Thirty Years of Digital Libraries Research at the University of Padua: The User Side

Maristella Agosti, Giorgio Maria Di Nunzio, Nicola Ferro, Maria Maistro, Stefano Marchesin, Nicola Orio, Chiara Ponchia and Gianmaria Silvello

University of Padua, Italy
{name.surname}@unipd.it

Abstract. For the 30th anniversary of the Information Management Systems (IMS) research group of the University of Padua, we report the main and more recent contributions of the group that focus on the users in the field of Digital Library (DL). In particular, we describe a dynamic and adaptive environment for user engagement with cultural heritage collections, the role of log analysis for studying the interaction between users and DL, and how to model user behaviour.

1 Introduction

Traditionally, Digital Libraries (DL) are considered places where information resources can be stored and made available to end users, but in the last decade they have also become systems that need to support the user in different information centric activities. Indeed, in the context of DL we need to take into account heterogeneous information sources with different community background such as libraries, archives and museums; but, DL are central also for research purposes and they provide the infrastructures able to gather, manage and grant access to scientific data at large. When it comes to interacting with the system for discovering, retrieving, re-using or citing cultural heritage objects or scientific data, users are the main focus for DL systems.

In the last thirty years, the Information Management Systems (IMS) research group at the University of Padua has contributed substantially to the DL field with special attention to the role of users. In this paper, we report the more recent and relevant contributions that focus i) on user engagement with cultural heritage collections in the context of the CULTURA project (Section 2), ii) on the role of log data analysis in the context of multilingual DL and search engines (Section 3), and iii) on user behaviour modelling (Section 4).

2 Features of Adaptivity in the CULTURA Environment

The main goal of CULTURA, a European project co-funded under the 7th Framework Programme which ran from 2011 to 2014[1], was to increase user

[1] http://www.cultura-strep.eu/

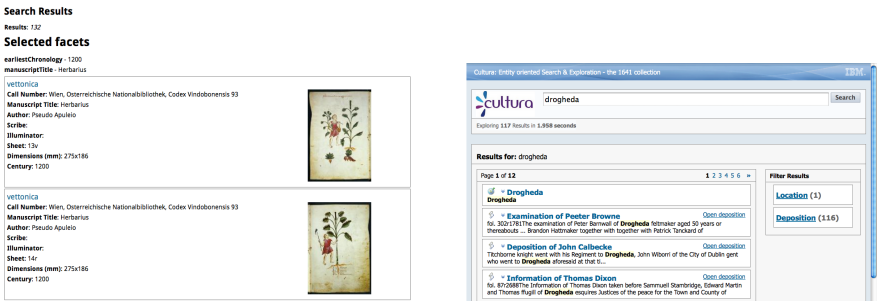
engagement with digital cultural heritage collections through the development of a new adaptive and dynamic environment, that is a Virtual Research Environment (VRE), and specifically developed tools. The CULTURA consortium had a strong emphasis on meeting real end-user needs, maximizing societal impact and laying a foundation for successful commercialization. To this end, the environment went beyond the traditional search-based exploration, providing natural language processing technologies, entity-oriented search and a comprehensive set of logging, bookmarking and annotating tools that make it a powerful aid to both extensive and intensive work on content collections. For the validation of the project, two pre-existing cultural heritage collections were used: The *1641 Depositions*, a collection of accounts by victims of the Irish Rebellion of 1641, which constitutes a textual corpus that has been augmented by manually generated metadata[2]; and *IPSA*, a collection of illuminated scientific manuscripts (including herbals and astronomical-astrological codices), which is a purely visual collection with extensive metadata[3].

2.1 Narratives

Narratives have been introduced in the CULTURA environment as a novel tool to engage the different types of users [10]. Narratives are implemented as threads through the document collection, linking artifacts and tools related to a particular topic. Expert researchers, guided by the use cases and user requirements outlined during user consultations, designed narratives. They used their specialist knowledge of the collections to create a series of paths through the content that can engage users from all the groups in the exploration and use of specific content. To address the different level of expertise of each category of users, each narrative has a number of levels. Less expert users are offered a relatively high level narrative, but as users interact with the resources that are presented to them, the system dynamically discloses additional material, resulting in a more complex and more captivating user experience. These narratives allow for an open-ended and developing engagement with the resource collections. An example of a high level narrative are the ones designed for less expert users. These describe the different steps of a short course on a specific topic that is encompassed within the collection. Typically, the relevant material will be spread across the diverse parts of the different components of a collection. Adaptive narratives provide structured routes through the collection, exposing the user to artifacts that are relevant to their topic of interest. Furthermore, at certain steps of the narrative, users are given the possibility to access some extra steps which will provide them with additional information on that particular topic, also using external resources such as Wikipedia. Once they have gone through all the additional steps, users will come back to the point where they left the narrative and will be able to proceed.

These narratives were implemented for both collections in the environment, and the user experience of these narratives validated their usefulness for both

[2] http://1641.tcd.ie/
[3] http://ipsa.dei.unipd.it/



(a) (b)

Fig. 1: Faceted search and Entity-oriented search.

types of content collection. In the case of the historical textual corpus, a series of narratives were developed for use with secondary school students who were encountering the collection for the first time. These users were presented with a sidebar containing a brief explanation of the context of the individual part of the text being viewed, along with prompts for further research. Users can move backwards or forwards within the narrative, or branch off into further narratives, covering particular areas that especially pique their interest. At any point, the user can leave the pathway to carry out their own detailed investigations. Users are able to resume the narrative when desired. In addition to tying together chains of documents, these narratives can include the other tools that contribute to the environment. For instance, a user of the IPSA collection who is following a narrative based on herbals of a given historical period can be presented with the results of a faceted search based on chronology and title (see Figure 1a) or a user who is following a narrative based on a particular location can be presented with the results of an entity-oriented search displaying the relationship of the entity in question to other entities within the collection (see Figure 1b). The visualization features are also available to users of the textual collection, and, in a similar way, allow the collection to be explored on the basis of the links and relationships between people, places and events. Individually, these tools offer a range of exciting new ways of looking at the collection.

The narratives add a further layer of richness, integrating these views within a coherent structure. Narratives help in fostering a more involving experience of the collections, more complete and extended through time. Therefore, other kinds of static collections can benefit from narratives too, e.g. archaeological collections.

2.2 An Adaptive Cross-Site User Modelling Platform for Cultural Heritage Websites

Websites need information about their users in order to deliver personalized experiences [20, 23]. Previous activity across the Web provides an opportunity

to gather further data about the user from beyond the website, enhancing user profiles with overarching information from different websites. In cross-site user modelling, web personalization techniques are based upon information provided by an independent third-party user modelling service, in order to assist users in addressing information needs.

In the domain of cultural heritage, this approach can provide relevant information to users that attempt to answer cross-site information needs within their browsing space. In those situations where the user's need spans topics that are not confined to a single website, the introduction of such cross-site service [24] might improve the effectiveness of the website personalization, along with the user's level of satisfaction. In [9], a parallel and complementary approach to the Virtual Research Environment (VRE) for the Digital Humanities of the CULTURA[4] project [32, 33] has been proposed. Introducing the cross-site service, the user model can gain a higher precision in those topics that are more relevant to the user and provide more tailored personalization to help addressing the user's cross-site information needs. An example of this could be an overarching user's interest in the living conditions of the Irish middle class during the Irish rebellion, which cannot be addressed by a single website of the CULTURA VRE but requires the user to navigate across different websites of the VRE browsing space.

The general process that identifies the cross-site approach is as follows: (1) The user lands on a website related to an information gathering task in the cultural heritage domain; (2) The user authenticates a first time with the third party service; (3) The website tracks all user activities in the webpages along with the relevant text entities identified by a term identification component; (4) The user triggers the information exchange function, in anticipation of subsequent personalisation by the target website, which provides relevant user data (based on the selected communication pattern) to the website and newly tracked user information to the service; (5) The user surfs to a second website and, depending on whether they are already authenticated or not, authenticates or directly triggers the information exchange function, which should provide more tailored information to the website; (6) Steps (1)-(5) can then be re-iterated many more times, without a strict order of execution. The high-level architecture of the cross-site service is presented below and can be found in Figure 2.

Term Identification Component: The main purpose of the term identification component is to identify text entities related to the body of the current webpage the user is viewing. Text entities indicate the meaning of the underlying content from websites belonging to the cross-site browsing space. An additional responsibility of the term identification component is to ensure the creation of a shared conceptualization of the user's cross-site browsing space. This conceptualization is represented as a text entity space and it is based on the contents the user has browsed within the websites of the cross-site browsing space.

[4] `http://www.cultura-strep.eu/`

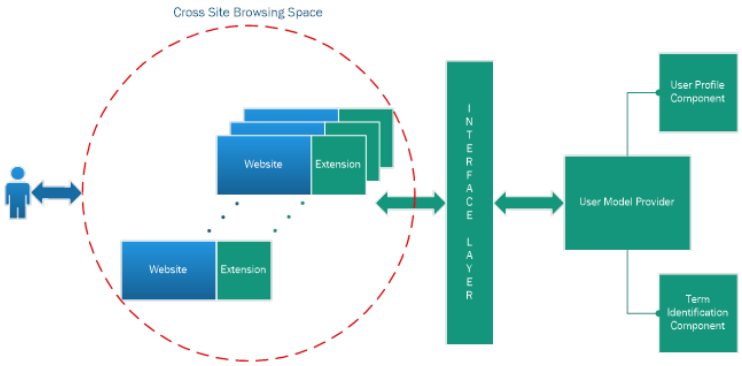


Fig. 2: High-Level Service Architecture.

User Profile Component: The user profile consists of activities and text entities related to the current task of the user and that were identified by the term identification component.

Interface Layer: The interface layer provides an abstraction from the specific implementation of the different websites within the user's cross-site browsing space. It implements a RESTful API[5] that facilitates the communication between the cross-site service and the websites the user is browsing.

WCMS - Web-based Content Management System Module Extensions: WCMS module extensions allow a simple and limited-impact integration of cross-site information exchange techniques into existing website implementation. The responsibility of the WCMS module extensions is twofold: (1) To facilitate communication between the website and the API of the cross-site user modelling platform and (2) to provide non-intrusive information exchange techniques to the user, within the website the user is currently browsing.

Cross-Site Browsing Space: The application of information exchange techniques to independently hosted websites introduces a cross-site browsing space. Within the browsing space, target websites receive user data through information exchange techniques and return novel user's data to the cross-site service through the same mechanism.

3 Log Analysis

The interaction between users and information access systems can be analyzed and studied for different goals, for example to personalize the presentation of results. User preferences can be learned either explicitly with surveys or questionnaires, or implicitly by studying the actions that the user performs when using the system. These actions are saved on log files that can be used to study the usage of a specific application, and to better adapt it to the objectives the

[5] `http://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm`

users were expecting to reach. The analysis of log data can be roughly divided according to the nature of the system, as suggested by [4], into: Web search engines (WSE) log analysis and Digital Library Systems (DLS) log analysis. In fact, both WSE and DLS can be used as information access systems accessible through the Web; however, the collections of documents the user can access are drastically different: WSE retrieve documents that are in general Web pages, while DLS retrieve documents that have been chosen after a quality control performed by professionals.

3.1 The European Library (TEL)

The European Library[6] originates from the TEL project that was partly founded by the European Commission and successfully ended on January 2004. A new project called TEL-ME-MOR[7] was born with the purpose to expand The European Library in order to support the ten national libraries from the New Member States of the European Union and later to offer access to the resources of 45 European National Libraries.

In [7], a general methodology for gathering and mining information from Web log files was proposed and a series of tools to retrieve, store, and analyze the data extracted from Web logs was designed and implemented. The aim of this work was to form general methods by abstracting from the analysis of the Web logs of TEL in order to give advice about the development of the portal from the point of view of both the security and the improvement of personalization. The solution proposed by this methodology for the problem of storing the information contained in a log fie was the use of a DataBase Management System (DBMS) in such a way that it was possible to perform queries easily, for example to obtain statistical information useful for the development of the site and to allow subsequent mining of the managed data. Figure 3 presents the Entity–Relationship conceptual schema that was designed.

In [8], an experimental analysis was performed on the available log data at that time, corresponding to eleven months of TEL Web log files from October 31st 2005 to September 25th 2006. In this analysis, we used heuristics to identify users and, as a result, we suggested that authentication would be required since it would allow TEL servers to identify users and create profiles to tailor specific needs. Moreover, authentication would have also helped to solve the problem concerning crawlers accesses, granting access to some sections of the Web site only to registered users, blocking crawlers using faked user agents. A second batch of seven months of TEL Web log files, from October 1st 2006 to April 30th 2007, was analyzed in [1]. This new set of log files contained richer information about: the Internet Protocol (IP) address and the user-agent which allowed the identification of single users, as well as the referrer field, a Uniform Resource Locator (URL) address which communicates the last page viewed by the user which could be used to know the way visitors get to TEL service. These

[6] `http://www.theeuropeanlibrary.org/`
[7] `http://telmemor.net/`

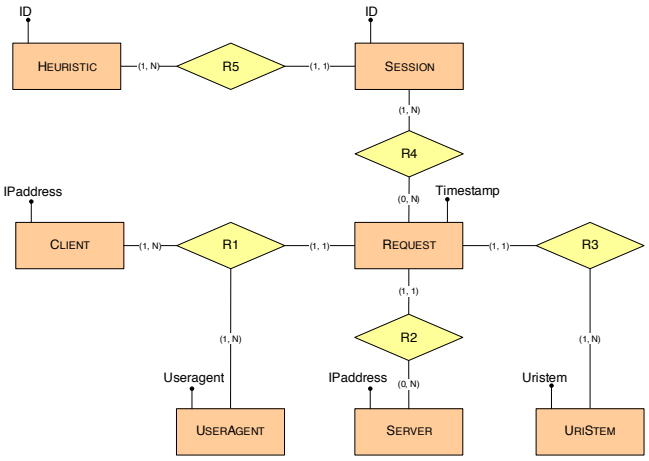


Fig. 3: The European Library Log Data Storage Entity-Relationship Scherma. Rectangles show the main objects (entities) of the data contained in the log files. Diamond shapes describe the relationships among the objects.

logs also contained the information saved by the cookie file such as: the language selected by the user during the navigation of the service; the collections of documents selected during the query or query refinement, the identifier of the session assigned by the server to a specific user. These richer data allowed us to better understand the profile of the TEL users in terms of geographical areas, willingness to spend some time to perform advanced searches and to select collections of documents different from the default ones. Further analysis to combine the results of the HTTP logs analysis with that of the user study to give a better understanding on the usage of The European Library service were presented in [3, 6]. In this work, 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. One interesting finding of this implicit and explicit data analysis was that short sessions found in the HTTP logs could be explained by the answers collected by questionnaires during user studies; in particular, users were not satisfied by the presentation of the results and by the content of the results. The sources of data used and the interrelation among them are depicted in Figure 4 where an example of a user session is drawn.

Implicit data on user interaction with a new Digital Library portal were useful for detecting typical usage patterns of different user groups, as well as their possible evolution or stability over time. They provided key information for identifying possible interaction problems on a large-scale (e.g. the whole user population), but little or no insight as to why these problems occurred and how they could be solved, as discussed in [5]. They undoubtedly were influential in the design of more focused survey studies (based on explicit user data) able to

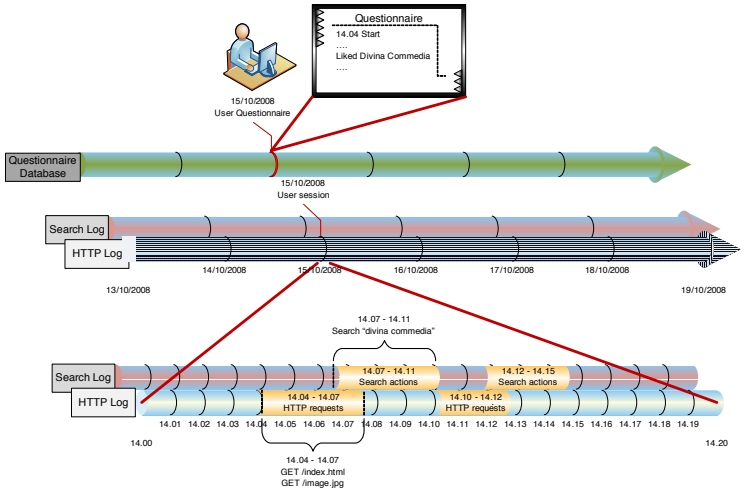


Fig. 4: The three arrows represent the three sources of data that are collected and combined together by the proposed method. An example of a user session is shown to highlight the way the three sources are generated in an interleaved way during the user activities.

investigate which relevant user expectations, habits, motivations, preferences or difficulties were responsible for the interaction patterns observed. Some of the insights gathered in these surveys were used as requirements for the design of an advanced query suggestion tool focussed on and tailored to TEL named i-TEL-u. While traditional query suggestion tools could not recognize different contexts from which the suggested queries are produced, i-TEL-u leveraged a variety of data sources and allowed users to seamlessly move from one context to another according to their evolving information needs during the search session [2].

3.2 Multilingual Log Analysis - LogCLEF

Despite being a very important resource to study user information needs and preferences, very few log datasets have been made available for research experiments. This lack of log data have made the verifiability and repeatability of experiments very limited since it is difficult to find two researches on the same system (for example on the same search engine or digital library), as well as using the same period of time to make results comparable across different studies. In the context of the Cross-Language Evaluation Forum (CLEF) Initiative[8], a self-organized body whose main mission is to promote research, innovation, and development of information access systems with an emphasis on multilingual and multimodal information with various levels of structure, LogCLEF was

[8] `http://www.clef-initiative.eu`

an evaluation initiative for the analysis of queries and other logged activities used as an expression of user behaviour [17,25]. An important long-term aim of the LogCLEF initiative was to stimulate research on user behaviour in multilingual environments and promote standard evaluation collections of log data. From 2009 until 2011, LogCLEF released collections of log data with the aim of verifiability and repeatability of experiments [16,27,28]. In the three years of LogCLEF editions, different data sets have been distributed to the participants: search engine query and server logs from the Portuguese search engine Tumba![9] and from the German EduServer[10]; digital library systems query and server logs from The European Library (TEL) and Web search engine query logs of the Chinese search engine Sogou[11].

LogCLEF promoted different tasks such as:

- Language identification task: participants are required to recognize the actual language of the query submitted.
- Query classification: participants are required to annotate each query with a label which represents a category of interest (for example, geographic entities, historical events, people, etc.).
- Success of a query: participants are required to study the trend of the success of a search. Success can be defined in terms of time spent on a page, number of clicked items, actions performed during the browsing of the results;
- Query refinding: when a user clicks an item following a search, and then later clicks on the same item via another search;
- Query refinement: when a user starts with a query and then the following queries in the same session are a generalization, specification, or shift of the original one.

4 User-System interaction

The huge increase in the number of documents published after WWII called for the design and development of new and fully automatic techniques to handle and search this increasing amount of information content. Web search [15], biomedical search [30], patent retrieval [26], and enterprise search [13], as well as DL, started to rely on Information Retrieval Systems (IRS) to perform automatic indexing and searching. However, with Internet becoming pervasive and accessible to everyone, IRS have begun to exploit more and more complex techniques to satisfy the user information need. Among them, particularly advantageous was the use of click log data to adapt systems to users behaviour.

Nowadays, it is a pretty common practice for commercial search engines, to record user interactions with their interface, such as query keywords, clicked URLs, impressions and clicks timestamps and many others. Click log data offer several advantages [22]: they are easy and inexpensive to collect, they are

[9] `http://www.tumba.pt/`
[10] `http://www.eduserver.de/`
[11] `http://www.sogou.com`

available in real time and they are user centered, i.e. they directly represent user preferences. However, even if click data have proven to be a valuable resource of implicit feedback, they are biased and noisy and they are intrinsically difficult to interpret [21]. Moreover, the user behaviour is tightly related to the information need and task that she is performing. As an example, consider that solely in Web search user queries can be classified as navigational, when the user searches for a particular website that she has in mind, and Informational, when the user explores the result page looking for some information on a specific topic [11]. When the scope is widened and heterogeneous tasks are considered, then the differences in user behaviour will be even more pronounced, as presented in [31] where Web, job and talent search are compared. In [31] a thorough analysis of log data is conducted, showing that substantial differences in Web, job and talent search can be detected, when considering the click frequency for each rank position, query frequency and query diversity.

The heterogeneity of tasks, information needs and users require IRS to adjust their output by considering the context in which the search is performed. This can be achieved by modelling the user behaviour in order to remove the noise and bias incorporated in log data. In [18] a new user model based on Markov chain is presented. Differently from other state of the art models [14,29], the exploitation of Markov chains makes it possible to describe a user who scans the ranked list of documents according to possibly complex paths. Therefore, the model can handle users that go forward and backward, jump from one document to any other in the list and visit the same document multiple times.

Furthermore, the proposed model can be embedded in IRS to account for user interactions. In [19] the Markovian model is used to describe the user dynamic, which is successively integrated in the *Learning to Rank (LtR)* algorithm LambdaMart [12,34]. The proposed approach stems from the observation that users behave differently depending on the query type, i.e. navigational vs. informational, and two different dynamics are calibrated on a click log dataset to resemble the real user dynamic. The algorithm accounts for the user dynamic as discounts for the objective function, assigning different discounting accordingly to the query category.

Finally, personalized IRS require involvement of the user even during the evaluation process. Indeed, if the system returns a different output depending on the user and the task, even the evaluation measure should consider these features. The Markovian model proposed in [18] defines a new family of evaluation measures, called *Markov Precision (MP)*, which injects user models into precision. MP can be considered as an offline measure, by using predefined transition matrices, or as an online measure, by tuning the transition matrix on click log data. In the latter case, the measure attempts to face the challenge of adapting evaluation to the user behaviour.

References

1. Agosti, M., Angelaki, G., Coppotelli, T., Di Nunzio, G.M.: Analysing HTTP logs of a european DL initiative to maximize usage and usability. In: Proc. of the 10th

International Conference on Asian Digital Libraries, ICADL 2007, pp. 35–44 (2007)
2. Agosti, M., Cisco, D., Di Nunzio, G.M., Masiero, I., Melucci, M.: i-tel-u: A query suggestion tool for integrating heterogeneous contexts in a digital library. In: Proc. of the 14th European Conference, ECDL. pp. 397–400 (2010)
3. Agosti, M., Crivellari, F., Di Nunzio, G.M.: Evaluation of digital library services using complementary logs. In: Proceedings of the Workshop on Understanding the User - Logging and Interpreting User Interactions in Information Search and Retrieval (2009)
4. Agosti, M., Crivellari, F., Di Nunzio, G.M.: Web log analysis: a review of a decade of studies about information acquisition, inspection and interpretation of user interaction. Data Min. Knowl. Discov. 24(3), 663–696 (2012)
5. Agosti, M., Crivellari, F., Di Nunzio, G.M., Gabrielli, S.: Understanding user requirements and preferences for a digital library web portal. Int. J. on Digital Libraries 11(4), 225–238 (2010)
6. Agosti, M., Crivellari, F., Di Nunzio, G.M., Ioannidis, Y.E., Stamatogiannakis, E., Triantafyllidi, M.L., Vayanou, M.: Searching and browsing digital library catalogues: A combined log analysis for the european library. In: Post-PROC. of the Fifth Italian Research Conference on Digital Libraries - IRCDL 2009. pp. 120–135 (2009)
7. Agosti, M., Di Nunzio, G.M.: Gathering and mining information from web log files. In: Digital Libraries: Research and Development, First International DELOS Conference. pp. 104–113 (2007)
8. Agosti, M., Di Nunzio, G.M.: Web log mining: A study of user sessions. In: 10th DELOS Thematic Workshop on Personalized Access, Profile Management, and Context Awareness in Digital Libraries (2007)
9. Agosti, M., Lawless, S., Marchesin, S., Wade, V.: An adaptive cross-site user modelling platform for cultural heritage websites. In: Digital Libraries and Archives. CCIS, Springer Verlag, in print (2017)
10. Agosti, M., Orio, N., Ponchia, C.: Guided tours across a collection of historical digital images. In: Proc. of the Third AIUCD Annual Conference. pp. 7:1–7:6. AIUCD 2014, ACM Press (2015)
11. Broder, A.: A Taxonomy of Web Search. SIGIR Forum 36(2), 3–10 (2002)
12. Burges, C.J.: From ranknet to lambdarank to lambdamart: An overview. Tech. rep. (2010)
13. Burnett, S., Clarke, S., Davis, M., Edwards, R., Kellett, A.: Enterprise Search and Retrieval. Unlocking the Organisation's Potential. Butler Direct Limited (2006)
14. Chapelle, O., Metzler, D., Zhang, Y., Grinspan, P.: Expected Reciprocal Rank for Graded Relevance. In: Proc. 18th International Conference on Information and Knowledge Management (CIKM 2009). pp. 621–630. ACM Press (2009)
15. Croft, W.B., Metzler, D., Strohman, T.: Search Engines: Information Retrieval in Practice. Addison-Wesley (2009)
16. Di Nunzio, G.M., Leveling, J., Mandl, T.: Logclef 2011 multilingual log file analysis: Language identification, query classification, and success of a query. In: CLEF 2011 Labs and Workshop, Notebook Papers (2011)
17. Di Nunzio, G.M., Leveling, J., Mandl, T.: Multilingual log analysis: Logclef. In: Proc. of the 33rd European Conference on IR Research, ECIR 2011. pp. 675–678 (2011)
18. Ferrante, M., Ferro, N., Maistro, M.: Injecting User Models and Time into Precision via Markov Chains. In: Proc. 37th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2014). pp. 597–606. ACM Press (2014)

19. Ferro, N., Lucchese, C., Maistro, M., Perego, R.: On Including the User Dynamic in Learning to Rank. In: Proc. 40th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM Press (2017)
20. Gauch, S., Speretta, M., Chandramouli, A., Micarelli, A.: User profiles for personalized information access. The adaptive web pp. 54–89 (2007)
21. Joachims, T., Granka, L., Pan, B., Hembrooke, H., Gay, G.: Accurately Interpreting Clickthrough Data as Implicit Feedback. In: Proc. 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2005). pp. 154–161. ACM Press (2005)
22. Joachims, T., Swaminathan, A., Schnabel, T.: Unbiased Learning-to-Rank with Biased Feedback. In: de Rijke, M., Shokouhi, M., Tomkins, A., Zhang, M. (eds.) Proc. 10th ACM International Conference on Web Searching and Data Mining (WSDM 2017). pp. 781–789. ACM Press (2017)
23. Keenoy, K., Levene, M.: Personalisation of web search. In: Intelligent techniques for web personalization, pp. 201–228. Springer (2005)
24. Koidl, K., Conlan, O., Wade, V.: Cross-site personalization: assisting users in addressing information needs that span independently hosted websites. In: Proc. of the 25th ACM conference on Hypertext and social media. pp. 66–76. ACM (2014)
25. Leveling, J., Di Nunzio, G.M., Mandl, T.: Logclef: Enabling research on multilingual log files. In: Proceedings of the First Workshop on Personalised Multilingual Hypertext Retrieval. pp. 55–56. PMHR '11, ACM Press (2011)
26. Lupu, M., Hanbury, A.: Patent Retrieval. Foundations and Trends in Information Retrieval (FnTIR) 7(1), 1–97 (2013)
27. Mandl, T., Agosti, M., Di Nunzio, G.M., Yeh, A.S., Mani, I., Doran, C., Schulz, J.M.: Logclef 2009: The CLEF 2009 multilingual logfile analysis track overview. In: 10th Workshop of the Cross-Language Evaluation Forum, CLEF 2009, Revised Selected Papers. pp. 508–517 (2009)
28. Mandl, T., Di Nunzio, G.M., Schulz, J.M.: Logclef 2010: the CLEF 2010 multilingual logfile analysis track overview. In: CLEF 2010 LABs and Workshops, Notebook Papers (2010)
29. Moffat, A., Zobel, J.: Rank-biased Precision for Measurement of Retrieval Effectiveness. ACM Transactions on Information Systems (TOIS) 27(1), 2:1–2:27 (2008)
30. Müller, H.: Medical (Visual) Information Retrieval. In: Agosti, M., Ferro, N., Forner, P., Müller, H., Santucci, G. (eds.) Information Retrieval Meets Information Visualization – PROMISE Winter School 2012, Revised Tutorial Lectures. pp. 155–166. LNCS 7757, Springer (2013)
31. Spina, D., Maistro, M., Ren, Y., Sadeghi, S., Wong, W., Baldwin, T., Cavedon, L., Moffat, A., Sanderson, M., Scholer, F., Zobel, J.: Understanding User Behavior in Job and Talent Search: An Initial Investigation. In: Proc. of the 2017 SIGIR workshop On eCommerce (eCom 2017). CEUR-WS.org (2017)
32. Sweetnam, M., Siochru, M., Agosti, M., Manfioletti, M., Orio, N., Ponchia, C.: Stereotype or spectrum: Designing for a user continuum. In: the Proceedings of the First Workshop on the Exploration, Navigation and Retrieval of Information in Cultural Heritage, ENRICH (2013)
33. Sweetnam, M.S., Agosti, M., Orio, N., Ponchia, C., Steiner, C.M., Hillemann, E.C., Siochrú, M.Ó., Lawless, S.: User needs for enhanced engagement with cultural heritage collections. In: TPDL. pp. 64–75. Springer (2012)
34. Wu, Q., Burges, C.J.C., Svore, K.M., Gao, J.: Adapting boosting for information retrieval measures. Information Retrieval 13(3), 254–270 (2010)