User Activity Characterization in a Cultural Heritage Digital Library System

Cyrille Suire, Axel Jean-Caurant, Vincent Courboulay, Jean-Christophe Burie, Pascal Estraillier

User Activity Characterization in a Cultural Heritage Digital Library System

Cyrille Suire
cyrille.suire@univ-lr.fr

Axel Jean-Caurant
ajeanc01@univ-lr.fr

Vincent Courboulay
vcourbou@univ-lr.fr

Jean-Christophe Burie
jcburie@univ-lr.fr

Pascal Estraillier
pestrail@univ-lr.fr

L3i laboratory, University of La Rochelle
La Rochelle, France

ABSTRACT

Digital access to large amount of heterogeneous data can create methodological biases regarding the discovery and exploitation of resources, particularly when it comes to Social Sciences. In order to provide relevant adaptivity for social scientists, it is important to fully consider their research practice diversity. To do so, we consider an activity-based approach for researchers' information search behavior. We have also conducted an experiment in a Cultural Heritage use case. The main result shows us that social scientists have the same research behaviors as those observed in exact Sciences.

Categories and Subject Descriptors

H.3.7 [**Information Storage and Retrieval**]: Digital libraries

Keywords

Information Seeking; User Modeling; User Behavior; Cultural Heritage; Task Models; Humanities

Paper accepted in JCDL16 Conference
DOI: http://dx.doi.org/10.1145/2910896.2925459

1. CONTEXT AND MOTIVATION

Social scientists and humanists have an increasing number of web platforms to access heterogeneous resources. The new challenge for researchers is the discovery and exploitation of these data. The tools commonly used to access this huge amount of data can create methodological problems [5]. Moreover, we know that the practices of social scientists are very diverse [3]. They need to be analyzed and understood to avoid most methodological issues. To investigate this, we present an activity-based approach. Following Marchionini's conceptual framework [4], users' research activities can be distinguished in two main groups: **lookup tasks** and **exploratory search tasks**. A recent paper by Athukorala et. al. [2] has shown that it is possible to distinguish the type of task the users are engaged in, when they are using a *scientific search engine*. The purpose of our work is to determine if this stands in another context — using a *digital library platform* and experimenting with a different type of public.

2. HYPOTHESE

As proposal, we want to prove that user behavior evaluation based on activity type is valid even if several points of our approach differ from [2]. Indeed, our system is significantly different and is built around the following considerations: its design and implementation is closer to a digital library web portal than a search engine like Google Scholar. Considering the work of [1] this kind of systems requires a specific evaluation. We implemented precise observers in our system to log users' behavior and compute features in two different ways. First, we focus on the first query iteration. A query is composed of every user events from the first input in the search field to the next one or to the end of the task. Then we focus on the overall task by averaging the values of each feature over all queries from the task.

- ***Query Length*** (f_1)***:*** this feature is defined as the number of terms in the query.
- ***Duration*** (f_2)***:*** duration in seconds.
- ***Maximum scroll value*** (f_3)***:*** This corresponds to the proportion of the page seen by the user. It reflects the number of items observable in the result list.
- ***Number of clicked items*** (f_4)***:*** how many resources were clicked on in the results page. That corresponds to an access to more precise metadata.
- ***Position of those items in the result list*** (f_5)***:*** this is a mean of the position of the items selected by the user inside the search engine results page.
- ***Number of viewed documents*** (f_6)***:*** how many original documents were actually opened by the user.
- ***Duration dwelling*** (f_7)***:*** time the user spent investigating documents outside the results page (in seconds).

3. EXPERIMENT

We conducted a first experiment with forty master students who are a the beginning of their specialization in

the Cultural Heritage field of study. They have a medium level of topic knowledge and are able to adapt their behavior to the task type. Our experimental corpus must be precise enough to allow participants to perform realistic tasks. It contains 240 documents related to the Cultural Heritage field of study which the students are relatively familiar with. The resources of the corpus are heterogeneous: general works, scientific articles, iconographic documents, multimedia resources and even primary documents.

To carry out our experiment, we chose to make the students perform tasks in each Marchionini's overall category [4]. These tasks were designed by a specialist of the Cultural Heritage field of study. We defined the tasks such as follows.

Lookup category: Tasks T_1 and T_2 referred to the *fact retrieval* subcategory. For these two first tasks we suggested two straightforward questions. Participants had to find the date of a particular event and where an event took place. The search goal was simple and participants were able to decide when they had found the correct information. During the third task *(T_3)*, students had to retrieve a precise document using given information. This task belongs to the *known item search* subcategory.

Exploratory category: We have identified one task for Marchionini's *learn and investigate* group referring in particular to the *knowledge acquisition* and *accretion* low level subcategories. The exploratory task *(T_4)* is an important and regular work for Humanities researchers and students. Participants had to identify and write a research problem on the Cultural Heritage topic in one of its aspect covered by our data set.

At the beginning of our experiment, all the participants received a fifteen minutes training on the use of our experimental platform. In order to control technical factors that could affect user behavior, all participants did the experiment under the same conditions. They used the same web browser connected to our experimental platform on a desktop computer and a 22-inch display. We gave participants the same instructions concerning the four required tasks. Users had to select the type of task inside a list and click on the start button to begin each task. At the end of the task, users had to click on the end button and provide the answers and results required by the current task.

4. RESULTS AND DISCUSSION

For the four tasks, features were computed regarding the first query iteration (f_n) and the entire task (f'_n). Because the raw data does not follow a normal distribution, we used non parametric methods. We performed a Friedman test on each feature, followed by a pairwise Wilcoxon signed rank test to evaluate the differences between two different tasks. Our results are consistent with the analysis made by [2]. Indeed, most of the features are fundamentally different when it comes to exploratory tasks. However, we find ourselves in a different context than the study led by Athukorala. Our system is a digital library storing not only scientific articles but more heterogeneous resources. Conducting the experiment with a different public — participants coming from social sciences and Humanities — leads us to the same results as Athukorala who took an interest in Computer Science researchers. Yet, these two types of researchers have very different habits in terms of information search.

Going further than the confirmation of previous results, we made the following additional observations. We can observe that some of the features presented in Table 1, computed on the first query and the overall task behave similarly. Identifying the type of task the user is engaged in as early as possible is key to the system adaptation. To discriminate user's task as soon as possible, features f_1, f_2, f_3 and f_5 are significant. When looking at the results in Table 1, it seems that T_1 behave differently than T_2 and T_3 when compared to T_4. This is particularly true for f_2, f_3, f_6 and f_7, in the case of the first query iteration. We made another statistical test comparing T_1 and T_2 that confirmed this observation. It seems that when users discover the platform for the first time, some features are clearly impacted. This could be used to evaluate how difficult it is to get familiar with the platform and should be taken into account to efficiently contextualize user behavior traces.

	first query (f)			overall task (f')		
	T4			T4		
	T1	T2	T3	T1	T2	T3
	query length					
p	***	***	***	*	***	***
Z	-4.28	-3.33	-4.16	-2.56	-4.05	-4.08
	duration					
p	0.1	***	***	***	***	***
Z	1.61	-3.75	-3.78	-4.33	-4.55	-4.53
	max scroll value					
p	0.06	**	**	**	0.18	*
Z	-1.87	-2.97	-3.74	-2.64	-1.34	-2.09
	clicked items position					
p	*	*	0.05	***	***	***
Z	-2.25	-2.31	-1.96	-3.39	-3.92	-3.51

Table 1: Statistical differences between an Exploratory task (T4) and three different Look-up tasks (T1, T2 and T3). The left-hand side of this table describe the results during the first query iteration, while the right-hand side concerns the overall task. We use a Wilcoxon signed rank test. *p*-values with *s are statistically significant, with * for $p < 0.05$, ** for $p < 0.01$ and * for $p < 0.001$.**

5. REFERENCES

[1] M. Agosti, F. Crivellari, G. M. D. Nunzio, and S. Gabrielli. Understanding user requirements and preferences for a digital library Web portal. *Int J Digit Libr*, 11(4):225–238, Sept. 2011.

[2] K. Athukorala, D. Glowacka, G. Jacucci, A. Oulasvirta, and J. Vreeken. Is exploratory search different? A comparison of information search behavior for exploratory and lookup tasks. *J. Assoc. Inf. Sci. Technol.*, 2015.

[3] N. Audenaert and R. Furuta. What Humanists Want: How Scholars Use Source Materials. In *Proc. of the 10th JCDL*, pages 283–292, NY, USA, 2010. ACM.

[4] G. Marchionini. Exploratory Search: From Finding to Understanding. *Commun. ACM*, 49(4):41–46, 2006.

[5] I. Milligan. Illusionary Order: Online Databases, Optical Character Recognition, and Canadian History, 1997-2010. *The Canadian Historical Review*, 94(4):540–569, 2013.