Jamshid Beheshti, Andrew Large, Marni Tam
School of Information Studies, McGill University, Montréal, Quebec

Search Patterns on a Children's Portal

Abstract: More than two years of transaction logs on a children's portal are analyzed to investigate search patterns, when the users are presented with four search options. The results show that the hierarchical subject directory and alphabetic search options accounted for 83 percent of all the searches, indicating users' preference for browsing rather than keyword searching.

Résumé : Cette communication porte sur l'analyse de deux ans de journaux de transactions d'un portail pour enfants afin de dégager les méthodes de recherche des utilisateurs lorsqu'ils font face à quatre options de recherche. Les résultats démontrent que le répertoire thématique hiérarchique et les options de recherche alphabétiques représentent 83 pour cent de toutes les recherches, indiquant une préférence pour le furetage plutôt que la recherche par mots clés.

1. Introduction

In this paper, we report on the information searching patterns on a children's portal as recorded in the transaction logs. The objective of the research is to investigate the search patterns on a portal when users are presented with four search options: keyword, hierarchical subject directory, alphabetical, and advanced. The portal, *History Trek* (http://www.historytrek.ca), was created by an intergenerational team comprising of adult designers and young students who collaborated over thirteen one-hour sessions. The students were an integral and essential part of the team as they provided the perspective of the target audience – children (Large et al, 2007).

History Trek provides access to more than 2,300 web pages in English, French or both that are appropriate in content and language for elementary students and relate to some aspect of Canadian history. The portal includes the following retrieval mechanisms: keyword searching, topic taxonomy (a hierarchical subject directory), an alphabetical word search, an advanced search, and a scrollable timeline. The portal also includes help screens, and access to some web-based Canadian history quizzes (Figure 1). Each record includes a number of hyperlinked subject index terms assigned by professional librarians.

2. Methodology

Web traffic data have been stored since *History Trek* became publicly available in September 2007. From 792,173 hits between 2007 and June 2009, 92,226 transactions containing data on information searching were extracted. Selection of the transactions was based on unique IP addresses, indicating unique access points. While each IP may not represent one user (for example in a school the IP may be shared by many users), we can assume that the transactions demonstrate general patterns and trends of information searching patterns of young users on *History Trek*.

Figure 1. Four main search options

3. Results and Discussion

The results of the data analysis from transaction logs show that Topic taxonomy is the most popular search options at 42 percent of transactions (Table 1). Alphabetic searching is the second most popular search option at 41 percent, followed by the Google-style keyword search option at only 15 percent. The Advanced Search is the least popular among the four search options at only 3 percent. While we do not know the exact motivation of users for choosing one (or more) options over others, we can speculate that they choose the Alphabetic and the Topic Search to explore and browse for relevant information.

Table 1. Search Options

Options	Frequency	Percent
Topic	38633	42%
Alphabetic	37416	41%
Keyword	13727	15%
Advanced	2450	3%

a. Topic taxonomy

The Topic search comprises a taxonomy of around 1200 terms relating to Canadian history, created by the original *History Trek* research team, with eight main topics and up to four levels of depth within these topics. Table 2 shows the usage frequency of the second, third, and the fourth level of the taxonomy. The topic levels are not created equally; for some of the topics the taxonomy does not go beyond level 3, which may explain the much smaller usage frequency for level 4, at only 7.6 percent. Although browsing the Topic taxonomy may appear to be child-friendly it poses its own problems to effective information retrieval.

Table 2. Topic taxonomy

Taxonomy	Use Frequency
Level 2	38.8%
Level 3	53.5%
Level 4	7.6%

Ideally, the Topic taxonomy should allow children to recognize concepts that are relevant to their information needs, potentially involving less cognitive effort than retrieving from memory terms to be used for a keyword search. Taxonomies in children's web portals such as Yahooligans!http://www.squirrelnet.com/search/yk/yahooligans.asp), KidsClick (http://www.kidsclick.org/) and History Trek may, however, impose heavier cognitive loads than keyword searching. The child must recognize a suitable entry point into the taxonomy, which should be straightforward so long as the terms displayed in the taxonomy are relevant to his/her information needs. It would be challenging for a child if a relevant term is not displayed at the first level of the taxonomy; the greater the depth of the taxonomy, the greater the potential for failure (Large et al, In Press). In this study, the results demonstrate a willingness on the part of the users to explore the Topic taxonomy. Its usage is significantly higher than keyword searching and level 2 and level 3 of the taxonomy are visited frequently.

b. Alphabetical searching
In the Alphabetical search the first letter of a word is chosen, displaying a list of all subject descriptors beginning with that letter.
Children may be competent computer users, but they are not able always to retrieve information efficiently and effectively. Many of the problems typically encountered by children in information retrieval may also confront adults, including determining the appropriate concepts, translating these concepts to keywords that the information retrieval (IR) system can understand, finding appropriate synonyms, and distinguishing homonyms (Large, Nesset, Beheshti, 2008). Spelling is particularly a major obstacle for many young users (Bilal, 2000). The Alphabetical search option allows children to scan through letters of alphabet as if they are flipping through pages of a book. Shenton and Dixon (2003) found that when children were searching for precise information in printed sources, they often flipped through the pages of the entire book attempting to access sequentially the required information.
Another explanation for choosing the Alphabetic searching over the topic taxonomy may be the challenges facing the designers of the taxonomy's categories, which should try to reflect a child's categorization of any particular subject (Bilal, 2005).

c. Keyword searching
History Trek Keyword searching incorporates spell checking, right-hand stem truncation (using Porter's Stemming Algorithm) and some synonym control (for example, the keyword "automobile" will also find "cars"). While Google is the most widely used search engine by adults, and younger users, it is somewhat surprising that the Keyword

search option was utilized relatively infrequently. Druin (2005), however, found similar results for the children's digital library project – visitors used the graphical tools for searching about 90 percent of time versus using text keyword searching at approximately 10 percent of time.

d. Advanced searching
The Advanced search consists of three restrictive options: Title words; Subject descriptors (assigned to each record by professional librarians); and Words in a phrase. Considering that Advanced searching is based on keyword searching, the result of the transaction analysis showing very little use of Advanced searching is not surprising.

e. Other findings
A very important feature of the History Trek is the hyperlinked subject terms assigned to each record. More than 90 percent of users utilized this feature to retrieve information. Naturally, *History Trek*, covering the Canadian History is popular in Canada with 34.1 percent of transaction logs for searching generated in Canada. It is surprising, however, that the US accounts for 55.9 percent of transactions, while the remaining 10 percent are distributed among 83 countries from around the globe, with the bulk of the hits generating from China (2.7%).

4. Conclusion
History Trek is perhaps the only example of the children's web portal that is designed by an intergenerational team actually including children and fully functioning on the Web. This paper presents an analysis of the search patterns based on the transaction logs from a very large number of users from across the world. We are not aware of any other study on children's portal that utilizes more than 790,000 transaction logs to investigate search patterns. The results show that users (presumably in many cases young users) tend to prefer browsing through alphabetical lists as well as through different levels of a topic taxonomy than keyword searching. The overwhelming majority of information retrieval systems such as Google are designed to accommodate keyword searching, which requires the users to recall from memory appropriate keywords, placing a heavier cognitive load upon the young user than browsing. Young people, in principle seem to opt for browsing just because it is less demanding cognitively since it imposes an information structure that immediately restricts choice to a limited number of options (Borgman et al, 1995).

5. References

Borgman, C., Hirsh, S., Walter, V., and Gallagher, A. L. 1995. Children's searching behavior on browsing and keyword online catalogs: The Science Library Catalog project. Journal of the American Society for Information Science, 46: 663-684.

Druin, A. 2005. What children can teach us: developing digital libraries for children with children. Library Quarterly, 75: 20-41.

Large, A., Beheshti, J., Nesset, V., and Bowler, L. 2007. "Children's Web Portals: Can an Intergenerational Design Team Deliver the Goods?" In Youth Information-Seeking Behavior, Volume 2, ed. Mary K. Chelton & Colleen Cool (Lantham: Scarecrow Press), 279-311.

Large, A., Nesset, V., Beheshti, J. 2008. Children as information seekers: what researchers tell us. New Review of Children's Literature and Librarianship, 14: 121-140.

Large, A., Nesset, V., Tabatabaei, N., Beheshti, J. (In press). Bonded design revisited: involving children in information visualization design. Canadian Journal of Library and Information Science.