

# Digital Library Curriculum Development

## Module: 6-b: Search strategy, information seeking behavior, user modeling

(Draft, Last Updated: 2007/11/04)

1. **Module name:** Search strategy, information seeking behavior, user modeling

### 2. Scope

This module focuses on user perspectives in searching information in digital libraries. It covers the theories and practice related to online information seeking behaviors in different context of digital libraries.

### 3. Learning objectives:

Students will be able to:

- a. Identify the fundamental concepts, definitions, and theoretical models of online information seeking behaviors, as they apply to digital libraries
- b. Apply the models of information seeking behaviors in explaining the user behaviors that have been identified in empirical studies of digital libraries
- c. Utilize practical methods (e.g., transaction log analysis, interviews) for understanding information seeking behaviors in digital libraries

### 4. 5S characteristics of the module

- Stream, Structure, Space, & Society: N/A
- Scenarios: People look for information in digital libraries with their own histories or scenarios of information seeking.

### 5. Level of effort required:

- a. Class time: 3 hours
  - Session 1 (1 ½ hrs): Introduction to information seeking behavior models
    - Lecture on theories of information seeking behaviors
    - Brief review for the research methods (e.g., interviews and transaction log data) for the in-class activity and homework assignment.
    - In-class activity (See, 12. Exercises/Learning Activities, a. In class Exercise, below)
  - Session 2 (1 ½ hrs) : Empirical studies of the development of search strategies in the context of digital libraries
    - Discussion with the interview practice results from the homework assignment
    - Lecture on case studies
- b. Course preparation for students: 2 hours / session

1 · Mostly associated with readings (See the reading list for students in 11. Resources,  
2 below)

3 c. Course assignment completion: 2 hours

4 · User data collection practice (See 12. Exercises/Learning Activities, b.  
5 Homework Assignment, below)

## 6. Relationships with other modules

7 a. 6-a: Information Needs, Relevance

8 · *6-a: Information Needs, Relevance* covers the inner state of people's minds and  
9 information needs when they are about to start information seeking, and how that  
10 state evolves during their browsing and searching for information. *6-a Search*  
11 *Strategies, Information Seeking Behavior, and User Modeling* is different from  
12 this module in that it deals with the behavioral patterns and the characteristics of  
13 information seeking actions.

14 b. 6-d: Interaction Design, Info Summarization and Visualization, Usability Assessment

15 · *6-d: Interaction Design, Info Summarization and Visualization, Usability*  
16 *Assessment* identifies theoretical and practical knowledge related to the design  
17 and development of the system interfaces of digital libraries, while *6-a Search*  
18 *Strategies, Information Seeking Behavior, and User Modeling* provides clues and  
19 directions for how to use digital libraries in the overall context.

## 7. Prerequisite knowledge required:

20 · No prerequisite courses are required.

21 Taking a research method course prior to this course may help students to deal  
22 with the interviews in their homework assignments, but the instructor will provide  
23 basic guideline for the interviews for the students who don't have experiences  
24 with the method.  
25

26 · No technical programming or any other skills are necessary.

## 8. Introductory remedial instruction: None

## 9. Body of knowledge

### Session 1: Introduction to Information Seeking Behaviors

#### A. Basic Concepts & Definitions (Case, 2002)

30 a. Information need:

31 · A motivational state in people's minds that causes thoughts and actions through  
32 which they obtain something they desire.

33 · A state in which a person recognizes that his/her knowledge is inadequate to  
34 satisfy a goal that he/she has.  
35

36 b. Information seeking / Information seeking behaviors

37 · Actions that people do in response to their information needs

- 1           · Intentional activities to satisfy information needs
- 2           · What people do in response to information needs/tasks/problems
- 3                     “The purposive seeking for information as a consequence of a need to
- 4                     satisfy some goal” (Wilson, 1996)
- 5           · A process of discovering patterns or filling in gaps in patterns
- 6                     “A process in which humans purposefully engage in order to change their
- 7                     state of knowledge which is closely related to learning and problem
- 8                     solving” (Marchionini, 1998)
- 9       c. Information Browsing, Searching & Seeking
- 10           · Terms used interchangeably in literature
- 11           · Need to understand the appropriate meaning in the context of literature
- 12           · General meanings
- 13                     ○ Browsing: Informal and opportunistic behaviors moving across documents,
- 14                     examining them closely to see whether they satisfy the information needs or
- 15                     not
- 16                     ○ Searching: The actions taken when trying to find something
- 17                     The scope of meaning, sometimes, can be narrowed as being tied to specific
- 18                     searching techniques (e.g. Keyword-searching which enables people to
- 19                     retrieve information by inputting keywords in search engines)
- 20                     ○ Seeking: While browsing and searching are somewhat focused on behaviors
- 21                     or actions of people, information seeking covers more general categories and
- 22                     the broad nature of cognitive, perceptual, behavioral, and the environmental
- 23                     perspectives of people who are seeking information.

24

## 25       **B. Theoretical Models of Information Seeking in Online Environments**

### 26       a. Information Seeking in Online Environments

#### 27           **Allen’s Information Tasks (1996)**

- 28           · Scanning: Scanning tasks, resources, systems and acquiring general information
- 29           · Reviewing and Evaluating: Filtering information based on personal criteria (e.g.
- 30                     relevance judgments)
- 31           · Learning: Solving problems based on collected information and creating new
- 32                     knowledge
- 33           · Planning: Constructing search plans, evolving and developing them until the
- 34                     searches end

#### 35           **Marchionini’s Browsing Strategies (1999)**

- 36           · Scanning: Linear and/or selective scanning

- 1 · Observation: Capturing scenes, and making sense of what people see
- 2 · Navigation: Having specific goals or objectives when browsing, and developing
- 3 decision-making progress while organizing and structuring new knowledge
- 4 · Monitoring: Casual scanning of environments in order to sense what's in the
- 5 user's area of interest

6 **Choo, Detlor & Turnbull's Information Seeking on the Web (2000)**

- 7 · Starting: Identifying sources of information and initiating the search
- 8 · Changing: Modifying and re-directing the search process in order to locate
- 9 information
- 10 · Browsing: Scanning resources to understand the contents
- 11 · Differentiating: Evaluating and selecting appropriate resources
- 12 · Monitoring: Reviewing core sources to be acknowledged by news or changing
- 13 conditions
- 14 · Extracting: Working on a selected source to identify whether it is the most
- 15 appropriate source and to find connections to other related sources

16 **Model Summary**

Summary	Allen (1996)	Marchionini (1999)	Choo, et al. (2000)
· Initiating with information needs to solve problems, or make decisions	Information Needs		
· Scanning, learning about resources · Building general to specific ideas of what resources fit with information needs	Scanning	Scanning Observation	Starting Changing Browsing
· Narrowing down the scope of resources to what is really appropriate · Filtering out too broad, less related resources to information needs which evolve during the browsing.	Reviewing & Evaluating	Navigation	Differentiating Extracting
		Monitoring	Monitoring
· Knowledge building · Continues development of information seeking	Learning	-	-
	Planning		

- 18 · Allen (1996)
- 19 - Describing the overall process of information seeking from information
- 20 seeking to knowledge development

- 1           - Addressing the continuous nature of information seeking
- 2           · Marchionini (1999)
- 3           - Specifying processes of what people see and capture in their minds
- 4           · Choo, et al (2000)
- 5           - Focusing on how people deal with, manage and evaluate resources

6           **Common points**

- 7           · Models illustrate linear sequences of information seeking, although there are
- 8           loops and repetitive actions.
- 9           · It is assumed that people develop their search from broader topics to narrower
- 10           topics.
- 11           · Information needs evolve during the search.
- 12           · People continually evaluate information based on the judgments they make during
- 13           their search.
- 14           · People's knowledge structures evolve during the search.
- 15           · Although it is hard to identify the definitive end of a search which indicates that
- 16           the search is completed, people have strategies when closing their current search
- 17           session.

18           **b. Search Strategy Development in Online Environments**

19           **Chowdhury & Chowdhury's the Four-Phase Framework for Information**

20           **Search in digital libraries (2003)**

- 21           · Phase 1: Formulation:
- 22           - Triggered by an information need
- 23           - Early decision-making stages regarding what to search for and the search
- 24           variants
- 25           · Phase 2: Action
- 26           - Executing information searching with queries
- 27           - Associated with search queries and strategies development
- 28           · Phase 3: Review of Results
- 29           - Utilizing various tools to display the search results (e.g. sorting, highlighting)
- 30           · Phase 4: Refinement
- 31           - Modifying and reformulating search queries, conducting new searches

32           **Marchionini's Analytical Search Strategies (1999)**

- 33           · Building blocks

- 1                   - Starting searching with basic concepts and then expanding the scope of the
- 2                   search results by combining additional concepts from the search results
- 3                   obtained during the searches
- 4                   · Successive fractions
- 5                   - Using a large subset of the entire database and breaking it down into pieces of
- 6                   concepts relevant to a problem while conducting the searches
- 7                   · Pearl growing
- 8                   - Selecting a base document (a pearl) and using its characteristics such as index
- 9                   terms, title, text, citations, to extract and formulate search terms in order to
- 10                  locate documents relevant to the original information needs
- 11                  · Interactive scanning
- 12                  - Beginning the search with a large set of documents generally relevant to the
- 13                  problem and scanning them to extract their key common features and using
- 14                  them to formulate queries

15                  c. **Research Methods**

- 16                  · Interviews (Berg, 2004)
- 17                    o Types of Interviews
- 18                    o Interview Preparation
- 19                    o Tips for Interviewing
- 20                    o Interview data analysis
- 21                  · Transaction Log Analysis

22

23                  **Session 2: Case Studies of the Search Strategy Development in Digital Libraries**

24                  **A. Digital Library Users**

- 25                  · Quantitatively larger groups of people are involved, as compared to users of other
- 26                  types of libraries or systems
- 27                  · Geographically distributed
- 28                  · Individual differences (e.g., technical skills, learning / problem-solving, decision-
- 29                  making abilities, cognitive styles, personal characteristics, etc.)
- 30                  · Various levels of domain knowledge, computer skills, and familiarity with
- 31                  systems, information searching experiences

32

33                  **B. Information Seeking Behaviors in Various Contexts of Digital Libraries**

34                  **Case Study 1: Information seeking behaviors in databases in digital libraries**

1 Park, S. (2000) Usability, User Preferences, Effectiveness, and User Behaviors When  
2 Searching Individual and Integrated Full-Text Databases: Implications for Digital  
3 Libraries. *Journal of the American Society for Information Science*, 51(5), 456-68.

#### 4 **Research Purpose**

5 To compare the user characteristics in information seeking, such as usability, user  
6 preference, effectiveness and searching behaviors, towards the use of different  
7 types (integrated levels) of interfaces of a digital library with multiple databases,  
8 focusing on how people interact with multiple heterogeneous information  
9 resources.

#### 10 **Research Questions**

- 11 · User preferences: Which system did you like better, find more useful, easier to  
12 use, and easier to learn to use?
- 13 · Usability & User satisfaction: How easy is it to do the search on the topic?  
14 How satisfied are you with your search results? Did you have enough time to  
15 do an effective search?
- 16 · Effectiveness: Is there a statistically significant difference in performance  
17 (aspectual recall) between systems?
- 18 · Search behaviors: Is there a statistically significant difference in individual's  
19 searching behavior (e.g. the usage of frequency of various system features)  
20 between systems?

#### 21 **Methods**

- 22 · Text Retrieval Conference (TREC) resource use for searches  
23 (Topics of resources & search tasks: Federal Register, Wall Street Journal,  
24 Financial Times, & Congressional Record)
- 25 · Experimental system interfaces  
26 HERA: An integrated interface with multiple databases  
27 HERMIS: Common interface, individual interface access to multiple  
28 databases
- 29 · Participants: 28 graduate students
- 30 · Conducting searching experiments  
31 (Each participant conduct 3 searches with given tasks on each system)
- 32 · Post system and post search questionnaire completion
- 33 · Interviews

#### 34 **Findings**

- 35 · Searcher responses to each system regarding advantages and disadvantages to use
- 36 · Users preferred interacting with different databases through a common interface  
37 than interacting with databases through an integrated interface.

- 1           · User desires to control database selection, system features, relevance feedback

2

3           **Case Study 2: Scholars and professionals' information use in digital libraries I**

4           Nicholas D, Huntington P, and Watkinson A. (2005). Scholarly Journal Usage:  
5           The Results of a Deep Log Analysis. *Journal of Documentation*. 61(2), 248 -280

6           **Research Purpose**

7           To investigate the impact of digital roll-out and information seeking behaviors of  
8           academics and researchers, regarding their use of digital journals; Three main  
9           objectives of the study are;

10          1) To map the usage of the Blackwell scholarly database, including types of items  
11          viewed/requested, number of items viewed in a session, and repeat visits.

12          2) To categorize the characteristics of journal users according to their  
13          geographical location, job status, subject of journals searched, items bought  
14          online, their method of liking to the site.

15          3) To demonstrate the “deep log analysis” methodology in the field.

16          **Methods**

- 17          · Deep log analysis: Four step process

18               1) Statistical significance assessment of log data regarding how the data are  
19               defined and recorded in logs (e.g., who is a user, what is hit, etc) are analyzed and  
20               their statistical significance is assessed.

21               2) Data re-engineering to provide metrics

22               3) Adding user demographics to the data, obtained from user profile or online  
23               survey

24               4) Conducting additional surveys, interviews, or observation to explain the user  
25               satisfaction and the impact of the data to their seeking behaviors.

- 26          · Collecting raw server transaction logs for a month in 2003 of Blackwell Synergy;  
27          Log data created from the use of 500,000 people

28          **Findings; user data demonstration**

- 29          · Type of item viewed.

- 30          · Time spent viewing.

- 31          · Daily usage patterns.

- 32          · Subject of items requested.

- 33          · Place where the journal was published.

- 34          · Users defined by referral link used.

- 35          · User defined by use/non-use of Athens.

- 36          · Users defined by subject of journal searched.



- 1           · Users defined by whether they bought articles online.
- 2           · Further refinements on the site penetration metric.

3

## 4 **10. Resources**

### 5 **Reading list for students**

#### 6 **Session 1: Required**

7           Allen, B. L. (1996). Chapter 7, Information Tasks: Interacting with Information Systems,  
8           In *Information Tasks: Toward a User-Centered Approach to Information Systems*  
9           (pp.188-200). San Diego: Academic Press

10          Choo, W. C., Deltor, B., & Turnbull, D. (2000). Information Seeking on the Web: An  
11          Integrated Model of Browsing and Searching. *First Monday*, 5 (2).

12          Chowdhury, G. G. & Chowdhury, S. (2003). Chapter 8, Information Access and User  
13          Interfaces and Chapter 9, Information Retrieval in Digital Libraries. In *Introduction to*  
14          *Digital Libraries* (pp.152-213), London: Facet Publishing.

15          Marchionini, G. (1999). Chapter 6, Browsing Strategies. In *Information Seeking in*  
16          *Electronic Environments* (pp.100-138). NY: Cambridge University Press.

#### 17 **Recommended**

18          Borgman, C. L. (2000). Chapter 5, Why are Digital Libraries Hard to Use. In *From*  
19          *Gutenberg to the global information infrastructure* (pp. 117-141). Cambridge, MA: MIT  
20          Press Buttenfield

21          Case, D. O. (2002). *Looking for Information: A Survey of Research on Information*  
22          *Seeking, Needs, and Behavior*. New York: Academic Press.

23          Chang, S-J & Rice, R. E. (1993). Browsing: A Multidimensional Framework. In M.E.  
24          Williams (Ed.), *Annual Review of Information Science and Technology* (pp. 231-271), 28.  
25          Medford, NJ: Learned Information.

#### 26 **Session 2: Required (Case Studies)**

27          Park, S. (2000) Usability, User Preferences, Effectiveness, and User Behaviors When  
28          Searching Individual and Integrated Full-Text Databases: Implications for Digital  
29          Libraries. *Journal of the American Society for Information Science*, 51(5), 456-68.

30          Nicholas D, Huntington P, and Watkinson A. (2005). Scholarly Journal Usage: The  
31          Results of a Deep Log Analysis. *Journal of Documentation*. 61(2), 248 -280

32

### 33 **Readings for Instructors**

#### 34 **General Theories and Models of Information Seeking Behaviors in Online** 35 **Environments**

36          Allen, B. L. (1996). Chapter 7, Information Tasks: Interacting with Information Systems.  
37          In *Information Tasks: Toward a User-Centered Approach to Information Systems*  
38          (pp.188-200). San Diego: Academic Press

- 1 Case, D. O. (2002). *Looking for Information: A Survey of Research on Information*  
2 *Seeking, Needs, and Behavior*. New York: Academic Press.
- 3 Chang, S-J & Rice, R. E. (1993). Browsing: A Multidimensional Framework. In M.E.  
4 Williams (Ed.), *Annual Review of Information Science and Technology* (pp. 231-271), 28.  
5 Medford, NJ: Learned Information.
- 6 Choo, W. C., Deltor, B., & Turnbull, D. (2000). Information Seeking on the Web: An  
7 Integrated Model of Browsing and Searching. *First Monday*, 5 (2).
- 8 Marchionini, G. (1999). Chapter 6, Browsing Strategies. In *Information Seeking in*  
9 *Electronic Environments* (pp.100-138). NY: Cambridge University Press.
- 10 Wilson, T. D. (2000). Human Information Behavior. *Informing Science*, 3(2), 49-56.  
11 Retrieved September 6, 2007 from: <http://inform.nu/Articles/Vol3/v3n2p49-56.pdf>

## 12 **Research Methods for Information Seeking Behavior Studies**

- 13 Wang, P. (1999). Methodologies and Methods for User Behavioral Research. In M.E.  
14 Williams (Ed.), *Annual Review of Information Science and Technology* (pp. 53-99), 34,  
15 Medford, NJ: Information Today
- 16 Berg, B. L. (2004). *Qualitative Research Methods for the Social Sciences*. (5th ed).  
17 Boston: Pearson.
- 18 Burgess, R. G. (Ed.) (1986). Chapter 11, Elements of Sampling in Field Research. In  
19 *Field Research: A Sourcebook and Field Manual* (pp. 75-78). UK: Routledge.
- 20 Honigmann, J. J. (1986). Chapter 12, Sampling in Ethnographic Fieldwork. In Burgess, R.  
21 G. (Ed.), *Field Research: A Sourcebook and Field Manual* (pp. 79-90). UK: Routledge.
- 22 Jones, S., Cunningham, S. J., McNab, R., & Boddie, S. (2000). A Transaction Log  
23 Analysis of a Digital Library. *International Journal on Digital Libraries*, 3, 152-169.

## 24 **Search strategy development in online environments**

- 25 Aula, A., N. Jhaveri, and M. Käki. (2005) Information Search and Reaccess Strategies of  
26 Experienced Web Users. In *Proceedings of the 14th International Conference on the*  
27 *World Wide Web* (pp. 583-592), ACM Press.
- 28 Debowski, S. (2002) Wrong Way: Go Back! An Exploration of Novice Search  
29 Behaviours While Conducting an Information Search. *The Electronic Library*, 19(6),  
30 371-82.
- 31 Holscher, C. & Strube, G. (2000). Web Search behavior of Internet Experts and Newbies.  
32 *Computer Networks*, 33(1-6), 337-346
- 33 Kuhlthau, C. C. (1993). Chapter 3, The Information Search Process and Chapter 4,  
34 Verification of the Model of the Information Search Process. In *Seeking Meaning: A*  
35 *Process Approach to Library and Information Services* (pp.33-63). Norwood, New  
36 Jersey: Ablex Publishing Company.
- 37 Hsieh-Yee, I. (1993). Effects of Search Experience and Subject Knowledge on the Search  
38 Tactics of Novice and Experienced Searchers. *Journal of the American Society for*  
39 *Information Science* 44(3), 161-174.

1 Marchionini, G., Dwiggins, S., Katz, A, and Lin, X. (1993). Information Seeking in Full-  
2 Text End-User-Oriented Search Systems: The Roles of Domain and Search Expertise.  
3 *Library & Information Science Research*, 15(1), 35-69.

#### 4 **Developing searching and browsing strategies in the context of digital libraries**

5 Bates, M. J. (2002). The Cascade of Interactions in the Digital Library Interface.  
6 *Information Processing & Management* 38(3), 381-400.

7 Bates, M. J. (1998). Indexing and Access for Digital Libraries and the Internet: Human,  
8 Database, and Domain Factors. *Journal of the American Society for Information Science*,  
9 49(13), 1185-205.

10 Bishop, A. P. & Star, S. L. (1996). Social informatics of digital library use and  
11 intrastructure. In M. Williams (Eds.), *Annual Review of Information Science and*  
12 *Technology* 31 (pp.301-401). Medford, NJ: Information Today, Inc.

13 Bishop, A. P, Neumann, L. J., Star, S. L., Merkel, C., Ignacio, E. & Sandusky, R. J.  
14 (2000). Digital Libraries: Situating Use in Changing Information Infrastructure. *Journal*  
15 *of the American Society for Information Science*, 51(4), 394-413

16 Borgman, C. L. (2000). Chapter 5, Why are Digital Libraries Hard to Use In *From*  
17 *Gutenberg to the Global Information Infrastructure* (pp. 117-141). Cambridge, MA: MIT  
18 Press Buttenfield

19 Chowdhury, G. G. & Chowdhury, S. (2003). Chapter 8, Information Access and User  
20 Interfaces and Chapter 9, Information Retrieval in Digital Libraries. In *Introduction to*  
21 *Digital Libraries* (pp.152-213), London: Facet Publishing.

22 Das Neves, F.A. & Fox, E. (2000). A Study of User Behavior in an Immersive Virtual  
23 Environment for Digital Libraries. In *Proceedings of the Fifth ACM Conference on*  
24 *Digital Libraries*, 103-111.

25 Park, S. (2000) Usability, User Preferences, Effectiveness, and User Behaviors When  
26 Searching Individual and Integrated Full-Text Databases: Implications for Digital  
27 Libraries. *Journal of the American Society for Information Science*, 51(5), 456-68.

28 Paepcke, A. et al. (2000). Search Middleware and the Simple Digital Library  
29 Interoperability Protocol. *D-Lib Magazine*, 6 (10).

30 Schatz, B. R. (1997). Information Retrieval in Digital Libraries: Bringing Search to the  
31 Net, *Science*, 275(5298), 327-334.

32 Smith, A. G. (2000) Search Features of Digital Libraries. *Information Research*, 5(3).

33 Spink, A., Wilson, T., Ellis, D., and Ford, N. (1998). Modeling Users' Successive  
34 Searches in Digital Environments: A National Science Foundation/British Library  
35 Funded Study. *D-Lib Magazine*, April.

36 Borgman, C. L., et. al. (2005). Comparing Faculty Information Seeking in Teaching and  
37 Research: Implications for the Design of Digital Libraries. *Journal of the American*  
38 *Society for Information Science and Technology*, 56(6), 636-657.

39 Payette, S. & Rieger, O. (1998). Supporting Scholarly Inquiry: Incorporating Users in the  
40 Design of the Digital Library. *Journal of Academic Librarianship*, 4(2), 121-129.

1 Nicholas D, Huntington P, and Watkinson A. (2005). Scholarly Journal Usage: The  
2 Results of a Deep Log Analysis. *Journal of Documentation*. 61(2), 248 -280

### 3 **11. Concept Map**

### 4 **12. Exercises / Learning activities**

- 5 a. In-class Exercise: A Small Group Activity of the Transaction Log Analysis (See,  
6 Appendix A: Transaction Log Data Analysis Practice, as an exercise example)

7 The purpose of this exercise is to provide opportunities in which students look at a set  
8 of transaction log data from an information searching system and practice how to  
9 interpret and analyze the raw data and make them meaningful in understanding the  
10 user behaviors in searching.

11 The students will be assigned to groups of 3 or 4. The instructor will assign a sample  
12 transaction log data set to each group. The sample data is collected at the Davis  
13 Library at the University of North Carolina at Chapel Hill on December 14<sup>th</sup>, 2002.  
14 65 subjects participated in the study and log data from one subject is included here  
15 (Subject ID: x68). (Wildemuth, B. M, & Oh, S. (2006) Congruence between  
16 Information Known and Information Used in Conducting Known-Item Searches,  
17 Presentation at the Annual Research Forum of Librarian's Association of UNC-CH  
18 (LAUNC-CH), May 23.)

19 All groups will have the same data set for the class discussion later. The instructor  
20 provides the background information about the original information system, at least  
21 about the types of information systems, how and when the transaction log was  
22 captured, and if necessary, how the interfaces of the information system look. No user  
23 information needs to be released by the instructor, since the students would go  
24 through an inference process of analyzing the user characteristics from the exercise.

25 Each group will have 20 minutes to discuss the data. While reviewing the logs, the  
26 students will be asked to list the entries that they could identify, to extract any  
27 information they analyzed, and any descriptions about user behaviors in using the  
28 information system in as much detail as they can.

29 When the group discussion ends, the class will have a 10 to 20 minute session for a  
30 class discussion with the findings from the groups (See the Appendix A for the  
31 exercise example).

32 Here are examples of items that the student can capture from the transaction log data  
33 (Jones, 2000):

- 34 · User demographic information, if users are asked to provide their demographic  
35 information before starting searches
- 36 · Use of operators
- 37 · Search options of the queries
- 38 · Patterns in query construction and refinement
- 39 · Common mistakes in searching, like using incorrect spellings of the queries

- 1           ·   Distribution of query terms (e.g., popular terms, newly appearing terms, terms no  
2           longer in use, etc.)
- 3           ·   Frequency of query terms (e.g., ranking of the terms, the terms that the recent  
4           frequencies increase or decrease, etc.)

5

6           b. Homework Assignment: User Data Collection Practice

7           This exercise gives the students the opportunity of being able to meet users and  
8           investigate their information seeking behaviors while using a digital library.

9           This is a group exercise. The students will be grouped into 3 or 4. First, each group  
10          will pick a real or hypothetical digital library in any context. Since they need to find  
11          practice users (at least 1 per each group members) who may be interested in using the  
12          digital library, it would be easier if they chose an existing digital library and have  
13          interviews with people who have experiences using it. On the other hand, they could  
14          imagine that they plan to build a digital library for a specific target group (e.g.,  
15          children, students, health professionals, film makers, movie goers, etc) and interview  
16          the target group's information seeking behavior patterns.

17          Interviews will be conducted with at least 1 people per each group members on how  
18          they search or would search in the selected digital libraries. The interview  
19          questionnaires will be designed by the students, but they will include lists of  
20          questions that collect data about the measurement criteria that was discussed in class.

21          After the interviews, the students will analyze the interview results and evaluate  
22          whether the results would explain how people search for information.

23          The students will bring the interview results to the group in the next class session,  
24          discuss the issues that they found among group and share the results in the class  
25          discussion.

26

27       **13. Evaluation of learning outcomes (iterative to mastery)**

- 28          a. Do students understand well the fundamental concepts, definitions and theoretical  
29          models of online information seeking behaviors?
- 30          b. Did students build their own knowledge of the search strategies and the challenges of  
31          information seeking in digital libraries?
- 32          c. Did students successfully conduct the user interviews through the homework exercise  
33          and collect appropriate user data?
- 34          d. Do students understand well how the theoretical models of information seeking  
35          behaviors are applicable in digital library settings based on the analysis of the user  
36          data they collected?

37

38       **14. Glossary**

39

1 **15. Additional useful links**

2

3 **16. Contributors**

4 a. Initial author: Sanghee Oh

5 b. Team evaluators: Jeff Pomerantz, Barbara Wildemuth

6 c. Other evaluators: Ashwini Athavale (Indiana University), Stephanie Hass (UNC-CH),  
7 Maureen Henninger, (UTS, Australia), and Lindley Shedd (Indiana University).

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

## Appendix A: Transaction Log Data Analysis Practice

### Log Analysis Activity Guideline

The purpose of this exercise is to provide opportunities in which students look at a set of transaction log data from an information searching system and to practice interpreting and analyzing the raw data and in order to make them meaningful for understanding user behaviors in searching.

Make a group with 3 or 4 students. Each group will receive a set of log data samples and a questionnaire. Read the log data carefully, extract the necessary data for analysis and answer the questions.

### Log Data Reading Tips

There are various types of log data and they are displayed differently according to the system design. The sample data, here, is just one example. The following reading tips are not universal rules for reading the data, but they can be good practice for future analysis. Tips are provided for easy understanding for the data analysis in the limited class time. If there is enough time to manipulate the data, ask the students to figure out the function of certain items and extract data based on their configuration.

- Separate each session according to the time stamp.
- Check the “QUERY” item to extract the queries entered by users.
- Check the “button\_clicked” item to find which field was chosen for the search.
- Check the “filter” item to find whether the user select the filtering option.
- Remember that users sometimes need time to review the search results, traveling back and forth to the results pages.
- Use your highlight pen to mark the necessary data to use in the analysis of the data set.

### Log Data Information

- Collected at the Davis Library at the University of North Carolina at Chapel Hill on December 14<sup>th</sup>, 2002.
- 65 subjects participated in the study and log data from one subject is included here (Subject ID: x68).
- Reference: Wildemuth, B. M, & Oh, S. (2006) Congruence between Information Known and Information Used in Conducting Known-Item Searches, Presentation at the Annual Research Forum of Librarian’s Association of UNC-CH (LAUNC-CH), May 23.





## Answer Guideline for the Questions

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13

A1: About 12 to 14 times depending on whether the search result review sessions are counted or not.

A2: See Query in the table, below. There are several system languages shown in the queries. The spaces that the user put between queries are shown as “+”. “,” and “.” are shown as “%2C” and “%2E.”

A3: Check the log data written “button\_clicked” for the field selection, and the lines with “filter” information for the filtering selection.

A4: User behavioral characteristics: 1) Used the same keywords multiple times, 2) Aware of the field options, 3) Narrowed the results with the filtering options, 4) No missed spelling, 5) Bring relatively various information for the search (e.g., keywords, names, styles), etc.

Sessions	Time Stamp	Query	Field / Filtering Selection	Analysis
1	19:17:15	x68	Keyword	Entering ID
2	19:17:44	x68	Keyword	Entering ID again
3	19:19:50	x68++antique+clocks	Subject	Entering queries with ID, Conducting the subject search
4	19:20:48	antique+clocks	Subject	Entering queries without ID
5	19:21:30	antique+clocks	Subject	
6	19:24:42	Freund%2+Joan+Barzila	Filtering	Filtering the current results with additional queries
7	19:26:24	+clocks	Subject	
8	19:29:39	Astronomical+clocks+ Early+works+to+1800	Filtering	Filtering the current results with additional queries
9	19:30:46	+clocks	Subject	Conducting the search with the same keywords
10	19:32:31	+clocks	Subject	

11	19:33:00	+clocks	Subject	repeatedly for 1) simply regenerating the search results , 2) traveling back and forth to the search results pages, or something else. A clear reason is not known.
12	19:33:14	Hearn%2C+George+A%2E%2C+Mrs%2E	Filtering	Filtering the current results with additional queries

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24

## Transaction Log Data Sample

2002-10-14 19:17:15 152.2.71.24 - 152.2.181.16 GET /English/libraries.gif 304 141 1371 0 HTTP/1.1  
 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+4.0;+(R1+1.1))  
 http://web2.lib.unc.edu/web2/tramp2.exe/form/A0ab2jhd.000?QUERY\_SCREEN=Home.html&QUERY=x68&buttons=title%3Ddo\_authority\_search+\*search\_button%3Dtitle+index%3Dti&buttons=author%3Ddo\_a  
 uthority\_search+\*search\_button%3Dauthor+index%3Dau&buttons=subject%3Ddo\_authority\_search+\*se  
 arch\_button%3Dsubject+index%3Dsu&buttons=keyword%3Ddo\_ccl\_search+\*search\_button%3Dkeywor  
 d+index%3Ddefault&buttons=callno%3Ddo\_authority\_search+\*search\_button%3Dcallno+index%3D%25  
 &buttons=issn%3Ddo\_authority\_search+\*search\_button%3Dissn+index%3D%29&buttons=isbn%3Ddo\_a  
 uthority\_search+\*search\_button%3Disbn+index%3D%28&buttons=oclc%3Ddo\_authority\_search+\*searc  
 h\_button%3Doclc+index%3D%5E&buttons=lccn%3Ddo\_authority\_search+\*search\_button%3Dlccn+inde  
 x%3D%23&buttons=gov%3Ddo\_authority\_search+\*search\_button%3Dgov+index%3D%26&buttons=tech  
 %3Ddo\_authority\_search+\*search\_button%3Dtech+index%3D%26&buttons=music%3Ddo\_authority\_sea  
 rch+\*search\_button%3Dmusic+index%3D%26&button\_clicked=keyword&Search=Search&material\_filter  
 =all&language\_filter=all&date\_filter=all&servers=4home

2002-10-14 19:17:44 152.2.71.24 - 152.2.181.16 GET /web2/tramp2.exe 200 16089 1365 78 HTTP/1.1  
 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+4.0;+(R1+1.1))  
 http://web2.lib.unc.edu/web2/tramp2.exe/form/A0ab2jhd.000?QUERY\_SCREEN=Home.html&QUERY=x68  
 &buttons=title%3Ddo\_authority\_search+\*search\_button%3Dtitle+index%3Dti&buttons=author%3Ddo\_a  
 uthority\_search+\*search\_button%3Dauthor+index%3Dau&buttons=subject%3Ddo\_authority\_search+\*se  
 arch\_button%3Dsubject+index%3Dsu&buttons=keyword%3Ddo\_ccl\_search+\*search\_button%3Dkeywor  
 d+index%3Ddefault&buttons=callno%3Ddo\_authority\_search+\*search\_button%3Dcallno+index%3D%25  
 &buttons=issn%3Ddo\_authority\_search+\*search\_button%3Dissn+index%3D%29&buttons=isbn%3Ddo\_a  
 uthority\_search+\*search\_button%3Disbn+index%3D%28&buttons=oclc%3Ddo\_authority\_search+\*searc  
 h\_button%3Doclc+index%3D%5E&buttons=lccn%3Ddo\_authority\_search+\*search\_button%3Dlccn+inde  
 x%3D%23&buttons=gov%3Ddo\_authority\_search+\*search\_button%3Dgov+index%3D%26&buttons=tech  
 %3Ddo\_authority\_search+\*search\_button%3Dtech+index%3D%26&buttons=music%3Ddo\_authority\_sea  
 rch+\*search\_button%3Dmusic+index%3D%26&button\_clicked=keyword&Search=Search&material\_filter  
 =all&language\_filter=all&date\_filter=all&servers=4home

2002-10-14 19:19:50 152.2.71.24 - 152.2.181.16 GET /English/libraries.gif 304 141 1388 0 HTTP/1.1  
 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+4.0;+(R1+1.1))  
 http://web2.lib.unc.edu/web2/tramp2.exe/form/A0ab2jhd.001?QUERY\_SCREEN=Home.html&QUERY=x68  
 +++++antique+clocks&buttons=title%3Ddo\_authority\_search+\*search\_button%3Dtitle+index%3Dti&button  
 s=author%3Ddo\_authority\_search+\*search\_button%3Dauthor+index%3Dau&buttons=subject%3Ddo\_aut  
 hority\_search+\*search\_button%3Dsubject+index%3Dsu&buttons=keyword%3Ddo\_ccl\_search+\*search\_  
 button%3Dkeyword+index%3Ddefault&buttons=callno%3Ddo\_authority\_search+\*search\_button%3Dcalln  
 o+index%3D%25&buttons=issn%3Ddo\_authority\_search+\*search\_button%3Dissn+index%3D%29&butto  
 ns=isbn%3Ddo\_authority\_search+\*search\_button%3Disbn+index%3D%28&buttons=oclc%3Ddo\_authorit  
 y\_search+\*search\_button%3Doclc+index%3D%5E&buttons=lccn%3Ddo\_authority\_search+\*search\_butt  
 on%3Dlccn+index%3D%23&buttons=gov%3Ddo\_authority\_search+\*search\_button%3Dgov+index%3D%  
 26&buttons=tech%3Ddo\_authority\_search+\*search\_button%3Dtech+index%3D%26&buttons=music%3D  
 do\_authority\_search+\*search\_button%3Dmusic+index%3D%26&button\_clicked=subject&Search=Search  
 &material\_filter=all&language\_filter=all&date\_filter=all&servers=1home

2002-10-14 19:20:48 152.2.71.24 - 152.2.181.16 GET /English/libraries.gif 304 141 1383 0 HTTP/1.1  
 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+4.0;+(R1+1.1))  
 http://web2.lib.unc.edu/web2/tramp2.exe/form/A0ab2jhd.001?QUERY\_SCREEN=Home.html&QUERY=+  
 antique+clocks&buttons=title%3Ddo\_authority\_search+\*search\_button%3Dtitle+index%3Dti&buttons=aut  
 hor%3Ddo\_authority\_search+\*search\_button%3Dauthor+index%3Dau&buttons=subject%3Ddo\_authority  
 \_search+\*search\_button%3Dsubject+index%3Dsu&buttons=keyword%3Ddo\_ccl\_search+\*search\_button  
 %3Dkeyword+index%3Ddefault&buttons=callno%3Ddo\_authority\_search+\*search\_button%3Dcallno+ind  
 ex%3D%25&buttons=issn%3Ddo\_authority\_search+\*search\_button%3Dissn+index%3D%29&buttons=is  
 bn%3Ddo\_authority\_search+\*search\_button%3Disbn+index%3D%28&buttons=oclc%3Ddo\_authority\_se  
 arch+\*search\_button%3Doclc+index%3D%5E&buttons=lccn%3Ddo\_authority\_search+\*search\_button%3

Dlccn+index%3D%23&buttons=gov%3Ddo\_authority\_search+\*search\_button%3Dgov+index%3D%26&buttons=tech%3Ddo\_authority\_search+\*search\_button%3Dtech+index%3D%26&buttons=music%3Ddo\_authority\_search+\*search\_button%3Dmusic+index%3D%26&button\_clicked=subject&Search=Search&material\_filter=all&language\_filter=all&date\_filter=all&servers=5home

2002-10-14 19:21:30 152.2.71.24 - 152.2.181.16 GET /web2/tramp2.exe 200 11695 1394 10875 HTTP/1.1 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+4.0;+(R1+1.1))  
http://web2.lib.unc.edu/web2/tramp2.exe/form/A0ab2jhd.001?QUERY\_SCREEN=Home.html&QUERY=+antique+clocks&buttons=title%3Ddo\_authority\_search+\*search\_button%3Dtitle+index%3Dti&buttons=author%3Ddo\_authority\_search+\*search\_button%3Dauthor+index%3Dau&buttons=subject%3Ddo\_authority\_search+\*search\_button%3Dsubject+index%3Dsu&buttons=keyword%3Ddo\_ccl\_search+\*search\_button%3Dkeyword+index%3Ddefault&buttons=callno%3Ddo\_authority\_search+\*search\_button%3Dcallno+index%3D%25&buttons=issn%3Ddo\_authority\_search+\*search\_button%3Dissn+index%3D%29&buttons=isbn%3Ddo\_authority\_search+\*search\_button%3Disbn+index%3D%28&buttons=oclc%3Ddo\_authority\_search+\*search\_button%3Doclc+index%3D%5E&buttons=lccn%3Ddo\_authority\_search+\*search\_button%3Dlccn+index%3D%23&buttons=gov%3Ddo\_authority\_search+\*search\_button%3Dgov+index%3D%26&buttons=tech%3Ddo\_authority\_search+\*search\_button%3Dtech+index%3D%26&buttons=music%3Ddo\_authority\_search+\*search\_button%3Dmusic+index%3D%26&button\_clicked=subject&Search=Search&material\_filter=all&language\_filter=all&date\_filter=all&servers=5home

2002-10-14 19:24:42 152.2.71.24 - 152.2.181.16 GET /English/libraries.gif 304 141 553 0 HTTP/1.1 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+4.0;+(R1+1.1))  
http://web2.lib.unc.edu/web2/tramp2.exe/do\_authority\_search/A0ab2jhd.008?servers=5home&index=au&material\_filter=all&language\_filter=all&location\_filter=&location\_group\_filter=&date\_filter=all&query=Friend%2C+Joan+Barzilay%2E

2002-10-14 19:26:24 152.2.71.24 - 152.2.181.16 GET /English/libraries.gif 304 141 1375 0 HTTP/1.1 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+4.0;+(R1+1.1))  
http://web2.lib.unc.edu/web2/tramp2.exe/form/A0ab2jhd.001?QUERY\_SCREEN=Home.html&QUERY=+clocks&buttons=title%3Ddo\_authority\_search+\*search\_button%3Dtitle+index%3Dti&buttons=author%3Ddo\_authority\_search+\*search\_button%3Dauthor+index%3Dau&buttons=subject%3Ddo\_authority\_search+\*search\_button%3Dsubject+index%3Dsu&buttons=keyword%3Ddo\_ccl\_search+\*search\_button%3Dkeyword+index%3Ddefault&buttons=callno%3Ddo\_authority\_search+\*search\_button%3Dcallno+index%3D%25&buttons=issn%3Ddo\_authority\_search+\*search\_button%3Dissn+index%3D%29&buttons=isbn%3Ddo\_authority\_search+\*search\_button%3Disbn+index%3D%28&buttons=oclc%3Ddo\_authority\_search+\*search\_button%3Doclc+index%3D%5E&buttons=lccn%3Ddo\_authority\_search+\*search\_button%3Dlccn+index%3D%23&buttons=gov%3Ddo\_authority\_search+\*search\_button%3Dgov+index%3D%26&buttons=tech%3Ddo\_authority\_search+\*search\_button%3Dtech+index%3D%26&buttons=music%3Ddo\_authority\_search+\*search\_button%3Dmusic+index%3D%26&button\_clicked=subject&Search=Search&material\_filter=all&language\_filter=all&date\_filter=all&servers=4home

2002-10-14 19:29:39 152.2.71.24 - 152.2.181.16 GET /English/libraries.gif 304 141 569 0 HTTP/1.1 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+4.0;+(R1+1.1))  
http://web2.lib.unc.edu/web2/tramp2.exe/do\_authority\_search/A0ab2jhd.017?servers=4home&index=su&material\_filter=all&language\_filter=all&location\_filter=&location\_group\_filter=&date\_filter=all&query=Astrophysical+clocks+Early+works+to+1800%2E

2002-10-14 19:30:46 152.2.71.24 - 152.2.181.16 GET /web2/tramp2.exe 200 13597 1382 1343 HTTP/1.1 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+4.0;+(R1+1.1))  
http://web2.lib.unc.edu/web2/tramp2.exe/form/A0ab2jhd.001?QUERY\_SCREEN=Home.html&QUERY=+clocks&buttons=title%3Ddo\_authority\_search+\*search\_button%3Dtitle+index%3Dti&buttons=author%3Ddo\_authority\_search+\*search\_button%3Dauthor+index%3Dau&buttons=subject%3Ddo\_authority\_search+\*search\_button%3Dsubject+index%3Dsu&buttons=keyword%3Ddo\_ccl\_search+\*search\_button%3Dkeyword+index%3Ddefault&buttons=callno%3Ddo\_authority\_search+\*search\_button%3Dcallno+index%3D%25&buttons=issn%3Ddo\_authority\_search+\*search\_button%3Dissn+index%3D%29&buttons=isbn%3Ddo\_authority\_search+\*search\_button%3Disbn+index%3D%28&buttons=oclc%3Ddo\_authority\_search+\*search\_button%3Doclc+index%3D%5E&buttons=lccn%3Ddo\_authority\_search+\*search\_button%3Dlccn+index%3D%23&buttons=gov%3Ddo\_authority\_search+\*search\_button%3Dgov+index%3D%26&buttons=t

ech%3Ddo\_authority\_search+\*search\_button%3Dtech+index%3D%26&buttons=music%3Ddo\_authority\_search+\*search\_button%3Dmusic+index%3D%26&button\_clicked=subject&Search=Search&material\_filter=all&language\_filter=all&date\_filter=all&servers=4home

2002-10-14 19:32:31 152.2.71.24 - 152.2.181.16 GET /web2/tramp2.exe 200 13879 1382 391 HTTP/1.1 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+4.0;+(R1+1.1))  
http://web2.lib.unc.edu/web2/tramp2.exe/form/A0ab2jhd.001?QUERY\_SCREEN=Home.html&QUERY=+clocks&buttons=title%3Ddo\_authority\_search+\*search\_button%3Dtitle+index%3Dti&buttons=author%3Ddo\_authority\_search+\*search\_button%3Dauthor+index%3Dau&buttons=subject%3Ddo\_authority\_search+\*search\_button%3Dsubject+index%3Dsu&buttons=keyword%3Ddo\_ccl\_search+\*search\_button%3Dkeyword+index%3Ddefault&buttons=callno%3Ddo\_authority\_search+\*search\_button%3Dcallno+index%3D%25&buttons=issn%3Ddo\_authority\_search+\*search\_button%3Dissn+index%3D%29&buttons=isbn%3Ddo\_authority\_search+\*search\_button%3Disbn+index%3D%28&buttons=oclc%3Ddo\_authority\_search+\*search\_button%3Doclc+index%3D%5E&buttons=lccn%3Ddo\_authority\_search+\*search\_button%3Dlccn+index%3D%23&buttons=gov%3Ddo\_authority\_search+\*search\_button%3Dgov+index%3D%26&buttons=tech%3Ddo\_authority\_search+\*search\_button%3Dtech+index%3D%26&buttons=music%3Ddo\_authority\_search+\*search\_button%3Dmusic+index%3D%26&button\_clicked=subject&Search=Search&material\_filter=all&language\_filter=all&date\_filter=all&servers=4home

2002-10-14 19:33:00 152.2.71.24 - 152.2.181.16 GET /web2/tramp2.exe 200 15059 1383 1265 HTTP/1.1 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+4.0;+(R1+1.1))  
http://web2.lib.unc.edu/web2/tramp2.exe/form/A0ab2jhd.001?QUERY\_SCREEN=Home.html&QUERY=+clocks&buttons=title%3Ddo\_authority\_search+\*search\_button%3Dtitle+index%3Dti&buttons=author%3Ddo\_authority\_search+\*search\_button%3Dauthor+index%3Dau&buttons=subject%3Ddo\_authority\_search+\*search\_button%3Dsubject+index%3Dsu&buttons=keyword%3Ddo\_ccl\_search+\*search\_button%3Dkeyword+index%3Ddefault&buttons=callno%3Ddo\_authority\_search+\*search\_button%3Dcallno+index%3D%25&buttons=issn%3Ddo\_authority\_search+\*search\_button%3Dissn+index%3D%29&buttons=isbn%3Ddo\_authority\_search+\*search\_button%3Disbn+index%3D%28&buttons=oclc%3Ddo\_authority\_search+\*search\_button%3Doclc+index%3D%5E&buttons=lccn%3Ddo\_authority\_search+\*search\_button%3Dlccn+index%3D%23&buttons=gov%3Ddo\_authority\_search+\*search\_button%3Dgov+index%3D%26&buttons=tech%3Ddo\_authority\_search+\*search\_button%3Dtech+index%3D%26&buttons=music%3Ddo\_authority\_search+\*search\_button%3Dmusic+index%3D%26&button\_clicked=subject&Search=Search&material\_filter=all&language\_filter=all&date\_filter=all&servers=4home

2002-10-14 19:33:14 152.2.71.24 - 152.2.181.16 GET /English/Graphics/catalog-links-vert.gif 304 142 576 0 HTTP/1.1 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+4.0;+(R1+1.1))  
http://web2.lib.unc.edu/web2/tramp2.exe/do\_authority\_search/A0ab2jhd.022?servers=4home&index=au&material\_filter=all&language\_filter=all&location\_filter=&location\_group\_filter=&date\_filter=all&query=Hearn%2C+George+A%2E%2C+Mrs%2E

- 1
- 2
- 3
- 4