., Lam, S.K., Cosley, D., Frankowski, D., Osterhouse, J., Harper, F.M., Riedl, J. (2006) "tagging, communities, vocabulary, evolution". In: Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work. 181–190. http://www.grouplens.org/papers/pdf/sen-cscw2006.pdf
- Vuorikari, R., Ochoa, X., Duval, E. (2007) Analysis of User Behavior on Multilingual Tagging of Learning Resources. In Proceedings of the 1st Workshop on Social Information Retrieval for Technology Enhanced Learning, SIRTEL'07. CEUR, VOL-307, 6-17. http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-307/paper02.pdf

- Vuorikari, R., Van Assche F. (2007) "Collaborative content enrichment in multilingual Europe, European Schoolnet approach on educational resources". In Workshop proceedings of AXMEDIS2007, 85-88.
- Vuorikari, R., Poldoja, H. (2008) "Comparison of educational tagging systems - any chances of interplay?". In Proceedings of the 2nd Workshop on Social Information Retrieval for Technology Enhanced Learning, SIRTEL'08. CEUR, VOL-382 http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-382/paper1.pdf