To be presented at the 8th Int'l Conf. on Human-Computer Interaction – HCI'99, 25.8.99

Information Gathering from Hypermedia Databases

Joan C. Nordbotten
Dept of Information Science, University of Bergen,
N-5020 Bergen, Norway
hhtp://www.ifi.uib.no/staff/joan

HYPERMEDIA DATABASES

The Internet provides an opportunity to reach the general public with information traditionally disseminated through libraries, and museums. One consequence is that these institutions have begun to make large quantities of information available as hypertext/hypermedia databases of inter-linked images and/or documents.

The hypermedia format is particularly useful for presentation of museum data where information about a topic or artifact is normally presented using multiple media, such as images, text, tables, charts, film, sound tracks, and/or video. The French Ministry of Culture has perhaps the most ambitious museum project. It will contain more than 22.5 million documents (Mannoni 1996, 1997). J. Bowen maintains an updated index to museums with web pages (Bowen 1997). Off-line exhibits can be found in museums and other public buildings.

Hypermedia technology supports associative information retrieval and can facilitate information gathering (Bush 1945, Nelson 1967, Shneiderman 1992). However, researchers anticipate a number of problems with hypermedia system usage. As the number of inter-linked documents and path selections increases, *user disorientation* and *cognitive overload* may hinder users in gathering information (Conklin 1987, Preece 1994). Link structures may actually hinder location of specific information (MacKenzie 1996). And, it is uncertain if museum databases reach their intended public or whether their user's information requirements are satisfied (Day 1995, Futers 1997).

For the general public, gathering information entails location and retrieval of *interesting* document sets (Futers 1997). One measure of interest in a document set can be the time spent viewing/reading it. Providing effective support for

information gathering from hypermedia databases requires an understanding of how the intended public finds and accesses information, including:

- How much time the information gatherer is willing to invest.
- How many page selections the information gatherer is willing to use.
- User preference for serial or associative search strategies.
- How page layout affects the user's search for information.

As can be quickly observed, there is little standardization in the structure and presentation formats for hypermedia databases, indicating a lack of consensus as to what constitutes good information presentation.

INFORMATION GATHERING from a MUSEUM EXHIBIT

Museum visitors come from the general public. They can be considered similar to web users in the sense that both groups contain casual browsers, looking for something interesting, as well as goal oriented seekers of specific information. Thus studying usage patterns for off-line hypermedia exhibits, should provide information that can also be used to design effective web databases.

We have studied usage patterns for a small hypermedia exhibit presenting social sciences projects, developed as part of the 50th anniversary celebration for the University of Bergen (Nordbotten & Nordbotten 1999). The exhibit was implemented on an off-line PC with touch screen input using a WebSite™ server with a Netscape™ browser. The exhibit content, shown in Figure 1, was given in an overview page accessible from each exhibit page.

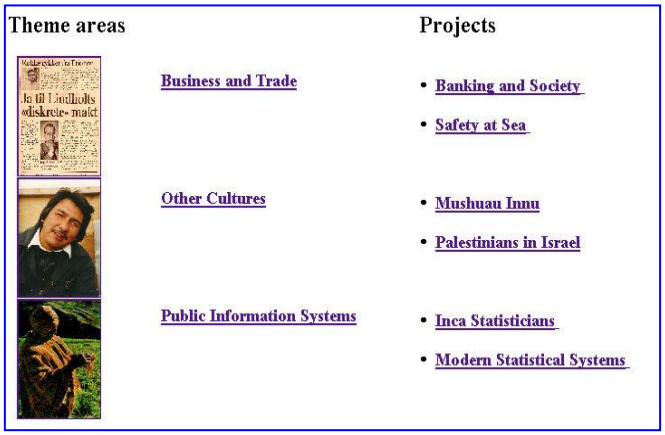

Figure 1. Exhibit content

The exhibit was first located as part of a traditional museum presentation of university research projects, made available to the general public by the Bergen Museum, in Bergen, Norway. It was later moved to the information area of the School of Social Science, at the University of Bergen. An English version is currently available at http://129.177.34.238/eng_museum/00sv-ut.htm.

Usage data has been collected using the server log. A summary, for 4 week periods for each of 3 populations, is given in Table 1. In these samples, 22-50% of the sessions contained only index pages, indicating that the visitor found nothing of interest in the exhibit. The last 3 columns of Table 1 give usage characteristics for those sessions containing at least 1 project selection.

Table 1: Session profile

Data Period	**Dominant visitor group in sample**	**Exhibit starts**	**project sessions #**	**projects/ session Avg.**	**Session length. Page #**	**session time Sec.**
Fall '96	School classes, University students	331	225	1.1	5.75	44.9
Summer '97	Adult visitor/tourist	374	187	1.6	7.0	54.6
Fall '98	Social science Student	322	250	1.4	8.2	78.8

As expected, the exhibit appears to be most interesting for the social science students, when interest is measured by length of time spent and number of pages viewed. The slight drop in projects selected can indicate that these visitors are more representative of goal searchers than general information gatherers.

We had assumed that the exhibit projects would be equally interesting to a varied, general public. As shown in Table 2, this was not the case. For each of the sample populations, project choice is highly correlated, >0.8, to its place in the index set. Note that, after the analysis for the 1st population, the theme index order shown in Table 2, was changed to the order shown in Fig.1, to test this observation.

Table 2: Project Selection

Index level 1 Theme area	***Index level 2 Projects within theme***	***Pop.'96 %***	***Pop.'97 %***	***Pop.'98 %***
Other cultures	Mushuau Innu	35	28	22
	Palestinians	14	12	14
Business and Trade	Banking & Society	22	23	24
	Maritime safety	8	8	8
Public information systems	Inca statisticians	18	20	18
	Modern IT systems	4	10	13

Project presentation length varies from 2 to 8 pages. Each initial page contains text links to its detail pages, which give different information about the project. Fig.2a shows that detail page selection falls from nearly 80% to 30% from selection of the 1st to 4 detail pages. Figure 2b shows a marked difference in the use of the embedded links for page selection. Here, the social science students were nearly twice as likely to choose an associative link, than the other museum visitors who selected exhibit pages using the directional buttons on each page. Possible explanations include:

- familiarity with hypermedia presentations.
- the longer time spent made it more likely that the links were seen.
- familiarity with the subject matter allowed preference selection.

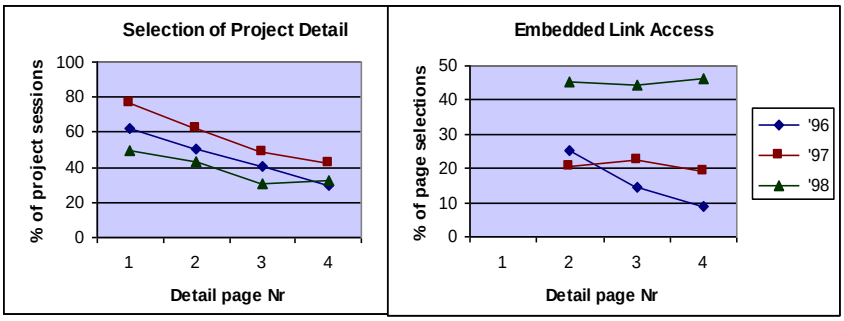

Figure 2. Detail page selection: a) by number b) use of embedded link

SUMMARY and CONCLUSIONS

We have studied viewing sessions for an exhibit of social science projects, for 3 populations, dominated by school and university students, adults, and social science students, respectively. Though social science students spent the most time and viewed most pages, they selected both fewer projects and detail pages than adult museum visitors, indicating index set review and interest selection.

Our visitors spent, on average, about 1 minute viewing 7 pages on less than 2 topics. Session length in time and page number varied from 4 to 103 seconds and 3 to 36 pages, increasing with the age and topic 'nearness' of the viewer. Index placement dominated topic selection for all groups, while use of embedded links appears related to familiarity with hypermedia presentations.

In conclusion, we would advise that hypermedia exhibits be organized as short theme presentations and that indexes be ordered by theme importance.

ACKNOWLEDGEMENTS

This project has been part of the celebration activities for the 50th anniversary of the University of Bergen and the 25th anniversary of the School of Social Sciences. Thanks are extended to staff, faculty, and students of the Bergen Museum, School of Social Science, and the Department of Information Science for their help in the construction and test of the electronic exhibits. Special thanks are extended to Professor Svein Nordbotten for all project support.

REFERENCES

Bowen,J. (1997). Virtual Library museums pages (VLmp). http://www.icom.org/vlmp/.

Bush, V. (1945). As we may think. *Atlantic Monthly* July 176:1, 101-108.

Conklin,J. (1987). Hypertext: An introduction and survey. *IEEE Computer*, 20:1,17-41.

Day,G., (ed.) (1995). Discussion. *Proceedings. Museum Collections and the Information Superhighway.* Science Museum, London. http://www.nmsi.ac.uk/infosh/discuss.htm.

Futers, K. (1997). Tell Me What You Want, What You Really, Really Want: a look at Internet user needs. *Mda*. http://www.open.gov.uk/mdocassn/eva_kf.htm.

MacKenzie, D. (1996). Beyond Hypertext: Adaptive Interfaces for Virtual Museums. http://www.dmcsoft.com/tamhpapers//evaf.htm.

Mannoni,B. (1996). Bringing Museums Online. *Communications of the ACM*, 39(6), 100-105.

Mannoni,B. (1997). A Virtual Museum. *Communications of the ACM*, 40(9), 61-62. See also http://www.culture.fr/lumiere/documents/files/imaginary_exhibition.html

Nelson, T.H. (1967). Getting it out of our system. *Information Retrieval: A Critical Review*. G.Schechter, ed. Thompson Books, 191-210.

Nordbotten,J. & Nordbotten,S. (1999) Search Patterns in Hypertext Exhibits. *Proc. 32nd Hawaii Int'l Conference on System Sciences - HICSS'32*. Maui, HI. Jan.4-9, 1999. IEEE. 1999. ISBN 0-7695-0001-3. 7p.

Preece, J. et.al. (1994). *Human-Computer Interaction*. Addison Wesley.

Shneiderman, B. (1992). *Designing the User Interface - Strategies for effective Human-Computer Interaction*, 2nd ed. Addison-Wesley.