Searching and Browsing Digital Library Catalogues: A Combined Log Analysis for The European Library

Maristella Agosti[1], Franco Crivellari[1], Giorgio Maria Di Nunzio[1], Yannis Ioannidis[2], Elefterios Stamatogiannakis[2], Mei Li Triantafyllidi[2], and Maria Vayanou[2]

[1] Department of Information Engineering – University of Padua
Via Gradenigo, 6/a – 35131 Padova – Italy
{agosti,crive,dinunzio}@dei.unipd.it
[2] Deparment of Informatics and Telecommunications – University of Athens
Panepistimiopolis, Ilissia – 15784 Athens – Greece
{yannis,estama,meili,vayanou}@di.uoa.gr

Abstract. The interaction between a user and a digital library catalogue can be analyzed and studied in order to gather user preferences and learn what the user likes the most and use this information to present the results in a personalized way. Since research on personalization of contents relies on actions performed by people, it is essential to identify user sessions in order to capture user intentions and preferences in a particular instant of time.
In this paper, we present a combined analysis of two different logs, action log and HTTP logs, of The European Library which is a free service giving access to resources of the 48 national libraries of Europe. A user study was conducted in order to collect into HTTP logs enough data to study the browsing activity and analyze possible relations between explicit preferences collected by online questionnaires and implicit actions recorded in the logs.

1 Introduction

Library catalogues, and in particular National library catalogues, are important sources for the preservation and diffusion of cultural material which can be stored on different media, for example books, files, pictures, and so on. The transformation of library catalogues into digital library catalogues has given the possibility to access library material remotely in an integrated way, and often includes a library catalogue module as the public interface to the system's inventory. The user is usually allowed to find items according to a set of choices, among them: author, title, subject.

The interaction between the user and the system can be analyzed and studied in order to gather user preferences and "learn" what the user likes the most, and use this information to present the results in different ways for each user. User preferences can be learned explicitly, for example asking the user to fill-in

questionnaires, or implicitly, studying the actions of the user which are recorded in the action log of a system. The second choice is certainly much less intrusive but requires more effort to reconstruct each search session a user made in order to learn his/her preferences.

The problem of identifying user sessions may become easier if authentication was mandatory for a user who wants to access a service. Authentication would provide the exact time of a session intended as a sequence of events like: login, actions and browsing, logout. However, authentication is not always required, especially when dealing with online free services. In these cases, since research on personalization of contents relays on actions performed by people, it is essential to identify user sessions in order to capture user intentions and preferences in a particular instant of time. Organizing the requests in a single session permits to have a better view of the actions performed by visitors. A procedure, named "session reconstruction", may be used in order to map the list of activities performed by every single user to the visitors of the site. A possible approach to isolate a single session of a user is the use of the pair IP address and user agent [1], and permits only a fixed gap of time between two consecutive requests [2]. This type of heuristic has been widely used in literature and works pretty well; it is important, however, to bear in mind that with this approach it is not possible to identify the specific instant in which a client leaves the site, therefore statistics like time spent on the Web site, or average of a session, are biased.

The main focus of this paper is to analyze the logs, for personalization purposes mentioned above, of The European Library (TEL)[1]. The European Library is a free service of the Conference of the European National Librarians (CENL)[2] for anyone world-wide seeking for library material. It was launched in 2005 to offer a common access point to the distributed resources of the 48 national libraries of Europe in 20 languages (for a comprehensive reference to The European Library and its history from project to operational service, please refer to [3]).

TELplus[3] is a project that is being taken forward by a consortium of 26 partners, national libraries and research centres to form a pool of research and innovation to provide value-adding services and products for The European Library. The project provides a major research and innovation hub for The European Library. Each work package aims at either significantly raising the amount of digital content in The European Library, or improving access to content and the overall usability of the service. Log file analysis can provide a rich source of information that can feed into improving the overall usability of the portal, creating particular services for specific user groups, offering true personalized search and retrieval according to a user individual profile [4]. The work here reported has been conducted in the context of Work Package 5, which is the work

[1] http://www.theeuropeanlibrary.org/
[2] http://www.cenl.org/
[3] http://www.theeuropeanlibrary.org/telplus/

package of the project devoted to the conceiving of user personalization services also through logs analysis[4].

The European Library portal does support authentication but in a soft way: a user can register and login if she wants, but she is free to use almost all the search services without authentication. However, there are pieces of information in The European Library we can use to partially reconstruct a session of a user: this information is stored in the cookie the The European Library Web server sends to the client when the first HTTP request is sent to the portal. In fact, the cookie contains an identifier, called TEL session identifier (or TELsessid), which should identify a session. Actually, this identifier stores information about user preferences and actions on a particular computer, the one a user is interacting with, and does not contain information about sessions. To clarify this point we describe some scenarios, supposing that nobody authenticates in the Web site:

- the user browses the TEL portal for a while, then closes the browser, leaves and reopens the browser for another search later. There are two distinct sessions but only one TELsessid;
- a user finishes his browsing activity and after a while another user starts to browse the portal on the same computer. There are two different users and sessions, but only one TELsessid;
- the user browses and performs some searches on the TEL portal, then for privacy issues he deletes the cookies stored in the browser and continues browsing. There is a single session, but two TELsessid are issued by the TEL Web server.

There could be many other examples of this kind, which demonstrates that this task is difficult and requires many different strategies to try to get close to the real session a user made.

To present the work conducted and the reached results on this difficult task the paper is organized as follows: Section 2 describes The European Library environment and the main results reached through log analysis and user studies are presented in Sections 3 and 4 respectively. Section 5 reports on results about the portal's general usage while specific focus is paid on the searching behavior in Section 6. Finally, collection related issues are investigated in Section 7 and significant correlations are highlighted, drawing a variety of conclusions. Final remarks are given in Section 8.

2 The European Library Environment

The European Library environment consists of three different components that must be analyzed together to produce a complete traffic overview: The European Library portal, action logs and HTTP logs.

In The European Library portal's home page, a user can initiate a simple keyword search with a default predefined collection list presenting catalogues from

[4] `http://www.theeuropeanlibrary.org/telplus/workplan.php`

national libraries. From the same page, a user may click *advanced search* link where she can enrich her query with Boolean operators and limit search to specific fields like author, language, or ISBN. Alternatively from the *simple search* page, she may change the searched collection by checking the theme categories below the search box. After search button is clicked the result page appears, where results are classified by collections and the results of the top collection in the list are presented with brief descriptions. We will call the presentation of result records description, a *result list* page. At this point, a user may choose to see *result lists* of other collections or move to the next page of records of current collection's results. While viewing a *result list* page a user may also click on a specific record to see detailed information about the specific record. Depending on the selected collection and record viewed, additional services may be available e.g. links leading to a library site. Additionally, while still being at the result page, a user can perform a new search or search within the current *result list*.

The European Library environment keeps detailed logs of users' actions within the portal. Action logs are stored in a table of a relational database, where a table record represents a user action. The most significant columns of the table are:

- *userid*: A numeric id, for identifying registered users or 'guest' otherwise.
- *userip*: User's IP address.
- *sessid*: An automatically generated alphanumeric, identifying sequential actions of the same user (sessions).
- *query*: Query contents.
- *action*: Name of the action that a user performed.
- *colid*: The corresponding collection's alphanumeric id.
- *date* : Date and time of the action's occurrence.

A detailed description of the possible values of action field is listed in Table 1. Action logs' analysis, presented in this paper covers the period from 1st January 2007 until 18th July 2008.

HTTP log files contain the records of all the HTTP requests made by clients to the Web server of The European Library. Here, we present the choice of those fields included in the log files we have analyzed:

- *date:* Date, in the form of yyyy-mm-dd.
- *time:* Time, in the form of hh:mm:ss.
- *cs-method:* The requested action. Usually GET for common users.
- *cs-uri-stem:* The URI-Stem of the request.
- *cs-uri-query:* The URI-Query, where requested.
- *c-ip:* The IP address of the client.
- *cs(User-Agent):* User-Agent of the Client. For a standard user this means the browser and other information about operating system.
- *cs(Referrer):* The site where the link followed by the user was located.

HTTP log analysis has been carried out on the data collected during the user studies described in Section 4.

Table 1: Description of The European Library logging actions and classification in action types.

Action type	Action value	Description
Search: all search related actions	search_sim	simple search
	search_adv	advanced search
	search_res	search initiated from results page
	search_url	search initiated from url query string
Result List Browsing: actions related to presentation of search result lists	view_brief	display result list page
	page_brief	navigation between result list pages using "next" or "previous" buttons
	jump_to_page	jump to result list page with user specified sequential number
Result clicking: actions indicating selection and view of an object's full record information	view_full	display/navigate between result record pages
Collection selection: actions to choose collections to be searched	col_set_theme col_set_theme_ country	*Theme collection selection*: Collections chosen from theme list (checkbox in homepage) or from country list (drop-down list in homepage)
	col_set_country col_set_subj col_set_desc	*Advanced collection selection*: Save link clicked after selecting collections in collections tab, that are browsed by country or subject or searched by description
	col_set_default	collections default list reinstantiated
Result retention: actions denoting a user's wish to retain information about a specified object.	option_save_ session_ favorite	result record saved in favorites
	option_send_ mail	sent by email result record
	option_print	print result record
	options_save_ reference	record saved for reference manager use
Outgoing: actions resulting in user redirection towards library sites or services	service_X	full record service link used of country X
	service_all	full record service link used
Show help file	show_help_X	displays help file X
Result clicking / Outgoing	available_at	"Available at Library" link clicked to view record in native interface
Result clicking / Outgoing	see_online	"See online" link clicked to see object in native interface

Table 2: Distribution of registered users declared profession.

Profession	Number of users	Profession	Number of users
Student - Undergraduate	522	Librarian - Administrator	93
Other	514	High School Teacher	65
Student - Postgraduate	304	Media/Journalist	51
Researcher	269	Librarian - Faculty Liaison	33
University Teacher/Professor	221	Librarian - Corporate	23
Professional/Practitioner	159	Bookseller	19
Librarian - Other	146	Total	2419

3 Registered Users

The European Library service offers users the opportunity to register and benefit from personalized services, e.g. registered users may store favorites for later sessions and save search sessions in their history. Until July 2008, the user registration procedure has been performed 2,419 times. However this number does not correspond to actual registered users, since there are firm indications, that some users have registered more than once. TEL's low registration numbers suggest that users are not motivated enough to register. Lack of user motivation is also supported by the registered user log-in frequency. In detail, approximately 84% of registered users have performed one single registered session since the beginning of 2007.

Apart from uniquely defining users, registration procedure constitutes also a valuable source of explicit user feedback. The optional and mandatory information, filled in the registration form, reveals the addressed user group characteristics. Approximately half of the registered users come from academic community and most of them are undergraduate or postgraduate students. The second most popular value chosen for the *profession* field is *Other* (21.2%), indicating that TEL is accessed by people, whose type of work is not reflected in choices available within the drop-down list.The distribution among the available professions is depicted in Table 2. Approximately half of the registered users come from the academic community and most of them are undergraduate or postgraduate students.

Also the results on The European Library portal's usage by registered users are reported in Section 5.

4 User Studies

Log data are not the only form of data which can capture user preferences; there are also explicit ways of gathering these preferences, such an example are user surveys. With the aim of gaining insights by analyzing log data together with

data from controlled studies, a user's study was conducted at the University of Padua, Italy, in order to study specific aspects of The European Library portal, focusing mainly on: parts of the Web site, clarity of the Web site, level of satisfaction.

This study was conducted in a controlled setting during the end of 2007 and the beginning of 2008, in the computer laboratories of different faculties of the university, with at least one person of the same university, present in the laboratory together with users that were all students of one of the faculties of the university. All students were attending one course that was taught by one person of the university involved in the TELplus project. This means that the students were all attending a course in some way pertinent to the contents of the questionnaire and the use of The European Library portal was coherent with the aim of the course's part that was taught at the time of the use of the portal and of the filling in of the questionnaire. In order to track the activity of each user and find him in the HTTP logs, we asked to change the user agent with a string which would have been easy to be identified in the logs. This identification played an important role because it gave us the possibility to reconstruct each user session in a clear way, without the use of heuristics.

At the end, a total of 216 students participated in these studies, but it was possible to clearly recognize in the logs only 151 of them. This is due to errors during the typing of the user agent which made it impossible to find user in the logs, or two identical user agents with overlapping time, in this case we preferred to discard this information so as to not introduce more bias.

It is important to underline that the sample of students cannot be considered as a significant sample of all the user of The European Library service; however, the results of the analysis on this group of students are useful to understand the behavior of undergraduate students of Humanities, specifically from Italy, and in particular from University of Padua. These are users who can be interested in using the portal and searching of bibliographic records and, for this reason, their judgements about the interface and the services have to be seriously taken into consideration for future improvements.

5 General Usage

In this Section, we analyze The European Library portal's general usage from the following points of view: action logs, HTTP logs, and HTTP Logs of User Studies.

5.1 Action Logs

The average session duration is estimated to approximately 6 minutes while median duration is limited to 2. In terms of actions, the average session length is almost 8 actions and median is 4 actions. Comparing with registered user sessions, where the median session duration reaches almost 5 minutes and median

Table 3: Number of actions per action category.

Action type	#	Type of action	#
Search	584,687	Outgoing actions	55,516
Result list browsing	478,585	Collection selection	200,009
Result clicking	626,007	Show help files	2,196
Result retention	13,154	Total actions	1,916,134

length is 8, one can safely conclude that registered users' sessions indicate a higher level of users' expertise than in the general case.

We have thematically classified actions into 7 categories, namely:

- *Search*,
- *Result list browsing*,
- *Result clicking*,
- *Result retention*,
- *Outgoing actions*,
- *Collection selection*, and
- *Show help files* actions.

Table 3 depicts the frequency of each action type. It is worth noting that some actions belong to more than one category, hence the sum of actions in Table 3 exceeds the total number of logged actions. The category label assigned to each individual action is reported in the *Action type* field of Table 1.

Ideally actions of type *result list browsing* should follow *search* actions, being at least the same in number as *Search* actions. The discrepancy's explanation is that users abandoned their searches before The European Library service returned any search results, probably due to The European Library service response time.

5.2 HTTP Logs

The HTTP log data were collected over the period of time which goes from January 2007 to June 2008. Table 4 reports descriptive statistics computed on the collected HTTP log data.

As reported in Table 4, in the analyzed period of time the number of unique visitors was more than 475 thousands, with an average of 870 visitors per day. More than 700 thousands visits where recorded with a daily average of about 1,300; these visits produced around 58 millions of hits (HTTP requests to the Web server) which corresponds to a daily average of 106 thousands of contacts. The number of Web pages requested were more than 19 millions, which corresponds to almost 35 thousands of pages per day. "Not viewed" traffic, both hits and bandwidth, includes traffic generated by robots, worms, or replies with special HTTP status codes.

Table 4: HTTP Log Synthetic Descriptive Statistics.

	Total	Daily Average	Per visit
No. of visitors	475,333	868.98	Not applicable
No. of visits	709,922	1,297.85	1.49 (per visitor)
Hits	58,053,293	106,130.3	81.77
No. of accessed pages	19,047,263	34,821.32	26.837
Bandwidth	1,059.51 GB	1.94 GB	1.53 MB
Not Viewed Hits	14,229,379	106,130.3	20.04
Not Viewed Bandwidth	2,876.63 GB	5.26 GB	4.15 MB

The two plots reported in Figure 1 show the trend of the total number of visitors of The European Library for each month of the period under analysis. In the first plot, the difference between the maximum value (January 2008) is reported, while in the second plot the numbers of visitors are reported.

If we consider the length of the sessions of people who connected to The European Library, we can see that the majority of the sessions last no more than two minutes. The distribution of the length of sessions can be summarized with the Pareto diagram reported in Figure 2.

The analysis of log data for each month, or even a shorter period like a week or a day, is fundamental for understanding specific and punctual analysis on a particular variable of interest, for example the study of the actions of a particular group of users during the day of a week. On the other hand, aggregated data are more suitable for presenting the results of long periods of analysis. In order to analyze the trend of the traffic of one year and a half of accesses to The European Library portal, we decided to aggregate the log data in group of three months to follow the possible seasonal variations. An example of this type of analysis is reported in Figure 3 where the variation of seasonal trends of the length of sessions is reported. The great majority of visits do not last more than a couple of minutes. It is interesting to study the trend of the percentage of visits duration compared to the total.

5.3 HTTP Logs of User Studies

We have analyzed the HTTP requests made by the 151 students of the University of Padua which have been recognized in the HTTP logs. Given the fact that this was intentionally a controlled experiment the number of sessions is obviously equal to the number of students. The analysis carried out were the following: statistics on the number of requests made by each user, statistics on the length of a session in terms of filling-in the questionnaire and of browsing the portal.

In Table 5 a summary of the statistics of the number of records is presented. Figures show that the distribution of these values is not symmetric: values are skewed to the lowest values, the median, is in fact closer to the first quartile. The mean, a measure highly affected by extreme values, is equal to 228.3 requests, this happens because there are a number of users who performed a high number

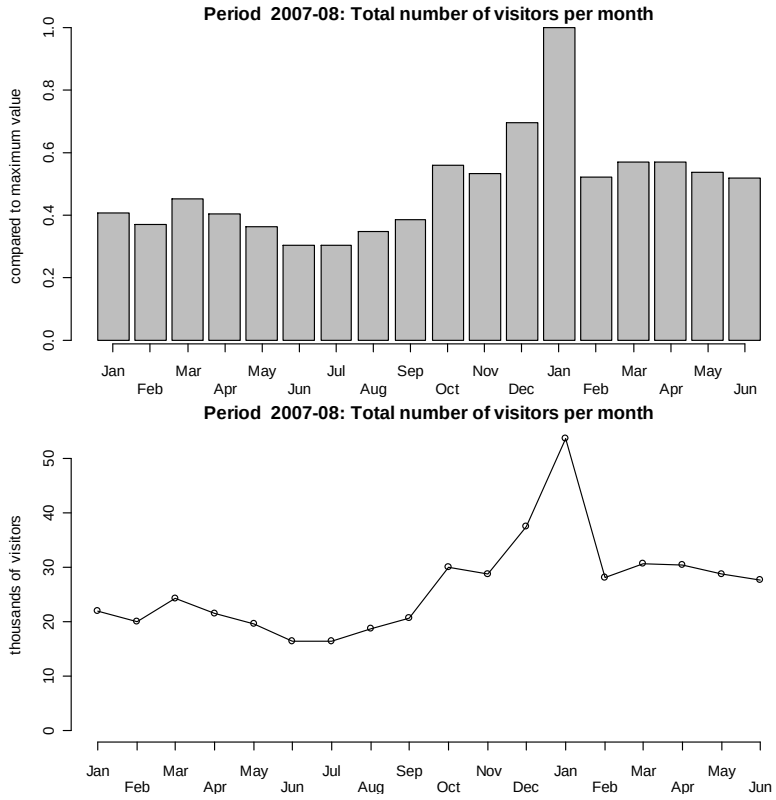

Fig. 1: Number of visitors per month.

of requests per session. The maximum number of requests was 866, while the minimum was 4 corresponding to a student who, at the end of the questionnaire, cleaned the cache of the browser and incidentally reopened the page of The European Library portal, which produced a new TEL session identifier and a sequence of 4 HTTP requests.

HTTP requests can also be divided in classes of frequencies to study modes of the distribution. It is interesting to note that about 60% of the HTTP sessions are classified sessions with less than 250 requests. If we consider that each single click on a link of a Web page produces multiple HTTP requests, this result means that the majority of sessions are very short in terms of browsing activity. This behavior reflects what has been found in previous studies on The European Library portal usage: sessions tend to be short and with few requests [5].

A study on the relation between the time a user spent to compile the questionnaire during the survey and the time a user spent on browsing the portal was carried out. For the first measure there were time constraint due to the available time slots of University computer labs. In Table 6 the statistics for this time

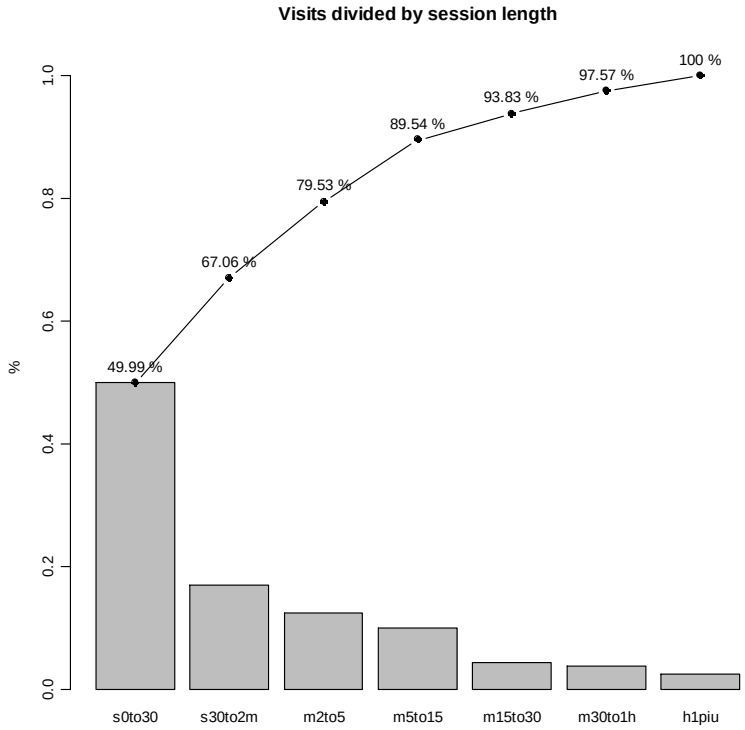

Fig. 2: Distribution of session lengths with a Pareto diagram.

are shown. The time is clearly bounded between 30 and 45 minutes which were the usual slots. In the same table, we also reported the number of requests per minute.

In Table 7 another statistic is shown, the one relative to the time spent for browsing calculating as the difference between the first and the last HTTP request and the average number of requests per minute. The numbers are similar to those calculated for the questionnaire in Table 6, this means that the students browsed the portal while filling in the questionnaire. These results are interesting when compared to the average time of a search session presented in Section 5.1 which presented a median of about 2 minutes and an average of 6 minutes.

6 Searching in The European Library

The European Library service provides users with two interfaces for formulating queries, namely simple search and advanced search. Simple search is similar to a traditional web search interface, where the user inputs terms separated

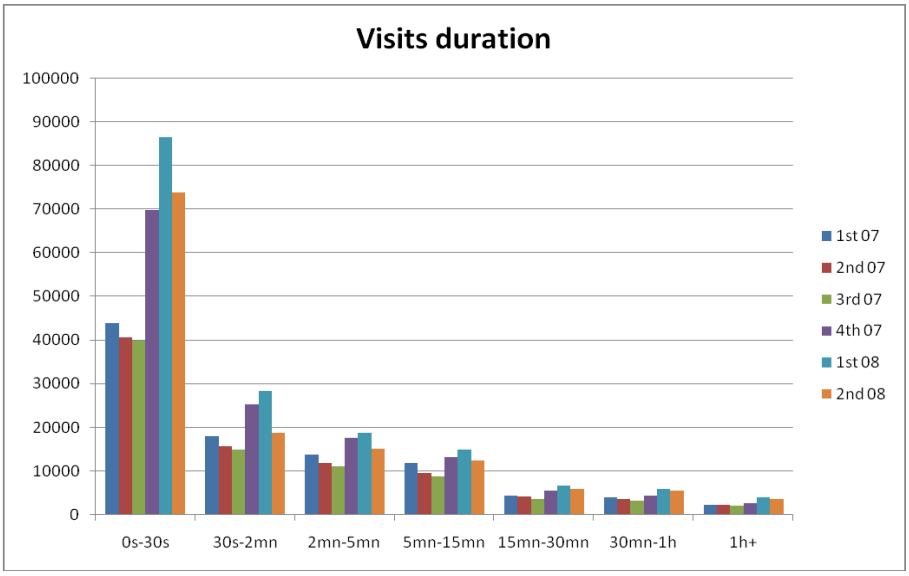

Fig. 3: Visit duration grouped quarterly.

Table 5: Summary of statistics for the number of HTTP requests per user.

Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
4.0	99.0	175.0	228.3	311.5	866.0

by space, having an implied conjunction between them. In advanced search, The European Library service provides attribute fields (Title, Creator, Subject, Type, Language, ISBN, ISSN) for query formulation, along with a set of Boolean operators (AND, OR, NOT). A user can chose from a drop-down list specific fields that should be searched for, meeting conditions provided as user input next to the list. We refer to queries containing operators or attribute fields as advanced queries, discriminating them from simple keyword queries.

A session in average contains 2.3 queries, having an average query length of 2.6 terms per query.

Approximately 50% of overall logged queries are unique queries while 10% of the query terms have occurred only once, thus emphasizing the need to handle queries at a term level, so as to efficiently leverage past users' search interactions. Furthermore, approximately 3% of simple keyword queries are empty.

Table 8 shows syntactic elements found in The European Library *advanced queries*, alongside their number of occurrences. Noteworthy is that *advanced* queries constitute 13.2% of all queries, firmly suggesting user preference for simple keyword search.

Turning to keyword queries, Table 9 depicts the frequencies of most popular queries alongside frequencies of most popular search terms.

Table 6: Summary of statistics for the time spent to fill-in the questionnaire during the user survey (first row), and the number of requests per minute (second row).

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
minutes	7.0	28.0	31.0	33.55	40.0	60.0
requests/min	0.14	2.85	5.56	7.43	9.92	56.0

Table 7: Summary of statistics for the time spent for browsing the portal (first row), and the number of requests per minute (second row).

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
minutes	0.03	25.58	30.25	31.80	38.90	57.38
requests/min	0.28	3.05	6.17	8.35	11.12	120.00

7 Collection Selection in The European Library

Since The European Library catalogues contain more than 300 collections, providing users with collection selection options is a necessity.

We have classified collection selection options in the following categories (descriptions of related actions are shown in Table 1):

- *Default collection selection* category consists of sessions not containing collection selection actions, thus using TEL's default collection selection. Overall, 65% of sessions fall in this category.
- *Theme collection selection* category, consists of sessions containing only *theme collection selection* actions. 33% of sessions follow the aforementioned collection selection pattern.
- *Advanced collection selection* category consists of sessions containing at least one *advanced collection selection* action. Overall, 2% of sessions are in this category.

Extensive use of theme collection selection feature is justified by its simplicity, requiring minimum user effort while being accessible within the search page.

Investigating collection selection methods other than those already provided by The European Library, we correlated nationalities of collections accessed by users with their geographic location (using IP geolocation). Figure 4 depicts a representative sample of this distribution. Strong preference of users towards collections having the same nationality as their location, is evident. Another noteworthy user behavior pattern is that users in Spain, Italy, France and Canada favor French collections against German ones, while users in Germany, Poland, Hungary and Croatia have the exact opposite behavior.

8 Conclusions

Studies on log files are essential for personalization purposes, since they implicitly capture user intentions and preferences in a particular instant of time.

Table 8: Frequencies of The European Library syntactic elements.

Syntax	#	Syntax	#
AND	26229	Type	880
OR	1268	Language	3722
NOT	415	ISBN	4695
Title	37024	ISSN	1171
Creator	31029	Advanced queries	77057
Subject	7002	Total queries	583422

Table 9: Frequencies of most popular queries (first column) and most popular terms (second column).

Query	#	Term	#
mozart	6634	mozart	7438
meisje met de parel	2001	history	2739
harry potter	1873	european	2549
van gogh	1730	journal	2349
pink floyd	717	potter	2308
nuremberg	657	harry	2271
rembrandt	609	international	2065
einstein	563	meisje	2032

The study carried out and presented in this paper took into account two different kinds of log files: action logs and HTTP logs. Action logs allowed to track registered users of the portal and, more in general, analyze search preferences. A user study was conducted in order to collect into HTTP logs enough data to study the browsing activity and analyze possible relations between explicit preferences collected by online questionnaires and implicit actions recorded in the logs.

In general, user sessions are short, on average less than ten minutes according to sessions reconstructed from both action logs and HTTP logs. There are at most a couple of search actions with a limited browsing activity of the results and of the Web site itself.

A possible comment to the length of a session could be that a user, that visits the portal for the first time, most of the times starts to conduct a Google-type search using the search form. Since the answer from the portal is different from a Google-type answer, the user goes away and never returns (or most of the times he does not return). This is in part reflected by the answers collected by questionnaires during user studies, where users were not satisfied by the presentation of the results and by the content of the results.

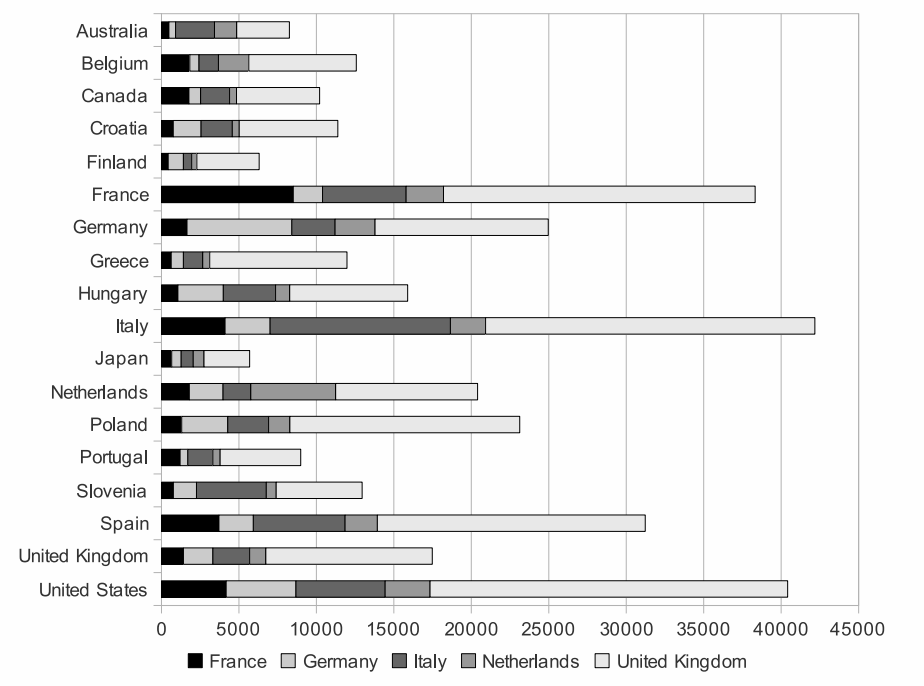

Fig. 4: Distribution of TEL traffic, group by originated country (country extracted from user IP geolocation) on 5 of most popular collection countries.

Some other evidences captured by log files is that there is a strong preference of users towards collections having the same nationality as their location. At the same time, users are willing to browse and search collections in different languages.

Acknowledgements

The authors would like to thank the staff of The European Library Office for their systematic support during the analysis of logs which is reported in the paper.

The work reported has been partially supported by the TELplus Targeted Project for digital libraries, as part of the eContent*plus* Program of the European Commission (Contract ECP-2006-DILI-510003).

References

1. Nicholas, D., Huntington, P., Watkinson, A.: Scholarly Journal Usage: the Results of Deep Log Analysis. Journal of Documentation **61**(2) (2005) 248–280

2. Berendt, B., Mobasher, B., Nakagawa, M., Spiliopoulou, M.: The Impact of Site Structure and User Environment on Session Reconstruction in Web Usage Analysis. In Zaïane, O.R., Srivastava, J., Spiliopoulou, M., Masand, B.M., eds.: Knowledge Discovery on the Web (WEBKDD) - MiningWeb Data for Discovering Usage Patterns and Profiles. Volume 2703 of Lecture Notes in Computer Science., Springer (2003) 159–179
3. Cousins, J., Chambers, S., van der Meulen, E.: Uncovering cultural heritage through collaboration. Int. J. on Digital Libraries **9**(2) (2008) 125–138
4. Angelaki, G.: Research and Innovation in The European Library; the TELplus Project. Globalisation and the Management of Information Resources `http://slim.emporia.edu/globenet/Sofia2008/`, Sofia, Bulgaria (Nov 2008)
5. Agosti, M., Di Nunzio, G.: Web Log Mining: A study of user sessions. In: Proc. 10th DELOS Thematic Workshop on Personalized Access, Profile Management, and Context Awareness in Digital Libraries (PersDL 2007), Corfu, Greece (Jun 2007)