Search or browse? Casual information access to a cultural heritage collection

Robert Villa, Paul Clough, Mark Hall, Sophie Rutter

Information School, University of Sheffield, Sheffield, UK

{r.villa, p.d.clough, m.mhall, sarutter1} @sheffield.ac.uk

Introduction

Providing public access to cultural heritage is an ongoing and challenging area of research. Previous work suggests that visitors to online cultural heritage collections are not necessarily motivated by an explicit task, and that interacting with cultural heritage collections is exploratory in nature. In this paper, we focus on how individuals explore a cultural heritage collection when given no task.

We consider three research questions:

RQ1: How do participants initiate their exploration?

RQ2: Do participants use browse or search in their exploration of the collection?

RQ3: How do participants decide to search or browse, when given no explicit task?

Interactive CHiC and Experimental Setup

This work is based on initial results from the Interactive CHiC (Cultural Heritage in CLEF) track of CLEF, based on the CHiC Europeana data set. No explicit task was provided to users. Instead instructions asked the user to explore freely as they wished, using the "Cultural Hertiage Interface" (right). Users were informed after they had been active for 10 minutes, and could then continue for a further 5 minutes if they wished, at which point they would be asked to stop.

In total 20 participants were recruited for the study, 11 male and 9 female. Eight participants were in the 18-25 year age band, nine in the 26-35 band; the other 3 between 36-45.

How did you start?

Over the whole data set four different actions were used by participants to initiate their session (Table 1, column 2). For the majority of users, the first action was to select one of the categories (15 out of the 20 users).

Action	#Users first action	#Users first search/browse action
Category select	15	20
Display item	3	-
Next search result page	1	-
Add to bookbag	1	-

Table 1: Number of users whose first action/first search or browse action were as column one.

Also investigated: how long each user spent before either clicking the interface, or starting a new search or browse (Table 2).

	Min	1st Qu.	Median	Mean	3rd Qu	Max
First action	7.00	19.00	25.00	30.50	38.75	90.00
First search/ browse	7.00	22.75	38.00	57.50	81.75	204.0
Total time	129	631.8	783.5	787.8	918.0	1544

Table 2: Time to first action, time to first search/browse action, and overall session time (all times in seconds)

Initiation of exploration

In a post-session interview two questions were asked of users: "how did you start?" and "Why did you choose to start with a [category/search query]?" The responses to the first question mentioned the category browser explicitly in 8 of the 12 answers. In most of these cases this was linked to exploring the interface. E.g.: *"I was drawn to the middle then decided to look around at the interface. I decided to look at categories first, picked politics"*

Responses from some users suggest that prior interests also played a part, e.g: *"I just look at the layout of the website and then found that I had a category browser so I went to what I study actually, and I study languages and I try to find something interesting."*

The design of the interface, with a relatively small search box, appears to also have had an effect, e.g. *"Because I only saw that [category]. I didn't see the search until a bit later on."* and *"I didn't really see this one at first [the search box] it was a bit obscure."*

For many, however, the fact that the category browser allowed easy exploration appeared to be the key, with some users making connections to physical museums.
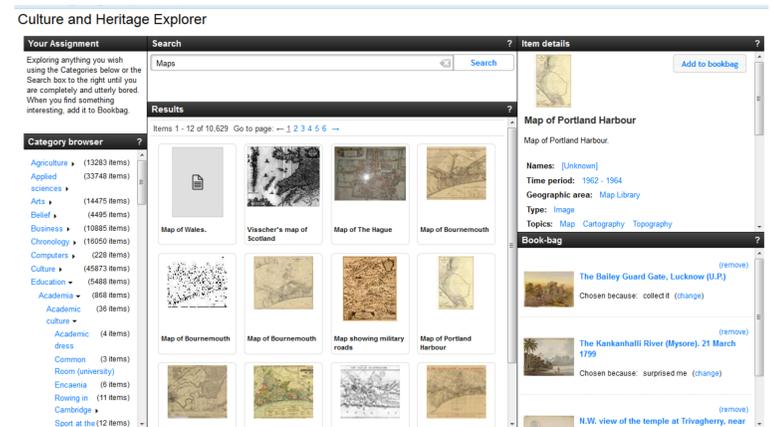
Search vs. browse

Figure 2 presents query and category counts across all users. A non-parametric Wilcoxon rank-sum test indicated that there was a significant difference between queries executed and categories selected (W = 50.5, p ≤ 0.001)

Time querying vs. browsing categories was estimated (Figure 3) by starting a timer when a query or category was selected, and taking all activity between this point and the next query or category select as the user either "querying" or "browsing categories".
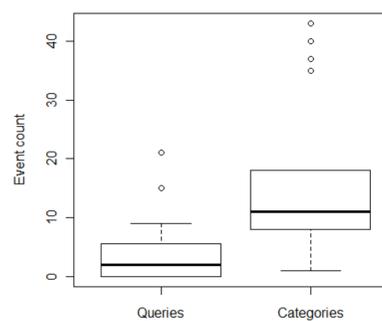

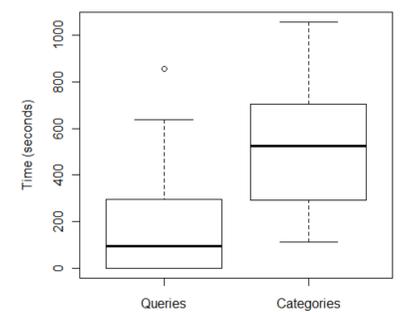**Figure 2: Comparison of query and category select counts**

Figure 3: Estimated time querying vs. browsing by category

Discussion and conclusions

For RQ1: from Table 1 all 20 participants started their exploration using the category browser, rather than a text search. The first action for the majority of users (75%) was to select a category. Quantitative data backs this up, with text transcripts frequently explicitly mentioning the category browser as a way of starting their exploration.

For RQ2 and RQ3, from Figure 2 and Figure 3 it can be seen that there is a general preference for browsing, e.g. from Figure 3 the median estimated time spent browsing using the categories was 524 seconds (IQR 399), compared to 77 seconds (IQR 394) for text queries. Looking at the participant comments, the lack of any explicit task would appear to have played a part in this preference.

The preliminary results reported here would suggest that providing browse functionality to cultural heritage collections is important for users arriving without a specific information need, as may be typical in casual search. For the majority of users, this preference for category browsing continues to hold for the session as a whole, with all but 5 users spending more time browsing than keyword searching.