

# Search Patterns in Hypertext Exhibits

Joan C. Nordbotten

*Dept of Information Science,*

*University of Bergen, N-5020 Bergen, Norway*

e-mail: [joan.nordbotten@ifi.uib.no](mailto:joan.nordbotten@ifi.uib.no)

ULR: <http://www.ifi.uib.no/staff/joan>

Svein Nordbotten

*Dept of Information Science,*

*University of Bergen, N-5020 Bergen, Norway*

e-mail: [svein.nordbotten@ifi.uib.no](mailto:svein.nordbotten@ifi.uib.no)

ULR: <http://www.ifi.uib.no/staff/svein>

## Abstract

*As access to hypermedia documents becomes generally available, it becomes increasingly important to understand how casual users search for information. We have studied search patterns used in a small hypermedia museum exhibit. Our goal has been to identify preferred search strategies that can be supported for effective information retrieval from web-based information systems. Our data indicates that users make little use of the embedded links enabling associative selection. Instead, topic selection was predominantly serially through first-in-index and next-page sequences. Further, sessions were short in both time and number of documents viewed, indicating little general interest in 'exploring' the exhibit.*

**Keywords** hypermedia, browsing patterns, electronic museum exhibits

## Hypermedia databases

Printed documents, containing administrative records, texts, tables, charts, images, music scores, and/or maps have long been used for information dissemination by libraries, museums, news media, and public and private organizations. The rapidly growing reach of the Internet has enabled the use of additional document forms including audio, video, and electronic databases. Web databases, composed of hypermedia documents (containing static and/or dynamic objects), are increasingly being used for information dissemination. Hypertext technology, which supports personalized, associative information retrieval, is expected to encourage information exploration and make knowledge acquisition enjoyable [Bush 1945, Nelson 1967, Shneiderman1992].

Two fundamental problems with using hypertext systems, *user disorientation* and *cognitive overload*, may

hinder users in acquiring desired information [Conklin1987, Preece 1994]. User disorientation, a feeling of being "lost in space", is expected to be aggravated as the quantity of inter-linked information increases. Cognitive overload is expected to lead to frustration as multiple information path choices must be made while retrieving information. Indeed, the embedded link structure of hypertext may actually hinder location of specific information [MacKenzie1996]. Further, it is uncertain if Web databases reach their intended public or whether their user's information requirements have been satisfied [Day1995].

Museum visitors can be compared 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 the search strategies used by museum visitors of off-line hypertext exhibits, should provide information that can be used to limit problems and support effective information retrieval from web databases.

## Usage characteristics of hypermedia exhibits

Museums worldwide are implementing hypertext exhibits. Perhaps the most ambitious of which is that of the Ministry of Culture in France which is to contain more than 22.5 million documents [Mannoni1996, 1997]. Smaller exhibits can be found within museum buildings and as web exhibits. See [Bowen,J. 1997] for an updated index to web-based museum exhibits.

Two studies of usage characteristics for museum based (non-web) hypertext systems are reported in [Shneiderman1989]. The 1<sup>st</sup> consisted of 30, 1-8 page articles giving supplementary information about the photographic exhibit within which the system was placed. This system was placed in 2 museums. An analysis of the usage data collected for 373 and 361 sessions, showed that an average of 3.2 and 5.7 articles were accessed and 41 and 64 seconds were spent per article, respectively.

Most sessions, >80%, started with the introductory article, followed by an article of individual interest selected from the alphabetical index. Session length averaged 6 minutes.

The 2<sup>nd</sup> system reported in [Shneiderman1989] was placed at the end of an exhibit on urban development based on archaeological findings. It consisted of 200 articles describing different archaeological sites that welcomed visitors. An analysis of 3561 sessions showed that an average of 8.5 articles were accessed for an average of 16.4 seconds each. Session length averaged 2.6 minutes.

Yamada studied a hypermedia system consisting of 55 linked video scenes supplemented with a hierarchic index [Yamada1995]. The system offered more detailed information about the exhibit in which it was placed. Sessions averaged 11 and 18 scenes for 2 versions of the system. Average session time was estimated at 5 minutes allowing less than 17 seconds per scene.

Two of the systems studied, the 1<sup>st</sup> Shneiderman and Yamada, provided more detail about the surrounding traditional museum exhibit. These appear to hold user interest longer than one that adds new topics. The 1<sup>st</sup> Shneiderman exhibit consisted of text articles, which may account for the longer time spent per topic/article.

### **A study of a 'context-free' hypermedia museum exhibit**

The exhibit we have studied was developed for presentation of social science projects and placed in a traditional museum exhibit of research projects from the faculties of the University of Bergen. The topics of the hypermedia exhibit were independent of the topics of the other exhibits in the room and the museum in general (which is a traditional natural science museum). As such, the users can be compared to users of web based exhibits, which are by definition independent of their physical environment.

The exhibit was implemented as an off-line system with touch screen input. It was available to the general public at the Bergen Museum, University of Bergen for 1 year from the summer of 1996. A web version, designed for usage testing, is scheduled for release in 1998. Usage data has been collected for two museum populations. These were taken during the fall of 1996 and summer of 1997, when the primary museum visitors were students from the university and local schools and adult tourists, respectively.

The objective of our study and this paper is to identify the search strategies used by casual users of independent exhibits in a general museum environment. We assume that these users will share information search characteristics with web browsers. Our aim has been to

provide improved support for effective information retrieval from web exhibits.

### **Exhibit design**

The original exhibit consists of 33 inter-linked pages: a cover page with usage instructions, a two level index, an overview index, and 6 project presentations of 2 to 8 pages. The exhibit was implemented using the Netscape™ browser and the WebSite Professional™ server. Cover page instructions for activating the exhibit were also posted on the lectern. Figure 1 shows the cover page of the web version of the exhibit, located at: <http://129.177.34.238/museum/00sv-ut.htm>.

**Figure 1. Exhibit cover page**

**The social sciences -**

### **Studies of people and societies**



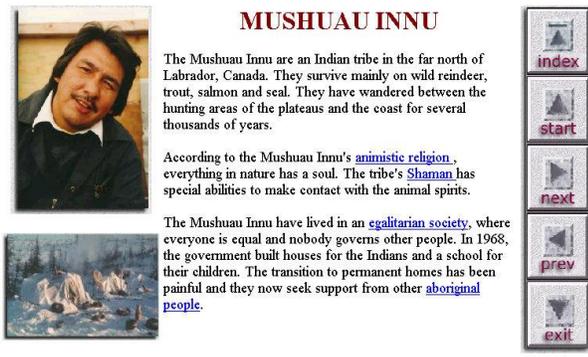
**Click the image to start the exhibit.**

**Thereafter, click the buttons or underlined text to continue.**

In the development of the exhibit, it was assumed that museum visitors would not have a priori knowledge of the exhibit contents and thus would be curious browsers. It was expected that these users would follow personal interest in topic selection and use an associative progression through the exhibit.

Primary design goals were to provide an interesting, holistic exhibit with multiple navigation tools. Since the six projects are independent and thematically quite different, a uniform presentation format, using a book layout metaphor, was chosen. The exhibit starts with a table-of-contents style index. The standard page layout defines title and text fonts, and text, image, and navigation bar placement. Each page is thematically complete (no scrolling). Each project introduction page contains embedded text links for associative page/topic selection to its detail pages as well as the navigation bar.

An example is shown in Figures 2a and 2b, which give the introductory and a project detail page from the Canadian, “Mushuau Innu” Indian project presentation.



Source and photographer: Georg Henriksen, Department of Social Anthropology, University of Bergen.

Figure 2a. Project introductory page

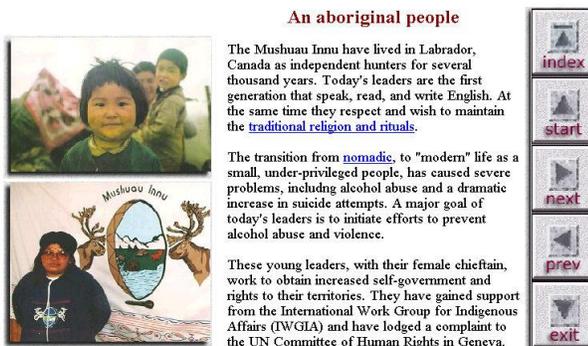


Figure 2b. Project detail page

Navigation through the exhibit is supported by embedded text links, active image links, and the navigation bar, which is included on each project page. The *index* button selects the overview page, shown in Figure 3, for the exhibit.



Figure 3. Exhibit overview index

A hierarchical exhibit structure was chosen to give equal access distance to the project presentations. In this structure, the path to any project requires 3 page selections, the 1<sup>st</sup> to activate the exhibit theme index, the 2<sup>nd</sup> to access the project list for the selected theme, and 3<sup>rd</sup> to select a project. It was assumed that this structure would support selection of interesting topics.

Finally, an automatic return to the cover page is made whenever the exhibit has been idle for more than 45 seconds, assuring that most users will start at the beginning of the exhibit.

### Usage indicators

Two aspects are expected to effect hypertext exhibit usage, user interest in the subject matter presented and information accessibility. A measure of user interest can be the time spent within the exhibit and on individual pages, as well as the number of pages viewed. Navigation patterns, defined as the sequence of pages selected by a viewer, can indicate both topic interest and accessibility. We expect substantial page access differences due to user interest and/or page accessibility.

One measure of information accessibility is the complexity and length of the access path. High use of orientation tools, such as backtracking, restart, and the overview page can be an indicator of exhibit complexity and/or user disorientation. Higher use of navigational buttons, with low use of embedded links can indicate user dependence on serial, predefined paths.

In our study of user activity, we have defined the following indicators for:

*User interest*

- The number of pages viewed by each visitor.
- The time was spent by each visitor.

*Topic interest*

- The selection frequency of the project/topic.
- Selection and time distribution for the project detail pages.

*Exhibit navigation*

- Path length and frequency through the exhibition.
- Selection frequency and time usage for the index pages.
- Selection frequency of associative links.

**Data collection**

Usage data, including time and source page identifier for each page change, was collected from the WebSite™ server log for 2 periods, as described in Table 1. The exhibit was installed at the natural science museum during the 1<sup>st</sup> and 2<sup>nd</sup> periods. The dominant museum visitors change through the year, as documented by the type of entrance fee they pay.

**Table 1. User profile**

<b>Data collection Period</b>	<b># sessions</b>	<b>Dominant user type</b>
Oct-Nov '96	450	School & university student
July-Aug. '97	374	Adult visitor/tourist

A user session is defined by a page transition from the cover page to the main exhibit index. The dominant user is assumed to be equivalent to the museum's general visitor profile, since no individual user characteristics have been available.

**Observations**

*User interest*

The exhibit was activated by about 20% of the museum's visitors, on average 16 per day, varying from 2 to 99. Given that the exhibit was only available for 4 to 5 hours a day and that it was not centrally placed in either the museum nor the university exhibit, we are satisfied with this interest level.

Session time was calculated from the first initiation of the theme index to the return to the cover page. The last page time was ignored if it equaled the system reset time under the assumption that the visitor had left the exhibit. Thus, total session time may be somewhat understated. Average session time was 52 seconds, varying from 3 seconds to 5.8 minutes. Session length averaged 4.6 and 6.5 pages for the two observation periods, respectively. Session length variation was from 3 to 50 pages.

Path analysis shows that about 50% of the sessions contained only index pages. Some were 'false starts', where the exhibit was activated and then immediately exited. We have assumed that the very short sessions can be attributed to any one of the following:

- testing that the system was functioning,
- small children who could not read,
- non-Norwegian tourists (the text was in Norwegian), or
- lack of interest in the projects presented.

*Topic interest*

There were 255 and 296 sessions that accessed at least one project in the two observation periods. These sessions averaged 1.2 and 1.6 project selections with an average of 2.5 pages displayed per project. The tendency to continue in the 2<sup>nd</sup> period may reflect the age difference in the two populations.

Average page display varied from 1.8 to 3.1 pages for the different projects. There was no significant difference in project and page selections between the observation periods. Table 2 gives the relative selection frequency for each project in the periods studied. The theme areas and projects are listed as they were presented to the user in the indexes for the initial period.

**Table 2. Project selection**

<i>Theme area</i>	<i>Project</i>	<i>P#1</i>	<i>P#2</i>
		%	%
Other cultures	Mushuau Innu	35	28
	Palestinians	14	12
Business and Trade	Banking networks	22	23
	Maritime security	8	8
Public information systems	Inca statisticians	18	20
	Modern IT systems	4	10

The frequency pattern for both theme area and project selection, in the 1<sup>st</sup> observation period, indicates a serial selection of the 1<sup>st</sup> index option at both index levels. In order to test this observation, the sequence of the theme list in the main index was changed so that the most popular theme "other cultures" was placed as the 2<sup>nd</sup> theme. (Figure 3, shows the 'new' index sequence.) In addition, the title of the last project was changed from "statistical information systems" to "modern IT systems". As shown in Table 2, the sequence change appears to have reduced the percentage of "other cultures" selections, though it remained the most popular theme. First index list choice was still the dominant choice for project selection. The title change for the 6<sup>th</sup> project seems to have increased its popularity.

**Exhibit navigation**

A session, which includes project selections, has 3 navigation path components: project selection using the indexes, 1-8 project page selections using the *next* button or an embedded link, and an exit from the project presentation and/or exhibit. More than 85% of the sessions exited from an index or project introductory page, indicating a high tendency to 'complete' project presentations.

Table 3, sorted by project selection frequency, shows the number of times each project was selected and the percentage of detail page access for each project. Note that:

- Introductory project pages were frequently reselected, either for more detail or perhaps on a "return-to-start" route. The initial selection count is shown in Table 3, column 2.
- Relative access to detail pages is given, in columns 3 to 6, as the percent of initial project selections.
- An empty cell indicates that there is no detail page at this level for the project.

In addition to the navigation buttons, three types of embedded links were included in the project pages: underlined keywords within sentences, list elements, or framed charts/pictures (see Figure 2a and b).

In average, less than 25% of page transitions within project presentations used embedded links. Both the Inca and Maritime project presentations contained image links, which seem to have invited a circular/repeat page selection.

The falling frequency of selection of detail pages, in average 67%, 53%, 37%, and 33%, for the 1<sup>st</sup> to 4<sup>th</sup> pages respectively, indicates a correlation between information accessibility and path length.

As shown in Tables 2 and 3, the most frequently accessed project pages were in the 3 primary index paths:

- Cultures 1, Innu
- Business 1, Bank
- Information systems 1, Incas

**Table 3. Project page accesses**

<i>ObservationPeriod</i> <i>1</i>	<i>Project</i> <i>selection</i>	<i>Detail page</i> Selection frequency %			
	<i>0 (ct)</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
<i>Project</i>					
Mushuau Innu	88	66	49	44	33
Banking networks	55	65	45	38	25
Inca statisticians	47	49	62		
Palestinians	35	46	49	23	
Maritime security	20	105			
Modern IT systems	10	40	50	20	20
<i>ObservationPeriod</i> <i>2</i>					
<i>Project</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
Mushuau Innu	82	77	52	43	34
Banking networks	67	66	57	45	46
Inca statisticians	60	72	78		
Palestinians	36	58	53	22	
Modern IT systems	28	25	18	25	14
Maritime security	23	108			

Average display time for the project pages was 7.8 seconds. In general, project introductory pages were viewed longest, 9 sec. Display times decreased by path length for the detail pages. There are several expansion pages for charts presented on detail pages. Display times for the enlargement pages averaged 8¾ sec., indicating that, when selected, they were also 'studied'.

Index selection accounts for nearly 60% of the page displays, which is as expected for a 2-level index structure. The main theme index and overview index were (re)selected in more than 55% of the sessions indicating usage of these as orientation and navigation tools. Index page display time averages showed little variation, averaging 8.8 seconds with variations from 8.4-10.4 sec.

## Summary and discussion

Table 4 gives a brief summary of the general characteristics of the four studies presented in this paper, S1 and S2 in [Shneiderman1989], Yamada in [Yamada1995], and N&N in the current paper. Session lengths are particularly difficult to compare in as much as data collection and analysis techniques differ and are not fully described in the papers referenced. Further, the studies differ in goal, and the exhibits differ in form and relationship to their surroundings. None the less, it appears that user interest in a hypertext/hypermedia exhibit is correlated to the degree that the exhibit supports and enhances the theme of its environment.

**Table 4: Usage study comparison**

<i>Exhibit report</i>	<i>Relationship to surroundings</i>	<i>Page type</i>	<i>Session length</i>	
			<i># pages</i>	<i>Time min.</i>
S1	Explanatory	Text	4.2	6
Yamada	Explanatory	Video	15	5
S2	Extension	Text + image	8.5	2.6
N&N	Independent	Text + image	5.4	1

### *Visitor interest*

The exhibit we studied consisted of introductory level, multi-topic presentations of social science projects, unrelated to the other themes in the exhibit room or to the themes of the natural science museum. Thus the exhibit visitors were unprepared (by the museum environment) for the thematic content of the hypertext exhibit. In this respect, these users can be compared to web surfers exploring a virtual museum exhibit.

Visitors spent on average, less than a minute at the exhibit and only half activated project presentations. Given that the visitors had come to a natural science museum, it is likely that the social science topics were not expected and/or not of general interest.

### *Topic interest*

Those visitors who selected project presentations spent 1.5 minutes, including index selections, viewing an average of 8-10 pages and 1.2 to 1.6 projects. Individual project presentations were viewed for about half a minute. The most frequently selected project, "Mushuau Innu – Canadian Indians", appears to have been of general interest. It was initially placed as the first theme and project in the index lists, but was still the most frequently

selected project when placed further down in the index structure.

Selection frequency for the detail pages fell from 67% to 33% for the 1<sup>st</sup> to 4<sup>th</sup> pages, respectively. Our analysis indicates that both project selection and the display frequency of individual pages depend significantly on project sequence in the indexes and sequence among the detail pages.

### *Navigation within the exhibit*

A 2-level index provided initial access to the project presentations. Further project selection could be made through the overview index or via backtracking to a higher level index. The overview index was intended as an orientation aid to provide support for topic change. About 60% of the accesses from the overview page were to the project indexes or presentations, indicating that the index functioned as intended.

In addition to the navigation bar, each of the project introduction pages had embedded links, as underscored words or phrases in the text, as list elements, and/or as highlighted images which allowed selection by interest (associative access). Associative page selection was used for only 25% of the detail page selections. The remaining 75% were accessed using the 'next-page' button in the navigation bar. It may be that the museum visitors were unfamiliar with embedded hypertext links. Alternatively, they may simply have chosen a familiar, serial navigation through the project presentations.

## Conclusions and suggestions

Our main objective has been to study usage characteristics of hypermedia exhibits. We have focused on user interest, as recorded in navigation paths and session length in time and page number, as well as the use of nonlinear navigation tools, particularly the use of embedded hypertext links. We had anticipated that topics would be selected by subject interest, but found that index placement dominates subject selection.

Based on the assumption that museum users of independent multimedia exhibits are similar in their information search behavior to that one can expect of web browsers, our study indicates that:

- *Visitor sessions* can be expected to be short in both time, 1-5 minutes, and number of pages, < 10. One can expect that 'normal' sessions will consist of 2-3 related topics of 2-3 pages. This indicates that care must be taken on topic choice, presentation, and length.
- *Project selection* and viewing time will depend significantly on project/page placement in the

indexes. Thus there is a continued need for careful index and exhibit construction.

- **Exhibit navigation** using a hierarchical index structure with an overview index will function well and is suitable for relatively small exhibits, as measured in number of topics. A useful supplement for larger exhibits may be keyword selection as used in traditional library and database systems.

Serial navigation, using the 'next-page' option, was chosen 3 times as often as embedded links indicating a need to support both serial and nonlinear navigation.

Our data do not indicate user disorientation or cognitive overload. It appears that the navigation tools provided sufficient orientation within the exhibit, as indicated by the relatively high use of the overview index and the tendency to exit from an index.

## Further studies

This has been a pilot study of the first use of an electronic exhibit for this museum. Our observations and experiences are expected to help improve the technical characteristics of coming exhibits. In addition, we expect that our observations of museum user behavior can be relevant for the design of web-based virtual exhibits.

The stand-alone exhibit has been moved from the natural science museum to the main building of the School of Social Science where it functions as an information directory for the school. Usage data have been collected for Jan-Feb'98 when new social science students began their studies. We are currently collecting exhibit usage data for the current semester students. We expect that analysis of these data will show increased usage length in time and number of project selections to the more theme compatible environment.

Design of a usage study for a web database version of the exhibit is underway with the goal of studying the search behavior of web users. The web version is available at <http://129.177.34.238/museum/00sv-ut.htm> (see also the above figures for page examples). We expect that web users will make more active use of the embedded links within the exhibit.

We lack demographic data about the exhibit users allowing us to distinguish "children at play" from general and/or specific information seekers. A user questionnaire will be included in the web exhibit.

## Acknowledgements

We would like to thank our colleagues at the School of Social Science, who submitted material on which the hypertext exhibit was based. We are also indebted to the staff at Bergen Museum for access to the museum facilities. Thanks also to Trond Pedersen and Bjarte Johnsen for their assistance in exhibit construction and data analysis.

This study has been performed within the framework of the Statistical Information Systems, SIS, project, as is also one of the projects in the Group for Interface Research, GIR, both headed by the authors. (see also <http://129.177.34.238/sis/sis.html-ssi> and <http://129.177.34.238/gir/html/gir.html-ssi>.)

## 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>.
- Henriksen, G. (1996). The Mushuau Innu Indians of Labrador. Dept. of Social Anthropology, University of Bergen.
- MacKenzie, D. (1996). Beyond Hypertext: Adaptive Interfaces for Virtual Museums. <http://www.dmcsoft.com/tamh/papers/evaf.php3>.
- 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](http://www.culture.fr/lumiere/documents/files/imaginary_exhibition.html)
- Nelson, T.H. (1967). Getting it out of our system. In *Information Retrieval: A Critical Review*. G.Schechter, ed. Thompson Books, 191-210.
- Preece, J. et.al. (1994). *Human-Computer Interaction*. Addison Wesley.
- Shneiderman, B., et.al. (1989). Evaluating Three Museum Installations of a Hypertext System. *Journal of the American Society for Information Science*, **40**(3), 172-182.
- Shneiderman, B. (1992). *Designing the User Interface - Strategies for effective Human-Computer Interaction*, 2nd ed. Addison-Wesley.
- Yamada, S., et.al. (1995). Development and evaluation of hypermedia for museum education: validation of metrics. *ACM Trans. of Computer-Human Interaction*, **2**(4), 284-307.