Assignment 2: Subject Analysis

Group D

April 24, 2005

Information Retrieval

LIBR 202-13

USER GUIDE

INTRODUCING OUR DATABASE

Welcome to the Group D Library Literature Database! This database is designed for use by LIS students interested in professional literature in the domain of Information Search, Storage and Retrieval. Focus areas include systems design and end-user research. We hope that students can use this database to locate useful and relevant articles to assist in their own research. The twenty articles in the database were written between 1985 and 1999. Seventeen come from refereed academic journals, and three from *Scientific American*. All of the articles were selected from electronic versions available from different web-based databases.

DATABASE RULES

In constructing a controlled vocabulary for this database, our group crafted the following rules in the interest of having a clearer resource. To a certain extent it will serve to limit ambiguity in term usage for database searchers. Our hope is that this will decrease the extent to which our user base needs to rely on overly subjective decisions when formulating searches.

1) **In our pre-coordinate descriptor list, "Education" is used in all cases that include the concepts of "Education" and/or "Training".** We feel that Education is the broader of the two terms and that Training is typically a component of Education. We think that if we had both terms in our vocabulary lists, it would cause confusion amongst users and ambiguity in the database searches.
2) **"Internet" is used instead of "World Wide Web" or other terms or abbreviations which refer to the online world.** The Internet is the broader field and the World Wide Web is just one of its many technologies/tools/platforms.
3) **No proper names are used in our pre-coordinate and post-coordinate vocabularies, with a few key exceptions:** George Boole, Claude Shannon, Mortimer Taube. One of the articles required these names as descriptors to accurately summarize the article.

4) **Concepts and terms are stated in the plural form rather than in singular form.** We only used singular terms when the plural form was rarely used (if at all) in the English language. For example, "information seeking" instead of "information seekings".
5) **Our controlled vocabularies contain no "Free-Floating" terms.** Only terms that relate to the aboutness of the 20 articles in our database are included in the pre-coordinate and post-coordinate descriptor lists. **[I'm not sure what this means. The term "free-floating" is usually applied to subdivisions in a pre-co vocabulary that are used after many different main headings. Subdivisions such as *–Models* or *–Design* are often "free-floating."]**
6) **In general specific names of important theories, studies or companies in the Information Search and Retrieval field are not used.** The two exceptions are "Dialog" and "Personal Construct Theory", which are used as post-coordinate terms and are respective subsets of "Information Retrieval Systems" and "Cognitive Studies" in our pre-coordinate vocabulary.

TIPS TO MAKE YOUR SEARCH MORE EFFECTIVE

- We attempted to keep our controlled vocabulary terms broad and consistent in order to allow for good aggregation in searches through a structure that users should find logical and simple. Because of this, users with more specific search requirements may prefer to run an abstract search to help with discrimination in searches. The abstracts have been word-indexed to capture more detailed keywords through a natural-language search. For example, if you want to search for "Resnikoff-Dolby 30:1 Rule", use an abstract search, whereas our descriptor vocabulary will accommodate those looking for "theories" in general. When performing an abstract search with very narrow terms, we recommend using a Boolean "or" and truncation to improve recall.
- Our pre-coordinate terms exhibit between one and three levels of specificity in their terminology. The first (main) heading relates to broad areas within the Information Search and Retrieval domain that are either action-oriented in definition or components upon which most modern librarianship depends. The following headings (sub-heading 1 and 2) progressively increase the specificity of the main heading topic.

- Some of the articles in our database are classics of the field. While they serve as pillars for later research in the information retrieval field, the language used in the abstracts may be outdated. Consider using the broader terminology used in the post-coordinate and pre-coordinate vocabularies.
- Only perform title searches when seeking with very broad terminology.
- Given the relatively small size of our database, we do not recommend using the author or citation fields in searches. **[That depends more on the particular user's information need than on the size of the database. If someone wants articles by Marcia Bates or published in the Journal of the American Society for Information Science, then the author or the citation field would be the best way to search.]**

Good job with the User Guide. Giving some specific searches as examples in the "tips" section would make it even more effective.

DATA STRUCTURE

Field Summary

Field	Field Type	Indexing	Validation
Doc_No:	Automatic Number (next avail=21, increm=1)	Term	
AUTHOR:	Text	Term & Word	
TITLE:	Text	Term & Word	
CITE:	Text	Term & Word	
ABSTRACT:	Text	Word	
PRECO:	Text	Term	valid-list
POSTCO:	Text	Term	valid-list

Log file enabled, showing 'Doc_No'
Leading articles: a an the
Stop words: a an and by for from in of the to

VALIDATION LISTS

Term index for field 'PRECO', textbase 'GroupD_Assn2DB', 4/16/2005 6:32:18 PM:

1 Cognitive Studies Personal Construct Theory
1 Digital Information Information Organization Automation
1 Digital information Preservation
1 Indexing
1 Indexing Coordinate Indexing
1 Indexing Taxonomies
3 Information Retrieval
1 Information Retrieval Artificial Intelligence Systems
1 Information Retrieval Education Librarians
2 Information Retrieval Education Users
1 Information Retrieval Evaluation Measures
1 Information Retrieval Evaluation Retrieval Effectiveness
2 Information Retrieval History Early Applications
1 Information Retrieval History George Boole
1 Information Retrieval Limitations
1 Information Retrieval Models Artificial Intelligence
1 Information Retrieval Models Automated Systems
2 Information Retrieval Models Cognitive
2 Information Retrieval Models Heuristic
2 Information Retrieval Models Hypertext
1 Information Retrieval Models Online Systems
1 Information Retrieval Models Repertory grid modeling
1 Information Retrieval Natural Language Processing
1 Information Retrieval Natural Language Processing Full Text Search
1 Information Retrieval Natural Language Processing Text Comprehension
3 Information Retrieval Relevance Judgments
1 Information Retrieval Relevance Psychological
1 Information Retrieval Search Strategies
5 Information Retrieval Studies
1 Information Retrieval Systems Database Design
1 Information Retrieval Systems Dialog
1 Information Seeking Search Results Statistical Ranking
7 Information Seeking Search Strategies
1 Information Seeking Search Strategies Berrypicking
1 Information Seeking Search Strategies Boolean Queries
1 Information Seeking Search Strategies Neutral Questioning
1 Interface BRAQUE Browsing Query Formulation
6 Interface Design
2 Interface Design Human Computer Interaction
2 Interface User
1 Interfaces Hypertext
1 Internet Organization
1 Internet Searching
1 Internet Searching Tools
1 Software Emulators

Total number of keys: 45

Information Retrieval–Artificial Intelligence–Models would be more consistent with *Information Retrieval–Artificial Intelligence–Systems* above.

Good job with the preco. The numbers show a fairly high level of aggregation, at least for certain headings – higher if you were to use truncation. You rely a little too much on certain main headings, especially *Information retrieval*; some of these topics could have stood on their own without the main heading.

Term index for field 'POSTCO', textbase 'GroupD_Assn2DB', 4/16/2005 6:33:29 PM:

1 Abstracting
2 Artificial Intelligence
1 Automated Search Tools **[since you have a term *Automation,* I would split this into 2 or even 3 terms]**
3 Automation
1 Boolean Algebra
1 BRAQUE Browsing Query Formulation
1 Circuit Switching
2 Classification
1 Claude E Shannon
5 Cognitive Studies
1 Conceptual Cues
1 Coordinate Indexing
1 Database
2 Design
1 Dialog
2 Digital Archives
2 Digital information
1 Digital Resources **[Consider using *Digital information* for this (or vice versa)]**
1 Emulators
1 George Boole
1 Hardware
2 Heuristics
1 Hypertext
1 Hypertext Models **[split this into 2 terms: you already have *Hypertext,* and *Models* would be a useful term in conjunction with many topics]**
2 Indexing
1 Information Networks
15 Information Retrieval
9 Information Seeking
1 Information Technology
3 Internet
1 Librarians
1 Links
1 Logic
1 Mortimer Taube
3 Natural Language Processing
1 Neutral Questioning
1 Online Searching
1 Organization
1 Perceptual Cues
1 Personal Construct Theory
1 Preservation
1 Psychological Relevance

1 Reference Interview
3 Relevance
2 Relevance Judgments
1 Repertory Grid Modeling
2 Retrieval Effectiveness
1 Scanning
9 Search Strategies
1 Software
1 Statistical Ranking
6 Studies
1 Taxonomies
2 Text Comprehension
3 Training Methodologies **[I would split this into two terms; you might need *Training* alone, and *Methodologies* could be useful in other contexts.]**
10 User Interfaces

Total number of keys: 56

Generally good work with the postco vocabulary. There were only a few terms that appeared similar enough in meaning that they could be merged. With postco, be careful about terms that coordinate two concepts, especially when one or both concepts are already represented by separate terms. It all goes back to aggregation and discrimination. Someone interested in hypertext would not aggregate everything by searching *=hypertext*; it would be necessary either to use truncation (*=hypertext**) or to OR in the term *=hypertext models* in order to aggregate everything.

RECORD LIST

Doc_No 1
AUTHOR Dervin, Brenda
Dewdney, Patricia
TITLE Neutral Questioning: A New Approach to the Reference Interview
CITE RQ, p. 506-513 (Summer 1986)
ABSTRACT Neutral questioning is a strategy for conducting the reference interview in a way that allows the librarian to understand the query from the user's viewpoint. Neutral questions are open in form, avoid premature diagnosis of the problem, and structure the interview along dimensions important to the users. Derived from extensive studies into information-seeking behavior, this strategy has now been taught to over a thousand practitioners through workshops developed in 1981. Empirically based research to test the effects of neutral questioning is in progress. Informal evaluation indicates that neutral questioning may become a useful component of in-service training for librarians.
PRECO Information Seeking - Search Strategies -
Neutral Questioning
Information Retrieval - Education - Librarians
POSTCO Neutral Questioning
Librarians
Training Methodologies
Reference Interview
Information Seeking

"Neutral questioning" is a term that is likely unique to this article, and thus won't be useful in aggregation. Your term *Reference Interview* probably would suffice.

Doc_No 2
AUTHOR Hearst, Marti
TITLE Interfaces for Searching the Web
CITE Scientific American, vol. 276 no. 3, p. 68-72 (March 1997)
ABSTRACT The rapid growth of the World Wide Web is outpacing current attempts to search and organize it. New user interfaces may offer a better approach.
PRECO Internet - Searching
Internet - Organization
Interface - Design
Interface - User
POSTCO Internet
Information Retrieval
User Interfaces
Search Strategies
Organization

Doc_No	**3**
AUTHOR	Farrow, John
TITLE	A Cognitive Process Model of Document Indexing
CITE	Journal of Documentation, vol. 47 no. 2, p. 149-166 (June 1991)
ABSTRACT	Classification, indexing and abstracting can all be regarded as summarization of the content of a document. A model of text comprehension by indexers (including classifiers and abstractors) is presented, based on task descriptions which indicate that the comprehension of text for indexing differs from normal fluent reading in respect of: operational time constraints, which lead to text being scanned rapidly for perceptual cues to aid gist comprehension; comprehension being task oriented rather than learning oriented, and being followed immediately by the production of an abstract, index, or classification; and the automatically of processing of text by experienced indexers working within a restricted range of text types. The evidence for the interplay of perceptual and conceptual processing of text under conditions of rapid scanning is reviewed. The allocation of mental resources to text processing is discussed, and a cognitive process model of abstracting, indexing and classification is described.
PRECO	Indexing Information Retrieval - Models - Cognitive Information Retrieval - Natural Language Processing - Text Comprehension
POSTCO	Indexing Abstracting Scanning Text Comprehension Cognitive Studies Perceptual Cues Conceptual Cues

Doc_No	4
AUTHOR	Frisse, Mark Cousins, Steve
TITLE	Models for Hypertext
CITE	Journal of the American Society for Information Science, vol. 43 no. 2, p. 183-191 (March 1992)
ABSTRACT	Features characteristic of hypertext are no longer the providence of specialized hypertext systems. Interfaces exhibiting the "power of linking" can be found in software ranging from document preparation programs to operating systems, providing users with many new approaches to individual and group information management. To take advantage of the new opportunities afforded by this migration, it is important to understand hypertext at representation levels beneath the more superficial aspects of the human- computer interface. Three abstract models for hypertext--representative members of a spectrum of popular formalisms--are presented to provide insight into the meaning and potential of

	hypertext. Each model represents a different level in the design-decision process necessary for effective hypertext development, and each model plays an important role in development and use of information management software exhibiting hypertext features.
PRECO	Information Retrieval Information Retrieval - Models - Hypertext Interface - Design
POSTCO	Hypertext Models Information Retrieval User Interfaces Links

---

Doc_No	5
AUTHOR	Ury, Connie Jo Johnson, Carolyn Meldrem, Joyce
TITLE	Teaching a Heuristic Approach to Information Retrieval
CITE	Research Strategies, vol. 15 no. 1, p. 39-47 (Winter 1997)
ABSTRACT	To become life-long learners, students must acquire information retrieval skills for future as well as current information needs. This article describes how the Library Use Instruction Program at Northwest Missouri State University incorporates a heuristic model in which students continually evaluate and refine their information seeking practices while progressing through all levels of courses in diverse disciplines. Collegial partnerships with departmental faculty and ongoing instructional assessment are essential to the success of the program.
PRECO	Information Retrieval Information Retrieval - Models - Heuristic Information Seeking - Search Strategies Information Retrieval - Education - Users
POSTCO	Information Retrieval Training Methodologies Heuristics Information Seeking Search Strategies

---

Doc_No	6
AUTHOR	Lynch, Clifford
TITLE	Searching the Internet
CITE	Scientific American, vol. 276 no. 3, p. 52-56 (March 1997)
ABSTRACT	This paper discusses, some of the issues involved in the challenge of organizing vast quantities of potentially useful information from the Internet and the World Wide Web so that it is more accessible and useful to people who may seek it. It briefly discusses search engines, web crawlers, and other automated indexing and cataloging tools that help to bring order to a chaotic online world. “Combining

	the skills of the librarian and the computer scientist may help organize the anarchy of the Internet.
PRECO	Information Retrieval Digital Information - Information Organization - Automation Internet - Searching - Tools
POSTCO	Information Retrieval Digital information Internet Automation Information Seeking

---

Doc_No	7
AUTHOR	Harter, Stephen P.
TITLE	Psychological Relevance and Information Science
CITE	Journal of the American Society for Information Science, vol. 43 no. 9, p. 602-615 (October 1992)
ABSTRACT	This article summarizes the theory of psychological relevance proposed by Dan Sperber and Deirdre Wilson (1986), to explicate the relevance of speech utterances to hearers in everyday conversation. The theory is then interpreted as the concept of relevance in information retrieval, and an extended example is presented. Implications of psychological relevance for research in information retrieval; evaluation of information retrieval systems; and the concepts of information, information need, and the information-seeking process are explored. Connections of the theory to ideas in bibliometrics are also suggested.
PRECO	Information Retrieval - Evaluation - Measures Information Retrieval - Relevance - Judgments Information Retrieval - Relevance - Psychological Information Retrieval - Studies
POSTCO	Information Retrieval Relevance Psychological Relevance Information Seeking Studies Relevance Judgments

---

Doc_No	8
AUTHOR	Blair, David C. Maron, M. E.
TITLE	An Evaluation of Retrieval Effectiveness for a Full Text Document Retrieval System
CITE	Communications of the ACM, vol. 28 no. 3, p. 289-299 (March 1985)
ABSTRACT	An evaluation of a large, operational full- text document-retrieval system (containing roughly 350,000 pages of text) shows the system to be reentering less

	than 20 percent of the documents relevant to a particular search. The findings are discussed in terms of the theory- and practice of full-text document retrieval.
PRECO	Information Retrieval - Studies Information Retrieval - Relevance - Judgments Information Retrieval - Evaluation - Retrieval Effectiveness Information Retrieval - Natural Language Processing - Full-Text Search
POSTCO	Information Retrieval Automation Relevance Studies Natural Language Processing Retrieval Effectiveness

---

Doc_No	9
AUTHOR	Gauch, Susan
TITLE	Intelligent Information Retrieval: An Introduction
CITE	Journal of the American Society for Information Science, vol. 43 no. 2, p. 175-182 (March 1992).
ABSTRACT	Researchers are exploring the application of artificial intelligence techniques to information retrieval with the goal of providing intelligent access to online information. This article surveys several such systems to show what is possible in the lab today, and what may be possible in the library or office of tomorrow. Systems incorporating user modeling, natural language understanding, and expert systems technology are presented.
PRECO	Information Retrieval - Artificial Intelligence - Systems Information Retrieval - Natural Language Processing Interface - Design Information Seeking - Search Strategies
POSTCO	Information Retrieval Artificial Intelligence User Interfaces Information Seeking Automated Search Tools Natural Language Processing

---

Doc_No	10
AUTHOR	Marchionini, Gary
TITLE	Interfaces for End-User Information Seeking
CITE	Journal of the American Society for Information Science, vol. 43 no. 2, p. 156-163 (March. 1992)

ABSTRACT Essential features of interfaces to support end-user information seeking are discussed and illustrated. Examples of interfaces to support the following basic information- seeking functions are presented: problem definition, source selection, problem articulation, examination of results, and information extraction. It is argued that present interfaces focus on problem articulation and examination of results functions, and research and development are needed to support the problem definition and information extraction functions. General recommendations for research on interfaces to support end-user information seeking include: attention to multimedia information sources, development of interfaces that integrate information-seeking functions, support for collaborative information seeking, use of multiple input/output devices in parallel, integration of advanced information retrieval techniques in systems for end users, and development of adaptable interfaces to meet individual difference and multicultural needs.

PRECO Information Retrieval - Models - Hypertext
Information Retrieval - Models - Cognitive
Information Retrieval - Models - Online Systems
Information Seeking - Search Strategies
Interface - Design
Interface - Design - Human-Computer Interaction

POSTCO Information Retrieval
Information Seeking
Studies
User Interfaces

Doc_No 11
AUTHOR Bates, Marcia J.
TITLE Indexing and Access for Digital Libraries and the Internet: Human, Database and Domain Factors
CITE Journal of the American Society for Information Science, 49, p. 1185-1205 (1998)
ABSTRACT Discussion in the research community and among the general public regarding content indexing (especially subject indexing) and access to digital resources, especially on the Internet, has underutilized research on a variety of factors that are important in the design of such access mechanisms. Some of these factors and issues are reviewed and implications drawn for information system design in the era of electronic access. Specifically the following are discussed: Human factors: Subject searching vs. indexing, multiple terms of access, folk classification, basic-level terms, and folk access; Database factors: Bradford's Law, vocabulary scalability, the Resnikoff-Dolby 30:1 Rule; Domain factors: Role of domain in indexing.

PRECO Information Retrieval - Models - Automated Systems
Interface - Design
Indexing - Taxonomies
Information Seeking - Search Strategies

POSTCO
Information Retrieval
Digital Resources
Internet
Automation
Indexing
User Interfaces
Classification
Information Seeking
Taxonomies

Doc_No 12
AUTHOR Smith, Elizabeth S.
TITLE On the Shoulders of Giants: From Boole to Shannon to Taube: The Origins and Development of Computerized Information from the Mid-19th Century to the Present
CITE Information Technology and Libraries, vol.12 no. 2, p. 217-226 (June 1993) [electronic version]
ABSTRACT This article describes the evolvement of computerized information storage and retrieval, from its beginnings in the theoretical works on logic by George Boole in the mid-nineteenth century, to the application of Boole's logic to switching circuits by Claude Shannon in the late 1930's, and the development of coordinate indexing by Mortimer Taube in the late 1940's and early 1950's. Thus, electronic storage and retrieval of information, as we know it today, was the result of two major achievements: the advancement of computer technology initiated to a large extent by the work of Shannon, and the development of coordinate indexing and retrieval by the work of Taube. Both these achievements are based on and are the application of the theoretical works of George Boole.
PRECO Information Retrieval - History - George Boole
Information Retrieval - History - Early Applications
Information Seeking - Search Strategies - Boolean Queries
Indexing - Coordinate Indexing

Why not headings for Shannon and Taube as you have in your postco?

POSTCO
George Boole
Information Retrieval
Logic
Boolean Algebra
Claude E. Shannon
Circuit Switching
Mortimer Taube
Coordinate Indexing
Search Strategies

Doc_No	13
AUTHOR	Belkin N. J. Marchetti, P. G. Cool, C.
TITLE	BRAQUE: Design of an Interface to Support User Interaction in Information Retrieval
CITE	Information Processing and Management, 29, p. 325-344 (1993)
ABSTRACT	We suggest that information retrieval is most appropriately considered as an inherently interactive process, and describe the design of an interface to a bibliographic information retrieval system that supports user interaction in an integrated fashion. A significant aspect of the interface design is its dependence upon a two-level hypertext model of information retrieval system databases, and the seamless support of a variety of information seeking strategies through the use of this model, and of a model of dimensions of information seeking behaviors.
PRECO	Interfaces - Hypertext Interface - BRAQUE (Browsing & Query Formulation) Information Seeking - Search Strategies
POSTCO	Information Seeking Search Strategies User Interfaces Information Retrieval BRAQUE (Browsing & Query Formulation) Cognitive Studies Hypertext Digital Archives

Doc_No	14
AUTHOR	Latta, Gail F. Swigger, Keith
TITLE	Validation of the Repertory Grid for Use in Modeling Knowledge
CITE	Journal of the American Society for Information Science, 43, p. 115-129 (1992)
ABSTRACT	The extrapolation of theories of cognitive modeling from psychological contexts to the domain of information systems design has raised empirical questions regarding the validity of reliability of these techniques in the new applications environment. Kelly's (1955) theory of personal construct psychology (PCP) is one such theory which, together with the repertory grid technique of cognitive modeling, is being explored for use in systems development. Since the repertory grid constitutes a mechanism for both the elicitation and the representation of cognitive models, it has been suggested for incorporation into intelligent front-end interfaces for information storage and retrieval systems (ISRS). Previous applications of the grid have emphasized the personal and ideographic characteristics of personal construct systems. New applications in information systems design require verification of the grid's utility and reliability for rendering faithful representations of impersonal and communal knowledge. The

research present here provides a test of the repertory grid for representing commonality of construing among members of a homogenous group of subjects. Commonality of construing is demonstrated on two sets of constructs generated by the subjects and those provided for them by an "expert." Implications of this effect for the application of PCP to ISRS design are discussed.

PRECO Information Retrieval - Models - Artificial Intelligence
Information Retrieval - Models - Repertory Grid Modeling
Information Retrieval - Studies
Cognitive Studies - Personal Construct Theory

POSTCO Artificial Intelligence
Cognitive Studies
Information Retrieval
User Interfaces
Personal Construct Theory
Repertory Grid Modeling
Search Strategies

Doc_No 15
AUTHOR Huston, Mary M.
TITLE Windows into the search process: an inquiry into dimensions of online information retrieval
CITE Online Review, 15, p. 227-243 (1991)
ABSTRACT From diverse users' points of view, contextual frameworks are elaborated for the nature of the information technology, the information universe, and the information search. Within these conceptual parameters, established theories on search strategy are reviewed and cognitive models of information- seeking are highlighted. Future directions for research on users' search processes are discussed in terms of the role for online retrieval in the future information environment.

PRECO Interface - User
Information Seeking - Search Strategies
Information Retrieval Systems - Dialog

POSTCO Online Searching
User Interfaces
Search Strategies
Information Technology
Information Retrieval
Information Networks
Studies
Information Seeking
Dialog
Cognitive Studies

Doc_No	16
AUTHOR	Rothenberg, Jeff
TITLE	Ensuring the Longevity of Digital Information
CITE	Scientific American, vol. 272 no. 1, p. 42-47 (January 1995) [electronic version]
ABSTRACT	Digital documents are replacing paper in the most dramatic record-keeping revolution since the invention of printing. Unlike texts printed on the paper, digital files cold not be read without adequate technology. The article explains basic facts about digital bit streams and problems of interpreting raw data without adequate hardware or software, and principles of bootstrapping and files compression. In order to be preserved digital records must be continuously migrated to the newer media carriers, what presents substantial logistic and financial costs. Article analyzes some possible solutions, among others the software emulators programs.
PRECO	Digital information - Preservation Software - Emulators
POSTCO	Digital Information Digital Archives Preservation Software Hardware Emulators

Doc_No	17
AUTHOR	Bates, Marcia J.
TITLE	Design of browsing and berrypicking techniques for the online search interface
CITE	Online Review, vol. 13 no. 5, p. 407-424 (1989)
ABSTRACT	First, a new model of searching in online and other information systems, called "berry picking", is discussed. This model, it is argued, is much closer to the real behavior of information searchers than the traditional model of information retrieval is, and, consequently, will guide our thinking better in the design of effective interfaces. Second, the research literature of manual information seeking behavior is drawn on the new model and research on information seeking, suggestions are made for how new search capabilities could be incorporated into the design of search interfaces. Particular attention is given to the nature and types of browsing that can be facilitated.
PRECO	Information Seeking - Search Strategies Information Seeking - Search Strategies - Berrypicking Interface - Design - Human-Computer Interaction
POSTCO	User Interfaces Database Design Information Retrieval Search Strategies

"Berrypicking" is like "neutral questioning"; I would use *Browsing*, or leave it with just *Information Seeking–Search Strategies*.

Doc_No	18
AUTHOR	Swanson, Don R.
TITLE	Historical note: Information retrieval and the future of an illusion
CITE	Readings in information retrieval, p. 555-561 (1997)
ABSTRACT	In this article author offer a personal perspective on automatic indexing and information retrieval, focusing on events and ideas over a 34-year period that have led to the view that information retrieval involved conceptual problems of greater subtlety than is generally recognized. Some experimental tests of information systems have yielded good retrieval results and some very poor results. Author thinks that poor results merit special attention and why we should reconsider a suggestion that Robert Fairthorne put forward in 1963 about development of the postulates of impotence – statements of what cannot be done. By understanding such limits we are led to new goals, metaphors, problems, postulates, and perspectives.
PRECO	Information Retrieval - Relevance - Judgments Information Retrieval - History - Early Applications Information Retrieval - Limitations
POSTCO	Information Retrieval Relevance Judgments Search Strategies Classification Retrieval Effectiveness Relevance Text Comprehension

Doc_No	19
AUTHOR	Herman, Donna
TITLE	User-friendly systems instead of user-friendly front-ends
CITE	Readings in information retrieval, vol. 43 no. 2, 413 - 423 (March 1992)
ABSTRACT	Most commercial online retrieval systems are not designed to service end users and, therefore, have often built "front-ends" to their systems specifically to serve the end- user market. These front-ends have not been well accepted, mostly because the underlying systems are still difficult for end users to use successfully in searching. New techniques, based on statistical methods, that allow natural language input and return lists of records in order of likely relevance, have long been available from research laboratories. This article presents four prototype implementations of these statistical retrieval systems that demonstrate their potential as powerful and easily used retrieval systems able to service all users.
PRECO	Information Retrieval Systems - Database Design Information Seeking - Search Results - Statistical Ranking Information Retrieval - Studies Interface - Design
POSTCO	Statistical Ranking

User Interfaces
Studies
Natural Language Processing
Design

Doc_No	20
AUTHOR	Borgman, Hristine L. **[OCR error? – it's Christine Borgman]**
TITLE	User's mental model of an information retrieval system: an experiment on a prototype online catalog
CITE	International Journal of Human-Computer Studies, vol. 51 no. 2, p. 435-452 (August 1999)
ABSTRACT	An empirical study was performed to train naïve subjects in the use of a prototype Boolean logic-based information retrieval system on a database of bibliographic records. The research was based on the mental models theory which proposes that people can be trained to develop a "mental model" or a qualitative simulation of a system which will aid in generating methods for interacting with the system, debugging errors, and keeping track of one's place in the system. As predicted, the model-based training had no effect on the ability to perform simple, procedural tasks, but subjects trained with a model performed better on complex tasks that required extrapolation from the basic operations of the system. A stochastic process analysis of search-state transitions reinforced this conclusion. Subjects had difficulty articulating a model of the system, and we found no differences in articulation by condition. More interestingly, those who dropped out were significantly more likely to be humanities or social science majors than science or engineering majors, suggesting important individual differences and equity issues. The sex-related differences ere slight, although significant, and suggest future research questions.
PRECO	Information Retrieval - Education - Users Information Retrieval - Studies Information Retrieval - Models - Heuristic Information Retrieval - Search Strategies
POSTCO	Training Methodologies Studies Search Strategies Heuristics Cognitive Studies

Ideally, your preco and postco terms would cover the same concepts. While yours overlap pretty well, sometimes a concept is brought out in one vocabulary but not the other.

DATABASE EVALUATION

Measurement

Query 1

Information need: I am interested in documents that discuss how interface design can aid in the information retrieval process.

Relevant documents: 2, 4, 6, 9, 10, 13, 15, 17, 19

Field	Query	Docs Retrieved	Recall	Precision	Effectiveness
Title	interface* / retrieval	2, 5, 8, 9, 10, 13, 15, 17, 18, 20	0.67	0.60	0.63
Abstract	(interface* / information retrieval) & design*	4, 13, 14, 17	0.33	0.75	0.49
Post-co	=user interfaces & =information retrieval	2, 4, 9, 10, 11, 13, 14, 15, 17	0.78	0.78	0.78
Pre-co	=interface design	2, 4, 9, 10, 11, 19	0.56	0.83	0.67

	Title	Abstract	Pre-co	Post-co
Relevant doc [A]	6	3	5	7
Doc retrieved [B]	10	4	6	9
All relevant doc [C]	9	9	9	9
Recall	**0.67**	**0.33**	**0.56**	**0.78**
Precision	**0.60**	**0.75**	**0.83**	**0.78**

Since interface seems to be the most important component of your information need here, I'm surprised you would OR it with an unrelated term in your title and abstract searches. (More commonly, OR is used between synonyms: interface* / interaction, for example.) I would expect poor precision as a result, but surprisingly your results are still pretty good.

Query 2

Information need: Approaches to training of users and librarians to handle information seeking and retrieval process.

Relevant documents: 1, 5, 20

With recall and precision both 0, effectiveness would also be 0.

Field	Query	Docs Retrieved	Recall	Precision	Effectiveness
Title	(user* & education*) /libr*	11	0	0	undefined
Abstract	(train* & user*) / libr*	1, 5, 6, 9	0.67	0.50	0.58
Post-co	=Training Methodologies & =Information Retrieval	5	0.33	1	0.53
Pre-co	=Information Retrieval Education*	1, 5, 20	1.00	1.00	1.00

	Title	Abstract	Pre-co	Post-co
Relevant doc [A]	0	2	3	1
Doc retrieved [B]	1	4	3	1
All relevant doc [C]	3	3	3	3
Recall	**0.00**	**0.67**	**1.00**	**0.33**
Precision	**0.00**	**0.50**	**1.00**	**1.00**

In a larger database, ORing a term like libr* would be a bad idea, since any title or abstract containing the word *library*, *libraries*, *librarian*, etc. would be retrieved, even if none of your other terms were present. With a very small database such as this one, results are harder to predict. A query that fits your stated need better might be something like:

(user / libr*) & (education* / train*)*

This wouldn't have retrieved any documents in a title search (I'm not sure about the abstract), but none of the titles of the documents you selected as relevant show a clear relationship to your information need.

Query 3

Information need: I want to learn more about how cognitive research studies have benefited the field of information retrieval.

Relevant documents: 1, 3, 7, 14, 15, 20

Field	Query	Docs Retrieved	Recall	Precision	Effectiveness
Title	psycholog* / cognitive*	3, 7	0.33	1.00	0.53
Abstract	(cognitive* / psycholog*) & information & retrieval	7, 14, 15	0.50	1.00	0.65
Post-co	=Cognitive Studies / =Training Methodologies	1, 3, 5, 13, 14, 15, 20	0.83	0.71	0.76
Pre-co	=Information Retrieval Models Cognitive / =Information Retrieval Education*	1, 3, 5, 10, 20	0.50	0.60	0.55

	Title	Abstract	Pre-co	Post-co
Relevant doc [A]	2	3	3	5
Doc retrieved [B]	2	3	5	7
All relevant doc [C]	6	6	6	6
Recall	**0.33**	**0.50**	**0.50**	**0.83**
Precision	**1.00**	**1.00**	**0.60**	**0.71**

Query 4

Information need: I am interested in documents which discuss how artificial intelligence can aid in information search and retrieval.

Relevant documents: 6, 9, 14, 16

Field	Query	Docs Retrieved	Recall	Precision	Effectiveness
Title	digital* / modeling* / *intelligence*	11, 14, 16	0.50	0.67	0.58
Abstract	(information retrieval / search*) & (intelligence* / modeling* / automation / digital*)	9, 11	0.25	0.50	0.36
Post-co	= information retrieval & = artificial intelligence	9, 14	0.50	1.00	0.65
Pre-co	=information retrieval *artificial Intelligence*	9	0.50	1.00	0.65

	Title	Abstract	Pre-co	Post-co
Relevant doc [A]	2	1	1	2
Doc retrieved [B]	3	2	1	2
All relevant doc [C]	4	4	4	4
Recall	**0.50**	**0.25**	**0.25**	**0.50**
Precision	**0.67**	**0.50**	**1.00**	**1.00**

Your measurements are very clearly presented and correctly calculated.

158 points (out of 170) for Parts A & B