

# Ecology of social search for learning resources

## Citation for published version (APA):

Vuorikari, R., & Koper, R. (2009). Ecology of social search for learning resources. *Campus-Wide Information Systems*, 26(4), 272-286. <https://doi.org/10.1108/10650740910984619>

## DOI:

[10.1108/10650740910984619](https://doi.org/10.1108/10650740910984619)

## Document status and date:

Published: 28/08/2009

## Document Version:

Peer reviewed version

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

<https://www.ou.nl/taverne-agreement>

## Take down policy

If you believe that this document breaches copyright please contact us at:

[pure-support@ou.nl](mailto:pure-support@ou.nl)

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 26 Nov. 2020



# Ecology of social search for learning resources

## Abstract

**Purpose:** This paper deals with user-generated Interest indicators (e.g., ratings, bookmarks, tags). We answer two research questions: can search strategies based on Social Information Retrieval (SIR) make the discovery of learning resources more efficient for users, and can Community browsing help users discover more cross-boundary resources. By cross-boundary we mean that the user and resource come from different countries and that the language of the resource is different from that of the user's mother tongue.

**Design:** We focus on a portal that accesses a federation of multilingual learning resource repositories. A measure for user's efficiency in finding relevant resources was defined. We then collected users' attention metadata and use this empirical data to answer two hypotheses.

**Findings:** We show that users are more efficient with Social Information Retrieval strategies, however, Community browsing alone does not help users discover a wider variety of cross-boundary resources.

**Practical implications:** By social tagging and bookmarking resources from a variety of repositories, users create underlying connections between resources that otherwise do not cross-reference. This is important for bringing them under the umbrella of SIR methods. Future studies should include testing SIR methods to leverage these user-made connections between resources that originate from a number of countries and are in a variety of languages.

**Originality:** The use of attention metadata to model the ecology of social search adds value to the actors of learning object economy, e.g. educational institutions, digital libraries and their managers, content providers, policy makers, educators and learners.

**Keywords:** learning resources, social information retrieval, social tagging, metrics, efficiency, Research paper

## 1. Introduction

Learning resource repositories and libraries make educational material and/or its metadata available in digital format, the sharing of which is their core *raison d'être*. Their reuse has been touted for enabling cost savings because the creation of high quality material is costly, hence the focus on standards that enable interoperability (Cambell, 2003) even across repositories (Ternier, *et al.*, 2008). Traditionally, metadata and/or web directories are used for searching and exploring educational content. Currently, novel exploratory search systems are developed for learning resources to assist users in obtaining content that meets their information needs. Such systems include social navigation and collaborative recommender systems, both of which belong to the family of techniques called Social

Information Retrieval (Goh *et al.*, 2007).

Social navigation involves using the behaviour of other people to help navigate online. It is driven by the tendency of people to follow other people's footprints when they feel lost (Dieberger *et al.*, 2000). Such footprints in an online environment are what Claypool *et al.*, (2001) define as Interest indicators which can be acquired either directly from the user (e.g. rating) or indirectly (e.g. time spent on an object). Collaborative recommender systems, on the other hand, use explicit ratings to find like-minded users (Adomavicius and Tuzhilin, 2005). Evaluation of recommender systems traditionally focuses on the algorithms and their performance (Herlocker *et al.*, 2004), similar to exploratory search systems (White *et al.*, 2008). Evaluating recommenders from the user perspective has received less attention (McNee, 2006).

Within the field of Technology Enhanced Learning (TEL) such systems exist. Rafaeli *et al.*, (2005) introduced a system to harness the social perspectives in learning where the learner could choose from whom to take recommendations (friend or algorithm). Koper (2005) used indirect social interaction in choosing a path that allows successful completion of a learning task. Drachsler *et al.*, (2008) took this research further showing that users employing a recommender system, that offers navigation support in self-organised Learning Networks, were more efficient time-wise in completing an equal number of Learning Activities. Farzan and Brusilovsky (2005) studied social navigation and found that adding the time spent reading each page provides more precise insight into the intention of the group of users and more accurate information about pages selected from search results. Jung *et al.*, 2007 studied implicit click data to increase both precision and recall of the feedback data on a university search portal. Tang and McCalla (2009) studied the pedagogical value while using collaborative filtering to recommend papers for learners, and Manouselis *et al.*, (2007) used multi-criteria ratings to recommend resources to teachers.

Both the field of recommender systems and social navigation, however, suffer from the same problems: how can Interest indicators be gathered without being too intrusive, and yet, at the same time, remain accurate enough in guiding users in their choice of product or navigational path. The sparse data and new items often are problematic too (Herlocker *et al.*, 2004, Adomavicius, Tuzhilin, 2005, Rafaeli *et al.*, 2005). Social bookmarking and tagging can offer new prospective thanks to tags, in which users describe their interest. This creates a triple (user, content, annotations) which indicates user's relationship between resources, users, and tags (Golder and Huberman, 2006, Marlow *et al.*, 2006, Sen *et al.*, 2006). Such underlying structure allows flexible social navigation (e.g. tag-item, tag-user, user-item), but could also be a source for collaborative recommender systems by linking like-minded users not only through resources, but also through tag-based interest sharing (Santos-Neto *et al.*, 2009). The idea of social tagging and bookmarking has been implemented in the TEL context (Maier, Thalmann, 2008, Vuorikari, Poldoja, 2008), in digital libraries (Puspitasari *et al.*, 2007) and for scientific papers (Farooq *et al.*, 2007, Gatutto, *et al.*, 2008).

Millen *et al.*, (2007) studied the use of social bookmarking at the enterprise level and suggest that integrated with traditional search engines, it has the potential to solve commonly known enterprise search problems, e.g. content from heterogeneous repositories that do not cross-reference via hyperlinks (Mukherjee, 2004). We study a similar implementation for multilingual learning resources within a federation that has social

bookmarking and tagging features (Vuorikari & Van Assche, 2007). Our aim is to study such a hybrid system to understand how it is used, how different variables are interconnected, and finally, how the behaviour of previous users could be leveraged to support and enhance the discovery process of educational resources for all users of the system.

From previous studies (e.g. McGormick *et al.*, 2004, Vuorikari and Koper, *submitted*) we have evidence that users of educational content use cross-boundary resources to a certain extent, but their reuse remains rather low.

In this study our specific questions focus on the efficiency of resource discovery and how we can support the users in discovering cross-boundary resources. We have defined our two hypotheses as following:

*H1*: The search methods that take advantage of Social Information Retrieval yield more relevant resources with less effort from the users than the methods based on conventional text based search.

*H2*: The users who take advantage of Community browsing discover more cross-boundary learning resources than those who use conventional text base search.

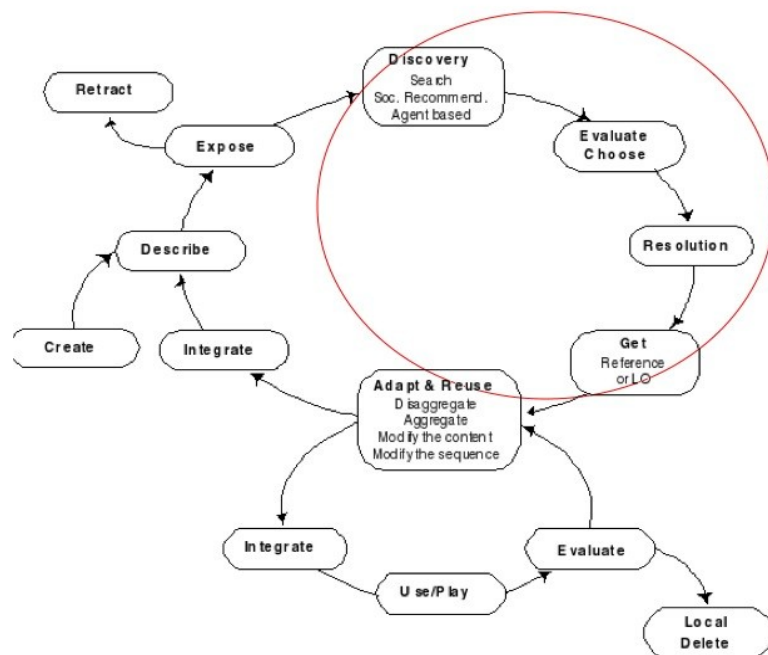


Figure 1. Learning resource lifecycle (Van Assche, Vuorikari, 2006)

We focus on a specific moment in the lifecycle of a learning resource, namely when the user discovers the learning resource and evaluates whether it matches with the information seeking need at hand (Figure 1).

In the following section we introduce the study methodology and the data set. The next section focuses on the results followed by the discussion. The paper is concluded with an

outlook for future work.

## 2. Context of the study and its method

The portal that we study makes open educational resources available from 19 content providers from Europe and elsewhere. These resources exist in different languages and conform to different national and local curricula. The portal, developed in the MELT project (Figure 2), offers three different categories of searches (Millen *et al.*, 2006).

The screenshot displays the MELT portal interface. At the top, a welcome message states: "Welcome to the Learning Resource Exchange (LRE) portal for schools where you can find open educational resources and assets from many different countries and providers, including 16 Ministries of Education." Below this, three boxes provide statistics: "What? 31333 Resources, 104798 Assets, 19 Providers", "How to register?", and "How to search?". The main content area features three tabs: "Travel well", "Most bookmarked", and "Featured". The "Featured" tab is active, showing three search results. Each result includes a star rating, a title, tags, keywords, resource type, description, and Creative Commons license. The results are: 1. "Hydroelectric power [no]" with a 3-star rating, tags "interactive melt selection interactive", keywords "hydroelectric power", resource type "exploration, application", and description "Learn the principles behind the production of electricity (...)". 2. "Addisjon [no]" with a 2-star rating, tags "melt selection interactive k nnendmurrud liittmine", keywords "addisjon", resource type "drill and practice, exploration, web page", and description " v p    legge sammen desimaltall. Hvor lenge klarer du 100%?". 3. "Loudspeaker [no]" with a 1-star rating, tags "interactive melt selection", keywords "loudspeaker, light, soundwaves, music", resource type "exploration, application", and description "How does a loudspeaker work? The illustration shows the (...)". On the left side, there are search filters for "Language" and "Subject", an "Age range" selector, and a "Find by tags" section with a cloud of tags including "allatok", "angol", "kicsiknek", "animals", "onice", "efl", "english", "fun", "games", "geschiedenis", "grammar", "interactive", "k mia", "kirjandus", "koolieleu", "Latijn", "melt", "selection", "picture", "promethean", "quiz", "Rome", "Romein", "Romeinen", and "vocabulary worksheet". There is also a "Find by subject" section with categories like "Art", "Biology", "Chemistry", and "Citizenship".

Figure 2. Melt portal.

Explicit search: comprises the traditional search box with text and filtering options based on multilingual metadata. "Find by subject" offers browsing through pre-defined categories. The results are shown on the Search Result List (SRL) with metadata and annotations, if available.

Community browsing: these are social navigation features such as accessing resources through tagclouds and specific lists of most bookmarked resources, but also "pivot browsing" which means using tags or usernames as a means to reorient browsing. Figure 2 shows examples such as the multilingual tagcloud and different tabs (e.g. most bookmarked).

Personal search: Looking for bookmarks from one's own personal collection of bookmarks (MyFavourites).

In this study, Social Information Retrieval methods mean all the Community browsing features, and also comprise of retrieved resources that contain user-generated Interest indicators. These are Interest indicators like a rating on a scale 1 to 5 (1="of no use" to 5="very useful") or a bookmark with tags (called Favourites). These public annotations are also called *contributing actions*. By clicking on the link the user plays a resource and

generates click-through which is an implicit Interest indicator. We use the term “cross-boundary discovery” when the user bookmarks or rates resources that come from different country than she does and that are in a language other than that of the user’s mother tongue.

We use the following metric for relevance: when a resource is added in the Favourites, or the resource is rated with the value of 3 or greater. Such relevance represents the relationship between the object and the information need, as perceived by the user. We do not focus on the other types of relevance, such as the query and object match or topical relevance (Borlund, 2003). With "less effort from the user" we mean that users will play and annotate relevant resources with less search effort (i.e. the number of executed searches).

## 2.1 Description of the data

We have defined our own logging scheme for users’ attention metadata, exportable in a Contextual Attention Metadata compliant format (Najjar *et al.*, 2007). It was first used on the Calibrate pilot and now we redefined the scheme to cover a range of contributing actions and social search. Table 1 describes the scheme that includes 3 main units: the resources, the user and user’s action. The latter includes three main categories: search, click and contributing actions. Each time a user arrives on the portal, a new session is started, regardless if the user logs in. Table 1, for example, shows that metadata about “languages” is gathered for the resource, the user (mother tongue and spoken languages) and for the tags.

**Table 1. Logging scheme on the portal to capture users' attention metadata. LO stands for Learning object (resource).**

Metadata	Unique ID	Languages	Location	Time stamp	Other metadata
<b>1. Resources</b>	x	x			LRE Application profile
<b>2. Users</b>	x	x		x	Country, school, topics of interest.
<b>3. Actions</b>					
3.1 Session	x			x	ID, if logged in.
3.2 Search				x	Type of search: advanced, browse topics, tag search
3.3 Click			x	x	The LO, tag, rating, other user clicked on
3.4 Bookmark	x		x	x	User ID, LO ID
3.5 Tag	x	x	x	x	User ID, LO ID
3.6 Rate	x	x	x	x	User ID, LO ID, value, comment

The registered users, mostly primary and secondary teachers, were invited to use the portal after an initial pilot. They came from 11 different European countries (Austria, Belgium, Estonia, Finland, France, Germany, Greece, Hungary, Slovenia, Sweden and United Kingdom).

The attention metadata logs were collected from the period of Oct 1 to Dec 18 2008. After excluding the project staff and users who did not search and click resources, we were left

with 82 users who had clicked at least on one resource. Table 2 presents the data used for this study which include registered users (82), but also sessions from unregistered users.

**Table 2. Data description of all recorded attention and actions included in the study.**

	Sessions	Executed searches	All plays	Distinct resources played	Distinct resources bookmarked and rated
Unregistered users	2036	7846	1854	1547	Not available to users
Registered users (82)	310	1863	974	687	394

During the pilot, a selection of more than 30 000 distinct resources was made available. 565 resources (less than 2% of all resources) had been annotated prior to the period of study by the project staff. Figure 3 shows that the growth of new users and tags follow one another. Farooq *et al.*, (2007) explain that linear tag growth suggests that the tag vocabulary is still maturing and has not yet reached its relatively stable stage, where the tag growth diminishes over time (Marlow *et al.*, 2006), a view that we also adopt for our system.

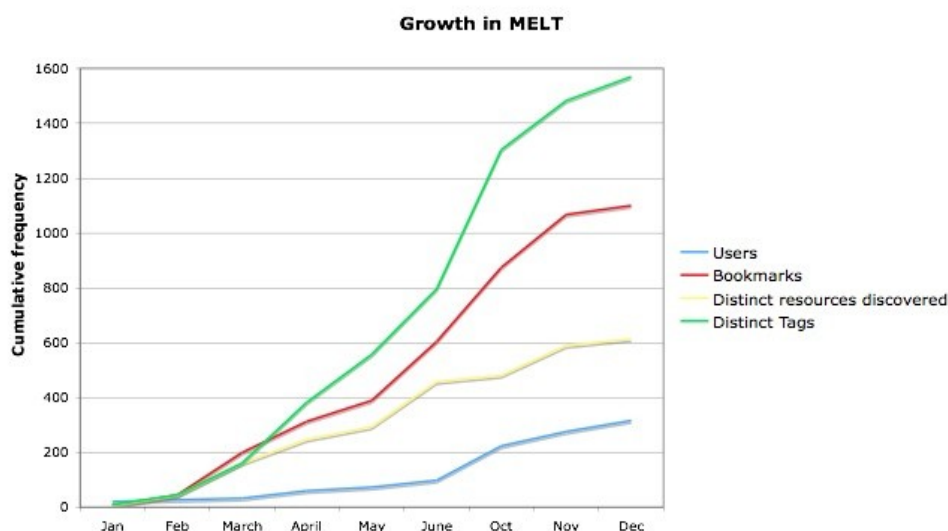


Figure 3. Tag and user growth in MELT. The portal was not in use of the summer months.

## 2.2 Data processing

We tested the normality of our data using the Kolmogorov-Smirnov test, which indicates a highly significant deviation from normality. As the assumption of normality was not tenable, we used methods for non-parametric data (e.g. Kendall tau). For testing the association between the type of search and the number of actions that followed, we used the Pearson Chi-Square test ( $p < 0.001$ ).

## 3. Results

We first give descriptive results on our three main actions: how do users search, play and contribute, and then look how new resources generated clicks and annotations. Following,

we present the model of the process that produced this data and then focus on our two hypotheses.

### 3.1 Descriptive results

#### Search

We found three main groups of search behaviour among the registered users; the ones who only used Explicit (47%) or only Community browsing features (6%) and about half of the users who used mixed methods (47%). Of all searches executed on the portal, 82% were Explicit methods (53% advanced searches; 29% browsing by discipline) and 18% Community browsing.

#### Click-through

Users generate click-through when they “play” a resource. Counting all click-through from all users, 1547 distinct resources were played 2828 times. Different user behaviour was observed among registered users: 52% played resources only on the Search result list, 10% only in Community browsing areas, and the rest (38%) in different parts of the portal.

Table 3 shows that 69% of the click-through was generated on Search result list on resources with no Interest indicators, whereas 31% of the click-through came from resources that had Interest indicators. As our Explicit search function does not take advantage of Interest indicators, we consider the latter to be an outcome of Social Information Retrieval.

**Table 3. Amount and percentages by areas of the portal where click-through was generated.**

<b>Click-through</b>	<b>Registered users</b>	<b>Other</b>	<b>All</b>	<b>%</b>
Search Result List	728	1235	1963	69.4%
Search Result List with Interest indicators	40	85	125	4.4%
Tagcloud	103	124	227	8.0%
“Travel well” list	68	300	368	13.0%
Favourites	20	34	54	1.9%
“Most bookmarked” list	15	76	91	3.2%
<i>Grand Total</i>	974	1854	2828	99.9%

#### Annotations

77% of registered users who played resources also annotated them: 44% both rated and tagged resources, whereas 33% only rated or tagged resources. Table 4 shows that users rated and bookmarked very similarly. Two users out-performed others with 120 and 108 annotations. In general, the ratings are positive; 84% voted 3 or greater ( $\geq 3$ ) and only 16% of ratings indicate a resource “of no use” (=1 or 2).



**Table 4. Descriptive statistics and distribution of bookmarks and tags.**

<b>Annotations:</b>	<b>Bookmarks and tags</b>	<b>Ratings</b>
Users	48	46
Number of actions	350	384
Number of tags	1507	Rating $\geq 3$ (84%)
Minimum of actions	1	1
Maximum of actions	65	56
Mean of bookmarks	7.29	8.34
Median	3	3
Mode	1	1
Standard Deviation of annotations	12.19	12.53

Annotations took place in different parts of the portal: 40% of users only annotated resources in the Search result list, 24% only in Community browsing parts of the portal and 36% of users annotated in both places, which resulted in most of annotations (70%). Most resources that were annotated (67%) had no previous annotations and were found through Explicit search (Table 5).

**Table 5. Place of the portal where annotations took place.**

	<i>Bookmarks obtained after..</i>	<i>Ratings done in..</i>
..Explicit search no annotations	236 (67%)	258 (67%)
..Explicit search with annotations	49 (14%)	26 (7%)
..Community browsing	65 (19%)	85 (22%)
..Personal Search	n/a	15 (4%)

### **Interest indicators, click-through and contributions**

We found a tendency that resources with Interest indicators generated more actions: 14% of all played resources that had Interest indicators generated 29% of all plays. Table 6 shows that we found a small correlation coefficient between resources that had existing Interest indicators and bookmarking action (0.329,  $p < 0.01$ ), similarly between actions play and bookmark (0.327,  $p < 0.01$ ) and some between bookmark-rate (0.300,  $p < 0.01$ ). These correlations between actions can be further studied for generating implicit Interest indicators that are less intrusive for users.

**Table 6. Correlation coefficient between different actions.**

<b>Correlation coefficient between</b>	<b>Kendall Tau</b>
Resource w/Interest indicators – bookmark action	(0.329, $p < 0.01$ )
Click-through -bookmark action	(0.327, $p < 0.01$ )
Bookmark action –rate action	(0.300, $p < 0.01$ )
Resource w/Interest indicators and click-through	(0.253, $p < 0.01$ )
Click-through –rate action	(0.233, $p < 0.01$ )
Resource w/Interest indicators and rate action	(0.196, $p < 0.01$ )

### 3.2 Modelling users' actions: ecology of social search

A model was created of the process that produced the above described data to study how processes are interlinked (i.e. ecology); for example, how newly created annotations become part of the search process where users interact and eventually annotate them.

Figure 4 shows authenticated and unauthenticated users on the portal and the percentage of their actions by each category (search-play-contribute). In principle, these two groups search rather similarly: Explicit search is preferred by both groups and Community browsing methods account for 21% of all search actions.

When looking into details, the main difference is that the unauthenticated users tend to explore resources through browsing (51%): 30% browse by discipline and 22% Community browsing, whereas the authenticated users browse less (38%), but additionally use Personal search (9%). The Pearson Chi-Square test was significant ( $p < 0.001$ ) for these differences.

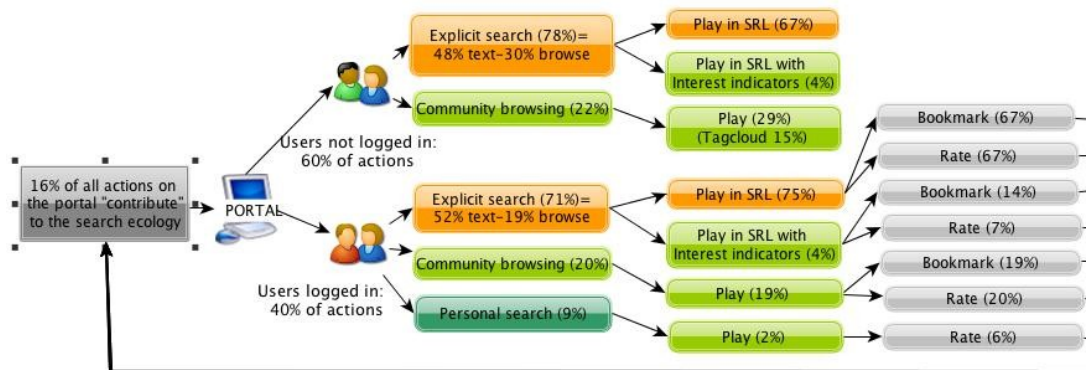


Figure 4. Users consume and contribute: orange indicates actions of play after Explicit search, whereas light green indicates actions of play after Community browsing, which are explained in Section 2. Grey actions contribute to the portal's social search ecology.

Users play resources differently, unauthenticated users play 71% of resources in the Search Result list (SRL) after an Explicit search and 29% in Community browsing areas. Authenticated users, on the other hand, play 79% of resources in the SRL and only 21% in Community browsing areas. The Pearson Chi-Square test was significant ( $p < 0.001$ ) for these differences.

Only users who are authenticated to the system can contribute. Out of all the actions recorded in the system, contributing actions amount to 16%. Most contributing actions are annotations (67%) on newly discovered resources, whereas 33% is generated in the Community browsing areas on resources that have previous annotations. The annotations are generated following a rather regular pattern (grey boxes in Figure 4), however, for ratings, a small amount is created in the user's Favourites.

The model shows that the annotation (i.e. Interest indicators) play an integral part in creating a social search ecology and offer more diverse ways to discover resources. 16% of all the search-play-annotation actions contribute to this ecology, which are used to create novel ways to discover resources. 21% of all search actions took advantage of these

features, and they supported 31% of plays. Our finding points to the same direction as Glahn *et al.*, (2008) who found that accessing of tagged resources is independent from the contribution level of a participant. However, we did not find that all participants use tags similarly while searching or accessing tagged resources: we observed different search preferences by users.

### 3.3 Measure of efficiency: Search-play-contribute ratio (H1)

*Hypothesis 1 (H1)* states that the search methods based on Social Information Retrieval yield more relevant resources with less effort from users. We use the above model to create a measure to test our hypothesis by studying the ratio between search, play and contributing actions from users who were authenticated on the portal. We are inspired by the Click-through rate that measures the success of an online advertising campaign. The rate can be obtained by dividing the number of users who clicked on an advertisement by the number of times it was delivered (e.g. Ward and Kirthi, 2007).

Our application of the efficiency measure shows how many search actions it takes to play and/or annotate a resource. For both search methods we created the following efficiency ratios: search-play, search-rate (3 or higher), search-bookmark, play-rate (3 or higher), and play-bookmark. The lower the figure the better, as it indicates the number of search actions that it takes the user to achieve the goal.

Table 7 presents the two baselines against which we measure. *Explicit search* indicates the results from a text-based search excluding the plays on resources that contained Interest indicators (i.e. ratings, tags). *Community browsing* indicates the results from tagcloud, lists and pivot browsing. *SIR methods*, against which we measure the efficiency, include results from *Explicit search* with Interest indicators (i.e. ratings, tags) and from *Community browsing*. On the top row, different ratios for actions are indicated, e.g. how many searches does it take to play a resource (2.1 Explicit searches to one play). An *Efficiency rate* for each search method is displayed, which is an average of all ratios. This allows for quick comparison. Lastly, *Explicit search (comparison)* reports the same ratios from another portal where no Interest indicators were made public (see Vuorikari, Ochoa, 2009). Comparing the two top rows we compare the efficiency of *Explicit search* with *Community browsing*. We find that the search-play, search-bookmark and play-bookmark ratio is almost identical using both methods, whereas Community browsing is more efficient for rating.

**Table 7. Users' efficiency with different search methods on the MELT portal.**

<b>Actions for authenticated users in MELT</b>	<b>Search: play ratio</b>	<b>Search: rate (<math>\geq 3</math>) ratio</b>	<b>Search: bookmark ratio</b>	<b>Play: rate (<math>\geq 3</math>) ratio</b>	<b>Play: bookmark ratio</b>	<b>Efficiency rate</b>
Explicit search (baseline)	2.1:1	7.0:1	6.4:1	3.3:1	3.1:1	4.4
Community browsing (baseline)	2.0:1	5.3:1	6.5:1	2.6:1	3.2:1	3.9
SIR methods	1.7:1	4.0:1	3.7:1	2.3:1	2.2:1	2.8
Explicit search (comparison)	0.5:1	21.6:1	11.3:1	42.6:1	22.2:1	19.6

Comparing the *Explicit search* baseline to *SIR methods* shows an efficiency gain, the *Efficiency rate* drops from 4.4. to 2.8. Search-rate ratio comes down from approximately 7 searches to 4 searches, whereas play-rate from 3 searches to 2 searches.

We conclude that on the MELT portal, *Explicit search* and *Community browsing* perform very similarly. However, *Explicit search* compared with *SIR methods* shows an efficiency gain of 1.6 units. For both the *Explicit search* data and SIR data, Pearson Chi-Square test was highly significant ( $p < 0.001$ ). We thus conclude that the findings support *H1* and we can accept *H1*.

### 3.4 Discovery of cross-boundary resources (H2)

*Hypothesis 2 (H2)* states that the users who use *Community browsing* discover more cross-boundary learning resources. In addition to using the above measures, we add a component indicating the cross-boundary nature of a bookmark. To do this, we compared the country of the resource to that of the user and the language of the resource to that of the user’s mother tongue as in Vuorikari and Koper (*submitted*). Additionally, we created a “search method” profile and “cross-boundary resource profile based on bookmarks” for each user. These profiles are comprised of percentages of different search methods by the user, and the percentage of resources that cross national and linguistic boundaries.

57% of all bookmarked resources by users, total of 198, are cross-boundary discoveries (Table 8). Out of 48 users who bookmarked these resources, 40 had cross-boundary resources (i.e. 83% with cross-boundary profile and 17% with only national profile). This split reflects the goal of the portal which promotes the discovery of educational material from different countries. Such cross-boundary profiles are usually also “cross-repository” profiles, information of which can be used to make link structures between content in different repositories that otherwise is not cross-referenced via hyperlinks nor metadata.

**Table 8. Cross-boundary discoveries by users.**

	Distinct LOs		Distinct LOs w/SI		“New” discoveries	
Cross-boundary discovery	198	(57% of all)	73	(37%)	125	(63%)

72% of these bookmarked cross-boundary resources were discovered in the Search result list as a result of *Explicit search*, whereas 28% were a result of *Community browsing*. Within *Community browsing*, 23% were discovered in the tagcloud and 5% chosen from the “Travel well” list. These findings show that most often users discover cross-boundary learning resources as a result of *Explicit search*, and when the resource is deemed relevant, they bookmark it in the Search result list. This finding from our current data points to the direction that *H2* cannot be supported, and prompts us to reject *H2*.

## 4. Conclusion

In this paper we used empirical data in the form of attention metadata to track users’ actions. We then use this data to model the ecology of social search of a learning resource

portal integrated with a social bookmarking and tagging system. We conclude that Interest indicators have an important role as a part of the social search ecology, and we expect that studies into inter-relations of these variables will offer interesting further insights. By studying the cross-boundary discoveries, we find that users create underlying connections between resources that come from a number of countries and are in a variety of languages, which is important for bringing them under the umbrella of SIR methods.

*Hypothesis 1* was accepted showing that the search taking advantage of Social Information Retrieval (SIR) methods yield more relevant resources with less effort from the user. Despite this edge, users have a strong search preference for Explicit search methods (2/3 of all executed searches). These conventional search methods strongly proved their role in discovering new resources and thus also alleviating the “cold start problem”. This also led us to reject our second hypothesis (H2): most cross-boundary resources are discovered using Explicit search. Encouraged by the H1, though, we believe that leveraging both implicit and explicit Interest indicators to support cross-boundary discovery (e.g. indicating the cross-boundary nature of resource discoveries, tag clouds filtered by language and by the country of users), and collaborative filtering methods for like-minded users based both on item and tag-based interests, are worth studying further.

A limitation in this research was that we studied a system that was still evolving. This has the positive effect of allowing us to revisit our SIR strategies for both cross-boundary and within-boundary discoveries to better support users. A second limitation is that as the study is based on server-side log-files, we have left out subjective measures such as user satisfaction or cognitive load while searching for cross-language content, which would also add important information in studying such system. Future studies should include testing wider range of SIR methods to leverage the underlying connections that users have created through social tagging and bookmarking between resources that originate from a number of repositories and are in a variety of languages.

## **References**

Adomavicius, Tuzhilin (2005), “Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions”, *IEEE Transactions on knowledge and data engineering*, Vol. 17 No.6.

Borlund, P. (2003), “The Concept of Relevance in IR”, *Journal of the American Society for Information Science and Technology*, Vol. 54 No. 10, pp. 913-925.

Campbell, L. (2003), “Engaging with the learning object economy”, Littlejohn, A. (Ed.) *Reusing Online Resources: A Sustainable Approach to e-learning*, Kogan Page Limited, London, pp. 35-45.

Cattuto, C., Schmitz, C., Baldassarri, A., Servedio, V. D., Loreto, V., Hotho, A., Grahl, M., and Stumme, G. (2007), “Network properties of folksonomies”, *AI Commun*, Vol. 20 No.4, pp. 245-262.

Claypool, M., Le, P., Wased, M. and Brown, D. (2001), “Implicit Interest Indicators”, *ACM Intelligent User Interfaces Conference (IUI 2001)*, ACM, Santa Fe: New Mexico, pp. 33-40.

Dieberger, A., Dourish, P., Höök, K., Resnick, P., and Wexelblat, A. (2000), "Social navigation: techniques for building more usable systems". *Interactions* Vol. 7 No.6, pp. 36-45.

Drachsler, H., Hummel, H., Berg, B., Eshuis, J., Waterink, W., Koper, R. (2008), "Effects of the ISIS Recommender System for navigation support in self-organised Learning Networks". In *Proceedings of Technology Support for Self-Organized Learners (TSSOL 2008)*, Salzburg, Austria.

Farooq, U., Kannampallil, T. G., Song, Y., Ganoe, C. H., Carroll, J. M., Giles, L. (2007), "Evaluating tagging behavior in social bookmarking systems: metrics and design heuristics", *Proceedings of the 2007 international ACM Conference on Supporting Group Work* (Sanibel Island, Florida, USA), pp. 351-360.

Farzan, R., Brusilovsky, P. (2005), "Social navigation support in e-learning: what are the real footprints", *Intelligent Techniques for Web Personalisation (ITWP'05)*.

Glahn, C., Specht, M., Koper, R. (2008), "Implications of writing, reading, and tagging on web for reflection support in informal learning", In M. Specht, P. Dillenbourg, E. Duval (Eds), *Third European Conference on Technology Enhanced Learning EC-TEL 2008*, Springer, Berlin, Heidelberg.

Goh, D.H., Foo, S., "Preface", in Go D.H. & Foo S. (2007), *Social Information Retrieval Systems: Emerging Technologies and Applications for Searching the Web Effectively*, Hershey, PA: Idea Group Publishing.

Golder, S., Huberman, B.A. (2006), "Usage patterns of collaborative tagging systems", *Journal of Information Science*, Vol. 32 No.2, pp. 198-208.

Herlocker, J. L., Konstan, J. A., Terveen, L. G., and Riedl, J. T. (2004), "Evaluating collaborative filtering recommender systems". *ACM Trans. Inf. Syst.* Vol. 22 No.1, pp. 5-53.

Jung, S., Herlocker, J. L., and Webster, J. (2007), "Click data as implicit relevance feedback in web search". *Inf. Process. Manage.* Vol. 43 No.3, pp. 791-807.

Koper (2005), "Increasing Learner Retention in a Simulated Learning Network using Indirect Social Interaction", *Journal of Artificial Societies and Social Simulation*. Vol. 8 No.2.

LRE Application Profile (2007). Learning Resource Exchange Application Profile. European Schoolnet. <http://fire.eun.org/>.

Maier, R., Thalmann, S. (2008), "Institutionalised collaborative tagging as an instrument for managing the maturing learning and knowledge resources", *International Journal of Technology Enhanced Learning*. Vol. 1 No.½.

Manouselis, N., Vuorikari, R., Van Assche, F. (2007), "Simulated Analysis of MAUT Collaborative Filtering for Learning Object Recommendation", *Proceedings of the 1st Workshop on Social Information Retrieval for Technology-Enhanced Learning*, Crete,

Greece. CEUR, 307.

Marlow, C., Naaman, M., boyd, d., Davis, M. (2006), "HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, To Read", *Proceedings of the seventeenth conference on Hypertext and hypermedia*, Odense, Denmark, pp. 31-40.

McNee, S., Meeting User Information Needs in Recommender Systems. *Ph.D. Thesis*, University of Minnesota, Minneapolis, MN, 2006.

McGormick, 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](http://celebrate.eun.org/eun.org2/eun/Include_to_content/celebrate/file/Deliverable7_2EvaluationReport02Dec04.pdf)

MELT, <http://info.melt-project.eu>, the public portal at <http://lreforschools.eun.org>

Millen, D., Yang, M., Whittaker, S., Feinberg, J. (2007), "Social bookmarking and exploratory search". In L. Bannon, I. Wagner, C. Gutwin, R. Harper, and K. Schmidt (eds.). *ECSCW'07: Proceedings of the Tenth European Conference on Computer Supported Cooperative Work*, pp. 24-28 September 2007, Limerick, Ireland

Mukherjee, R. and Mao, J. (2004), "Enterprise search: tough stuff", *Enterprise Search*, Vol. 2 No.4, pp. 36-46.

Najjar, J., Wolpers, M., Duval, E. (2007), "Contextualized Attentin Metadata - Personalised Access to Digital Resources", CAMA 2007 workshop report, *D-Lib Magazine*, Vol.13 No.9/10.

Puspitasari, F., Lim, E., Goh, D., Chang, C., Zhang, J., Sun, A., Theng, Y., Chatterjea, K., and Li, Y. (2007), "Social Navigation in Digital Libraries by Bookmarking", *Asian Digital Libraries, Looking Back*.

Rafaeli S., Dan-Gur Y., Barak M., (2005), "Social Recommender Systems: Recommendations in Support of E-Learning", *Journal of Distance Education Technologies*, Vol. 3 No.2, pp. 29-45.

Ryen W. White, Gary Marchionini and Gheorghe Muresan (2008), "Evaluating exploratory search systems: Introduction to special topic issue of information processing and management", *Information Processing & Management*, Vol. 44 No.2, pp. 433-436.

Santos-Neto, E., Condon, D., Nazareno, A., Ripeanu, M. (2009), "Individual and Social Behavior in Tagging Systems". In *20th ACM Conference on Hypertext and Hypermedia*. Torino, Italy.

Sen, S., Harper, F. M., LaPitz, A., Riedl, J. (2007), "The quest for quality tags", *Proceedings of the 2007 international ACM Conference on Supporting Group Work*, ACM, New York, pp. 361-370.

Tang, T., McCalla (2009), "The Pedagogical Value of Papers: a Collaborative-Filtering based Paper Recommender", *Journal of Digital Information*, Vol. 10 No.2.

Ternier, S., Massart, D., Campi, A., Guinea, S., Ceri, S., Duval, E. (2008), "Interoperability for Searching Learning Object Repositories: The ProLearn Query Language". *D-Lib Magazine*, Vol. 14 No.1/2.

Van Assche, F., Vuorikari, R. (2006), "A Framework for Quality of Learning Resources", U. Ehlers & J.M. Pawlowski (eds.), *European Handbook for Quality and Standardization in E-Learning*. Springer.

Vuorikari, R., & Van Assche, F. (2007), "Collaborative content enrichment in multilingual Europe, European Schoolnet's approach on educational resources". In *Workshop proceedings of AXMEDIS2007*. Barcelona, Spain.

Vuorikari, R., & Koper, R. (submitted). "Evidence of cross-boundary use and reuse of digital educational resources". doi: <http://dspace.ou.nl/handle/1820/1709>.

Vuorikari, R., Ochoa, X. (2009), "Tagging in the context of multiple languages", *Journal of Digital Information*, Vol. 10 No.2.

Vuorikari, R., Poldoja, H., Koper, R. (in print). Comparison of educational tagging systems - any chances of interplay? *International Journal of Technology Enhanced Learning*.

Ward A. Hanson and Kirthi Kalyanam (2007), *Internet Marketing and eCommerce*, Chapter 8, Traffic Building, Thomson College Pub, Mason, Ohio.

White, R., Marchionini, G., Muresan, G., (2008), "Evaluating exploratory search systems: Introduction to special topic issue of information processing and management", *Information Processing & Management*, 44(2), pp. 433-436.