Models for Reader Interaction Systems

Daniel Berleant
Dept. of Electrical and Computer Engineering
Iowa State University
Ames, IA 50011
(515) 294-3959

berleant@iastate.edu

ABSTRACT

In this paper we discuss models for systems that support reading. Our account identifies important models and presents a framework for organizing them. To evaluate this account, we show its ability to suggest a wide range of text presentation systems, many of them novel. This evaluation not only provides interesting ideas for future systems, it also shows the usefulness of the account, and further, exemplifies a general approach to evaluating meta-level discussions such as this, namely, evaluating by assessing ability to generate interesting implications.

Keywords

Reading, text, navigation, foraging, browsing, browsers, metaphors, paradigms, models.

1. INTRODUCTION

A variety of models, paradigms, metaphors, and componentizations exist that relate to reader interaction with text. We define *paradigms* here as broad and overarching computing themes that can contribute to the design of systems for reading but do not by themselves imply much about the specifications of such systems. Paradigms of this type include parallel computing, information retrieval and the contrasting paradigm of information foraging, and the spatial paradigm of information "navigation" (Waterworth and Chignell 1991 [40]; Dourish and Chalmers 1994 [10]). In comparison to paradigms, *models* provide significant guidance in system specification. By models for reader interaction with texts, we refer to models such as the hypermedia and (e-)book models that characterize different approaches to the specification of systems for helping humans to read text. *Metaphors* may be either paradigms (e.g. the spatial metaphor) or models (e.g. the book metaphor), though paradigms and models are not necessarily based on metaphors. By *componentizations*, we refer to reductionistic analyses that distinguish the stages of operation of such systems or their functional parts (e.g. Ellis 1989 [12], summarized in [40]; Shneiderman et al. 1998 [35]).

In this paper we build on previous work in this area, which might usefully be distinguished, and named *epistemology of information*

exploration. We review existing models and provide a framework for organizing them (Sections 2 & 3). We then evaluate by showing the ability of the framework and the models to support the generation of interesting ideas for text interaction systems (Section 4).

2. A FRAMEWORK FOR MODELS OF TEXT INTERACTION

In organizing a set of items such as models, one approach is to dichotomize – split a set into two subsets. For example, Chalmers (1999 [7]) characterizes his subsets as *positivist* (based on the objective content of the information) and *hermeneutic* (based on the evolving meaning of the information as interpreted by its users).

Additional organization may be imposed by then splitting the set again orthogonally. Two orthogonal splits result in four subsets, three result in eight subsets, etc. Waterworth and Chignell (1991 [40]) and Charoenkitkarn et al. (1994 [6]) do this in analyzing approaches to information exploration. An alternative to orthogonal splits is to employ a hierarchical taxonomy. This is more flexible because it allows a given subset to be split using a criterion that may not apply to another subset. In either case, splits that produce mutually exclusive subsets are better than splits with fuzzy boundaries. Fuzzy boundaries can occur due to either overlap (an item could be in both subsets) or an attempt to impose a dichotomy on what is really a continuum (an item could fall in a gray area in the middle of the continuum).

We organize models of reader interaction with text by exploiting the flexibility of a hierarchical taxonomy. In doing so, we incorporate the positivist vs. hermeneutic distinction of Chalmers (1999 [7]), the key string vs. browsing distinction analyzed by Waterworth and Chignell (1991 [40]) and Charoenkitkarn et al. (1994 [6]), and the prescriptive vs. non-prescriptive distinction we make in Berghel et al. (1999 [2]) which captures the notion that movement within an information space can be determined largely by the author (prescriptively) or, more flexibly, by the user (non-prescriptively). We do not rely fully on any one of those works because of the need for more than the one distinction made by Chalmers [7], the need in the present context to minimize consideration of the user's state of mind (explored by Waterworth and Chignell [40] and Charoenkitkarn et al. [6]) because that can vary from user to user regardless of the system or system model under consideration, and the concern here with interacting with information rather than Berghel et al.'s emphasis on customizing it [2]. Figure 1 shows the resulting taxonomy.

Models for Reader Interaction Systems

```
              Models for Reader Interaction Systems
                              /\
                             /  \
                            /    \
               Result oriented,    Process oriented,
               also Prescriptive,   also Non-prescriptive
               also Positivist
                  |                    |
                  |                    |
          Key string   Browsing    Key string    Browsing
          based        based       based, also   based, also
                                   Positivistic  Hermeneutic
              /\                      /
         Bounded  Unbounded          /

   Index,      Book,       Hypermedia,  Text       Composition
   Directory   Newspaper   Citation,    mining,    model
   models      models      Reuse        Query
                           models       models
```
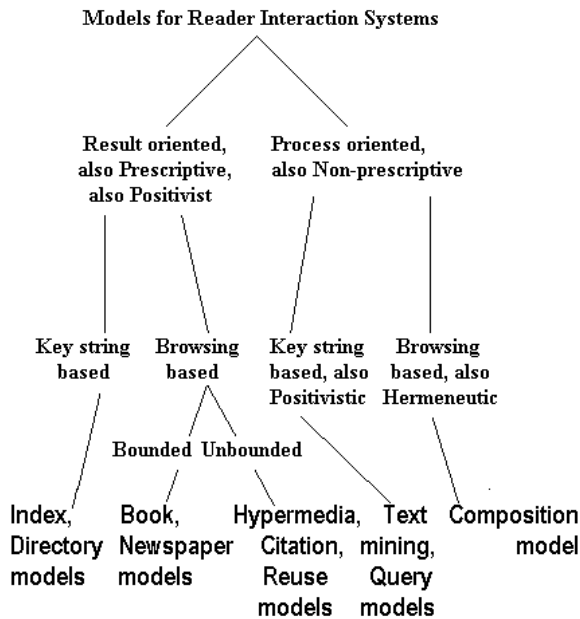
Figure 1. Framework for models of reader interaction systems. The models themselves appear in the leaf nodes. The index model refers to indexes as a focus for the system to support interaction. The directory model refers to non-alphabetized collections of pointers to locations in text, such as tables of contents, large Web directories, and small Web jump tables. The newspaper model (Kamba et al. 1993 [20]; Golovchinsky and Chignell 1997 [15]) emphasizes simultaneous presentation of different threads and an overarching organization that presents more important things ahead of less important things. The citation model refers to connections to remote information as a key element. The reuse model refers to such constructs as quotations, Web mirrors, and Nelson's transclusion [31] (which involves presenting passages in remotely stored documents as part of a given document without actually copying those passages). The composition model refers to interactive ways for users to construct composite text presentations distilled from larger quantities of text, including such primitive functions as Web browsers' history and bookmark lists.

3. THE MODELS IN MORE DETAIL

In this section the models are reviewed individually. Then in section 4 the models are employed as a foundation for focusing on mixed-model text interaction systems.

3.1 The Hypermedia Model

Before the Web, hypertext was usually applied to circumscribed bodies of related material, although from the beginning it was envisioned as a way to organize very large collections of information (Bush 1945 [5]; Nelson 1995 [31]). Hypertext tends to be author intensive since the author must create not only the text but the links as well, although automatically generated links are a continuing topic of investigation. The relatively static nature of hyperlinks (Halasz 1988 [16]) can be a disadvantage when the user's information-seeking goals do not match the author's information-presentation goals (Berghel et al. 1999 [2]).

Hypertext has limited capability to satisfy information exploration needs in large collections of information (like the Web), as the typical rather unstructured networks of texts and links make it easy for users to get lost – motivating such useful facilities as directories and the bot-generated index services called search engines. Differentiation between different types of links is typically via the descriptive labels that the user clicks.

3.2 The Book Model

The book is one of the oldest and most mature information presentation techniques. Aspects of books that contribute to its usefulness include tables of contents, indexes, sequential page numbers, and chapters. The book concept motivated such earlier electronic browsing systems as SuperBook (Egan et al. 1989 [11]) and the Book Emulator (Benest 1990 [1]), as well as recent commercial offerings of hand held devices of book size, even including a book-shaped device that opens into two displays (like two pages) around a hinged "spine" [13]. Excessive adherence to the book metaphor can guide system development away from the ways that computers could improve on books. The dual display screen device just mentioned is one example of this; another example is the emphasis in the ebook community on hardware design issues, such as electronic paper – actual physical film analogous to paper but electronically driven so it can be used as a display device [29].

3.3 The Newspaper Model

Newspapers present text in a distinctive style. The most important information is presented first, in that the most important articles are presented first, and in a particular article the most important material is presented first. Newspapers also present a number of discourse threads simultaneously; a page contains multiple articles. Thus, the newspaper embodies a good presentation strategy for busy readers with short attention spans. The newspaper model as a way to present information electronically is explored by Kamba et al. (1995 [20]) and Golovchinsky and Chignell (1997 [15]).

3.4 The Citation Model

The citation has long played a central role in the scholarly world as the mechanism for formally relying on existing foundations when building new material. On the Web, hyperlinks are the newly supported mechanism for referencing other work. Links to other documents can be regarded as citations in cyberspace, though the quality of such citations is uncontrolled [37][38] due to both the relative impermanence of Web content (links can break, and document content can change, unlike for paper documents) and because linking is affected by factors less frequently found in scholarly works (commercial competitors are relatively unlikely to link to each other even when this would be helpful to others, for example). Nevertheless, the hyperlink mechanism has caused citations to become vastly more widespread than before, when they were mostly limited to scholarly works. Even in scholarly works, hyperlinking from a paper's reference list to the papers it cites is radically more convenient than a trip to the bookshelf or library, suggesting that influential digital libraries such as those of ACM and IEEE would benefit users by providing papers, or at least their reference lists, in HTML format. Web technologies such as archiving the Web

(Kahle 1997 [19]), Web based document version retention (Simonson et al. 1998 [38] provides a survey; see also www.webdav.org), URNs (Sollin and Masinter 1994 [39]) and PURLs (Permanent URLs, see www.purl.org) take on additional significance when viewed as ways to increase the reliability of citations implemented as links in cyberspace.

3.5 The Reuse Model
New work can be strengthened by reusing (as contrasted with citing) previously developed material. Its value is apparent in such applications as quotes, and in the wide use of image inclusion in Web documents (as via the HTML `<img src=...>` command). The reuse model underlies Nelson's work on transclusion [30][31]. Transclusion is (for our purposes here) the process of dynamically importing into a document material from other, often remote documents. The material is typically some interior portion of a remote document, and is used as an interior portion of the importing document. If the material changes due to updates to the remote document, the changes are reflected in the importing document since the importing is done dynamically.

3.6 The Composition Model
By composition we refer to the incremental creation of a composite presentation from separately useful parts (Golovchinsky and Chignell 1997 [15]). Halasz's "browsers" and "fileboxes" (1988 [16]) were early examples in which the user interactively builds up a representation that abstractly describes a collection of passages. Other examples include Elastic Windows (Kandogan and Schneiderman 1997 [21]), sticky portals in Pad++ [36], the Krakatoa interactive electronic newspaper system (Kamba et al. 1997 [20]), and the CollageMachine (Kerne 1997 [22]). Common and basic applications of the composition model include Web browser bookmark lists and browsers' maintenance of a per-session history list, which may be traversed by back and forward buttons.

3.7 The Query Model
Queries are dynamic and defined by the user, in contrast to ordinary hyperlinks, which are static and defined by the author. Queries can retrieve multiple results, in contrast to ordinary hyperlinks which when followed lead to a single result. Yet hyperlinks have convenient point-and-click access whereas queries are typically typed. Queries need not be of the traditional information retrieval kind; ones based on descriptions of node-and-link structures in a hypermedia network are an alternative (Halasz 1988 [16]) with no obvious hyperlink counterpart. There are significant similarities between queries and links – Golovchinsky and Chignell (1993 [14]) title a paper "Queries-R-Links...". However the differences we have pointed out are significant.

Although the hyperlink model helped produce the Web, and information access on the Web was initially mainly by following hyperlinks around, queries (via the search engines) have since become an essential part of it. The Web without these query based search engine would be an impoverished environment in comparison.

3.7.1 Question Answering
A question is a kind of query, distinguished by the complexity of the process of determining what it means and how to respond. The best response is not necessarily stored text; it might be a meta-response intended to elicit further input from the user to in turn enable better retrieval of stored text. Question answering systems include FAQ finder (Burke et al. 1997 [4]) and closely related systems (e.g. Kulyukin et al. 1998 [23]), MURAX (Kupiec 1993 [24]), Pilkington (1992 [33]), and Question Master [34].

3.8 The Text Mining Model
Halasz (1987:358-360 [16]) early noted the importance of mining in hypermedia networks. The Web search engines use ranking algorithms to order the lists of links they return - based on an index derived by extensive mining of the Web. Other information can also be mined. For example google.com counts the incoming links to a page and uses that in its ranking algorithm. www.excite.com interactively suggests new terms that the user might add to the query to make it more specific, where those new terms are determined automatically by mining the documents returned by the original query for co-occurrences. A non-search engine example is the capability of amazon.com to inform the user, given a book of interest, what other books were sold most often to people buying a given book. Web robots that search the Web for useful items conforming to some user interest are other examples of text mining.

3.9 The Directory Model
Encyclopedias and other reference works rely on the directory model. Many Web search engine sites now provide directories. Yahoo provides an interesting combination of the directory and query models that allows users to query after moving down some distance into a directory tree. Directories have the attractive property that the "drilling" process is logarithmic in the total number of items cataloged. However they have the less attractive property of imposing a taxonomy that may not be suited to a given information seeking need.

4. HYBRIDIZING THE MODELS LEADS TO HYBRID SYSTEM CONCEPTS
One way to try to generate new system designs is to ask how two (or more) of the models we have discussed could be combined in a single text interaction system. Many such combinations lead to hybrid systems that have already been described in the literature. Others are novel. Guidance in combining them is implicit in Figure 1: models sharing the same leaf node in the taxonomy (e.g. the book and newspaper models) have basic underlying similarities, while entries in different leaf nodes have basic differences that make hybridizing them an interesting exercise which results in the outline of a system that draws on both models. Numerous pairs of entries from different leaf nodes are possible; inspection reveals 39 of them. More complex hybrids of three or more entries might also be discussed, but due to space considerations we limit the following list to brief discussion of the 39 pair hybrids.

The book model and the...

(1)...index model: e-books can easily have automatically generated indexes that are more extensive than is usual in printed books; such indexes should have entries that, when clicked, land the user at the corresponding location in the e-book.

(2)...directory model: having a clickable table of contents for an e-book is obvious. Automatic outline generation to facilitate rapid

browsing is almost as obvious. A hierarchical taxonomy tree with leaves that, when clicked, lead to locations in the e-book is somewhat less obvious.

(3)...hypermedia model: long HTML documents could benefit from some book-like features, such as an index at the end of the document, and page numbers.

(4)...citation model: book-length documents on the Web could be automatically populated with links to a major search engine. Each such link would contain search terms such that clicking on it would bring up the response of the search engine to the search terms. Quoted phrases from the document could be used as the search terms; such phrases could be automatically adjusted in length to retrieve, say, a list of at least 1 but no more than 10 URLs.

(5)...reuse model: all documents on the Web dealing with a particular narrow topic could be assembled into an on-line book-length compendium, with a common table of contents (index, etc.). The narrowness of the topic could be adjusted so that the result is, in fact, book length.

(6)...text mining model: automatic index creation for an on-line book could take into account the Web background frequency of the indexed terms, comparing those frequencies to the frequencies of terms in the book. Terms with significantly higher frequencies in the book are likely to be more important to index. Another combination of text mining and the book model would be to automatically generate extracts or abstracts [2][25][27][28] of an on-line book.

(7)...query model: an on-line book could be divided into separate paragraph-sized chunks, each in its own file. The resultant set of files could be indexed by a site-specific search utility, allowing users to browse the paragraphs of the book similarly to the way they browse documents on the Web. In fact, if each paragraph was tagged with the same special string, found nowhere else on the Web, then the set of files could be submitted to a standard search engine for indexing and, as long as queries AND that special string into the query, searching and browsing on the Web would automatically be constrained to that particular book.

(8)...composition model: as in 5), a system could create a book-length compendium of Web documents. The difference here is that the composition model suggests that the user have interactive input on which Web documents are included.

The newspaper model and the...

(9)...index model: an ordinary Web search engine, which relies on queries to an index, could be automatically queried using index terms of the user's choice. This is done periodically, perhaps every night, and any documents which are new or have changed since the previous access by the system are collected into a custom news package and emailed to the user as an on-line custom news service. While Web change monitoring has been described before (e.g. [9][32]) this service augments that by basing it on an ordinary Web search engine together with index terms characterizing what the user wants to receive news about.

(10)...directory model: like 9) but with the news service based on changes to documents descending from a particular node in a directory [32].

(11)...hypermedia model: online news sites such as www.cnn.com fall into this category. The news stories have links to maps, pictures, related stories from the past, and so on.

(12)...citation model: the popular press often reports on interesting scientific developments, but rarely provides definite citations and often doesn't provide even rudimentary ones. It would be an improvement if better citations were provided.

(13)...reuse model: Terms and phrases in on-line news articles could link to entries in an on-line encyclopedia or other reference resource.

(14)...text mining model: a bot running continuously in the background could travel the Web searching for documents which both satisfy some user-specific profile and have recently changed. The URLs found, or better, the changed passages within them, are then emailed to the user.

(15)...query model: see 9).

(16)...composition model: readers of online newspapers could interactively control newspaper layout and content (Kamba et al. 1993 [20]; Golovchinsky and Chignell 1997 [15]).

The index model and the...

(17)...hypermedia model: Web search engines use indexes, but unlike traditional indexes they do not show the user the actual alphabetized list of terms in the index. Sometimes, users might like to access parts of this list, such as terms in it that are in the neighborhood of terms in the user's query.

(18)...citation model: an alphabetical list of references at the end of a paper is almost an index, but unlike an index it typically lacks pointers back into the text. However, reference lists certainly could have such pointers. This would often make them more useful. Furthermore, if a reference points back to more than one place in the text, this suggests it is particularly important in the work. To illustrate these points, the reference list of this paper implements this combination.

(19)...reuse model: citations typically reference an entire work, not the portion of it that motivated the citation. It would be helpful if that actual portion was specified in the citation. In a hypertext style environment (though not, in general, the Web), links could point to the start and end points of the cited portion. In Nelson's work this corresponds to the concept of transclusion [30][31].

(20)...text mining model: see 17).

(21)...query model: see 17).

(22)...composition model: an electronically viewed index could have any given entry interactively expanded with a secondary index of terms it co-occurs with. Entries in such a secondary index could likewise be interactively expanded to the third level, and so on. Thus the index is interactively expanded consistently with the user's current needs.

The directory model and the...

(23)...hypermedia model: existing hierarchical Web directories like Yahoo exemplify this combination. Graphical "maps" of the network of links in a hypermedia environment are visually oriented, directory-like structures; numerous works have explored this on the Web environment.

(24)...citation model: this combination is exemplified by Science Citation Index and Social Science Citation Index [18], and by ResearchIndex (Lawrence et al. 1999 [26]) for the computing field, which tell what works cite a given work.

(25)...reuse model: a book of quotations [8] is such a combination.

(26)...text mining model: using text mining to develop a taxonomy for a Web directory automatically, based on term occurrence and co-occurrence frequencies would combine the directory and text mining models.

(27)...query model: a boolean query can be turned into a hierarchical directory. The top node consists of the OR of the query terms and the bottom (most specific) node consists of their AND. Each path through the directory encounters progressively more specific categories. This directory would be a DAG rather than a tree. A simple example is shown in Figure 2.
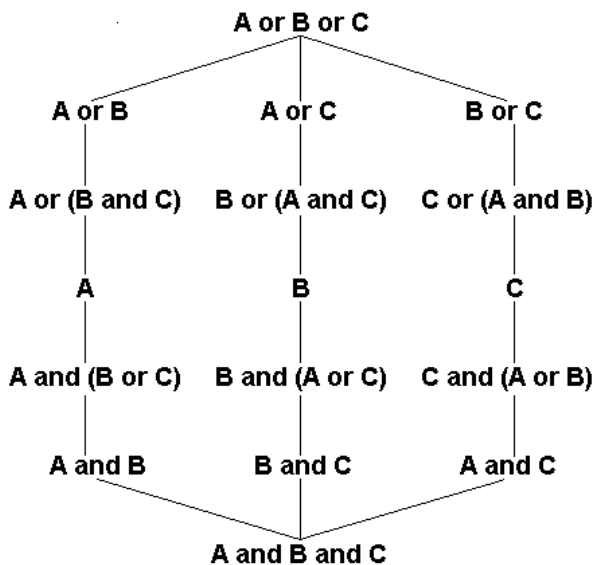

Figure 2. For query (A and B) or C, a configuration of categories in the corresponding directory structure is shown.

(28)...composition model: automatically organize the URLs a user visits over time, making the resulting structure accessible to the user as a resource to use when trying to locate a URL that was visited at some earlier time. Web browsers already maintain a history list, which could form the raw data for this service.

The hypermedia model and the...

(29)...text mining model: see 17).

(30)...query model: see 17).

(31)...composition model: support user annotation of on-line documents. While currently a document can be manually downloaded and then edited, this can hardly be said to support this potentially useful activity.

The citation model and the...

(32)...text mining model: A bot could go through the on-line resources of a particular field and generate a citation index for the literature therein. ResearchIndex exemplifies this (Bollacker et al. 1998 [3]).

(33)...query model: the CD-based version of Science Citation Index [18] allows query-based exploration.

(34)...composition model: since hyperlinks are a kind of citation, see 28).

The reuse model and the...

(35)...text mining model: have a Web bot mine for passages (e.g. sentences) meeting some criteria that are more complex than simply containing particular key terms. For example, sentences might be sought that contain terms that co-occur with the given query terms. The results could be organized via an index.

(36)...query model: see 5), except that results need not conform to the book model, particularly.

(37)...composition model: text browsers such as Web browsers could treat passages that the user highlights with the mouse specially. Such passages would be saved locally and indexed. Thus, whenever the user sees something of particular interest, a simple highlighting operation will cause it to be added automatically to the compendium for later retrieval as needed.

The text mining model and the...

(38)...composition model: Our work on multibrowsers involves building text browsers that take key terms and other inputs, then process documents to display multiple excerpts in separate, non-overlapping windows [17][41]. Then a new cycle of taking user input, locating relevant passages, and displaying them in separate windows occurs. To add composition to this, allow users to click on a window to freeze its contents. As browsing proceeds, additional windows may be frozen by the user. As this continues, the user is composing a screen of particularly useful excerpts.

The query model and the...

(39)...composition model: combining queries and composition could help with Web search engine usability. Currently, search engines typically list URLs at 10 or so per page. Clicking for the next page removes the previous 10 or so from view. It would be useful to allow the user to conveniently designate URLs to be brought forward from one page to the next page. After a few pages, the user would have gradually composed a list of particularly interesting URLs.

5. CONCLUSION

Building on previous work on the epistemology of text exploration, we have listed a number of major models used in systems for interacting with texts. We also have presented a framework for organizing them, shown in Figure 1. The framework was used as a cognitive aid to help identify diverse hybrid text interaction systems. These hybrid system concepts show the value of the exercise. They also exemplify an evaluation methodology for work in the epistemology of text exploration systems: evaluation by focusing on ability to generate ideas. While this methodology is not novel, it is under-recognized because it is so different from the evaluation approaches that are usually suited to results in fields like IR and HCI.

Further investigation of the topic of this paper could involve questions such as:

1. Accounting for more models. There are models that have been proposed that are (perhaps unfortunately) less widely known than the ones we have dealt with here.

2. Examination of the text interaction systems that could arise from combinations of three or more models at a time. Some of the two-model hybrids we have described actually do happen to incorporate three or more, but we have not investigated this systematically.

It seems likely that further work on the epistemology of text exploration would yield additional benefits.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Benest, I.D. A hypertext system with controlled hype. In McAleese, R. and Green, C. (eds.). Hypertext: State of the Art. Intellect Books, Oxford, 1990. (Cited in Sec. 3.2)

[2] Berghel, H., Berleant, D., Foy, T., and McGuire, M. Cyberbrowsing: Information customization on the Web. Journal of the American Society for Information Science (JASIS), 50, 10 (May 1999), 505-511.
(Cited in Sec's. 2, 3.1, & 4 item 6)

[3] Bollacker, K., Lawrence, S., and Giles, C.L. CiteSeer: An autonomous Web agent for automatic retrieval and identification of interesting publications. In Proceedings of the Second International ACM Conference on Autonomous Agents, ACM Press, 1998. (Cited in Sec. 4 item 32)

[4] Burke, R.D., Hammond, K.J., Kulyukin, V., Lytinen, S.L., Tomuro, N., and Schoenberg, S. Question answering from frequently asked question files: Exeriences with the FAQ FINDER system. Artificial Intelligence Magazine, 18, 2 (Summer 1997), 57-66. http://faqfinder.cs.uchicago.edu.
(Cited in Sec. 3.7.1)

[5] Bush, V. As we may think. Atlantic Monthly (July 1945), 101-108. (Cited in Sec. 3.1)

[6] Charoenkitkarn, N., Chignell, M., and Golovchinsky, G. Interactive exploration as a formal text retrieval method: How well can interactivity compensate for unsophisticated retrieval algorithms. In Overview of the Third Text Retrieval Conference (TREC-3) (1994), 179-199.
(Cited in Sec. 2)

[7] Chalmers, M. Comparing information access approaches. J. American Society for Information Science, 50th Anniversary Issue, 50, 12, (September 1999), 1108-1118.
(Cited in Sec. 2)

[8] Davidoff, D. The Pocket Book of Quotations. Simon and Schuster, Inc., 1952. (Cited in Sec. 4 item 25)

[9] Douglis, F., Ball, T., Chen, Y.-F., and Koutsofios, E. The AT&T Internet Difference Engine: Tracking and viewing changes on the Web. World Wide Web 1, 1 (January 1998) 27-44. (Cited in Sec. 4 item 9)

[10] Dourish, P. and Chalmers, M. Running out of space: Models of information navigation. In Ancillary Proceedings of BCS HCI '94 (Glasgow U.K., August 1994).
(Cited in Sec. 1)

[11] Egan, D. E., Remde, J.R., Gomez, J.M., Landauer, T.K., Eberhardt, J., and Lochbaum, C.C. Formative design-evaluation of SuperBook. ACM Transactions on Information Systems, 7, 1, 30-57. (cited in Sec. 3.2)

[12] Ellis, D. A behavioural approach to information retrieval system design. Journal of Documentation, 45 (1989), 171-212. (Cited in Sec. 1)

[13] Everybook Inc. http://www.everybook.net/.
(Cited in Sec. 3.2)

[14] Golovchinsky, G. and Chignell, M.H. Queries-R-Links: Graphical markup for text navigation. In INTERCHI Conference Proceedings: Bridges Between Worlds (April 1993), ACM Press, 454-460. (Cited in Sec. 3.7)

[15] Golovchinsky, G. and Chignell, M.H. The newspaper as an information exploration metaphor. Information Processing & Management, 33, 5 (1997), 663-683. See also Golovchinsky, From Information Retrieval to Hypertext and Back Again: The Role of Interaction in the Information Exploration Interface. Dissertation, University of Toronto, 1997.
http://anarch.ie.utoronto.ca/people/golovch/thesis/final.
(Cited in Fig. 1 & Sec's. 3.3, 3.6, & 4 item 16)

[16] Halasz, F.G. Reflections on NoteCards: Seven issues for the next generation of hypermedia systems. Commun. ACM, 31, 7 (July 1988), 836-852. Revised from a paper of the same title in Proceedings of Hypertext '87, ACM Press, 345-365. (Cited in Sec's. 3.1, 3.6, 3.7, & 3.8)

[17] Huang, J. A Server Based Search Engine-Like Interface for Individual Web Documents. May 1998. Mu, J., Http Client that Browses: A Stand-Alone MultiBrowser for the Web. May 1997. Potti, R., MultiBrowser: A Multiway Lookahead Based Information Customizer for Documents. May 1996. Master's theses, University of Arkansas, Fayetteville.
(Cited in Sec. 4 item 38)

[18] Institute for Scientific Information. Science Citation Index; Social Science Citation Index. Philadelphia.
(Cited in Sec. 4 items 24 & 33)

[19] Kahle, B. Preserving the internet. Scientific American, 276 (March 1997), 82-83. (Cited in Sec. 3.4)

[20] Kamba, T., Bharat, K., and Albers, M.C. The Krakatoa Chronical - an interactive, personalized newspaper on the Web. In Proceedings of the Fourth International World Wide Web Conference (WWW4) (Boston MA, November 1995). http://www.w3.org/Conferences/WWW4/Papers/93/
(Cited in Fig. 1 & Sec's. 3.3 & 3.6)

[21] Kandogan, E. and Shneiderman, B. Elastic Windows: A hierarchical multi-window World-Wide Web browser. In Proceedings of UIST '97 (October 1997), ACM Press.
(Cited in Sec's. 3.6 & 4 item 16)

[22] Kerne, A. CollageMachine: Temporality and indeterminacy in media browsing via interface ecology. In Proceedings of CHI '97 Extended Abstracts (Atlanta, GA, March 1997), ACM Press, 297-298. CollageMachine Web site: http://mrl.nyu.edu/ecology/collageMachine/index.html.
(Cited in Sec. 3.6)

[23] Kulyukin, V.A., Hammond, K.J., and Burke, R.D. Answering questions for an organization online. In Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98), AAAI Press, 532-538.
(Cited in Sec. 3.7.1)

[24] Kupiec, J. MURAX: A robust linguistic approach for question answering using an on-line encyclopedia. In Proceedings of SIGIR '93, ACM Press.
(Cited in Sec. 3.7.1)

[25] Kupiec, J., Pedersen, J., and Chen, F. A trainable document summarizer. In Proceedings of SIGIR '95 (Seattle WA, June 1995), ACM Press.
(Cited in Sec. 4 item 6)

[26] Lawrence, S., Bollacker, K., and Giles, C.L. Indexing and retrieval of scientific literature. In Proceedings of CIKM '99 (Kansas City MO, November 1999), 139-146. System is on-line at http://citeseer.nj.nec.com/cs.
(Cited in Sec. 4 item 24)

[27] Luhn, H.P. The automatic creation of literature abstracts. IBM Journal (April 1958), 159-165.
(Cited in Sec. 4 item 6)

[28] Mani, I. And Maybury, M. Advances in Automatic Text Summarization. MIT Press, Cambridge MA, 1999.
(Cited in Sec. 4 item 6)

[29] McCrary, V. Electronic Book 2000. Electronic Book '99. Electronic Book '98. Conferences. http://www.nist.gov/ebook2000, http://www.itl.nist.gov/div895/ebook99/, and http://www.nist.gov/itl/div895/isis/ebook98.html.
(Cited in Sec. 3.2)

[30] Nelson, T. Project Xanadu home page. http://xanadu.com/. Also see Literary Machines, 93.1 ed. Eastgate Systems, Inc., Watertown MA, 1993. (Cited in Sec. 3.5 & 4 item 19)

[31] Nelson, T. The heart of connection: Hypermedia unified by transclusion. Commun. ACM 38, 8 (1995), 31-33.
(Cited in Fig. 1 & Sec's. 3.3, 3.6, & 4 item 16)

[32] Netmind company. http://www.netmind.com.
(Cited in Sec. 4 items 9 & 10)

[33] Pilkington, R. Question-answering for intelligent on-line help: The process of intelligent responding. Cognitive Science, 16, 4 (1992), 455-491.   (Cited in Sec. 3.7.1)

[34] Richardson, J.V. Question Master. System, http://skipper.gseis.ucla.edu/faculty/jrichardson/html/FirstQUEST/.   (Cited in Sec. 3.7.1)

[35] Shneiderman, B., Byrd, D., and Croft, W.B. Sorting out searching: A user-interface framework for text searches. Commun. ACM 41, 4 (April 1998), 95-98.
(Cited in Sec. 1)

[36] Shneiderman, B. http://www.cs.umd.edu/projects/hcil/pad++/ has many links to papers and other relevant items. (Cited in Sec. 3.6)

[37] Simonson, J. and Berleant, D. Content permanence on the WWW: A decentralized approach. Submitted.
(Cited in Sec. 3.4)

[38] Simonson, J., Berleant, D., Zhang, X., Xie, M., and Vo, H. Version augmented URI's for reference permanence via an Apache Module design. Computer Networks and ISDN Systems 30, 1-7 (April 1998), 337-345. (Cited in Sec. 3.4)

[39] Sollin, K. and Masinter, L. Functional requirements for uniform resource names. RFC 1737, Internet Engineering Task Force, Dec. 1994. http://ietf.org.   (Cited in Sec. 3.4)

[40] Waterworth, J. A. and Chignell, M.H. A model for information exploration. Hypermedia 3, 1 (1991), 35-58.
(Cited in Sec's. 1 & 2)

[41] Zhou, X. Web-Based Document Browser (WBDB): A MultiBrowser for the Web. Master's thesis, University of Arkansas, Fayetteville, 1997.   (Cited in Sec. 4 item 38)