

## **Implications of Hypertext Theory for the Reading, Organization, and Retrieval of Information**

Megan Fitzgibbons, MLIS, BA

### **Introduction**

The study of reading originated with printed materials, but electronic transmission of the written word has long been ubiquitous. From a literary point of view, electronic formats have called into question previous conceptions of textuality. From a more practical standpoint, too, questions have arisen regarding the effect of digital formats on the retrieval and use of information. Theoretical and philosophical understandings of "text" have become intertwined with the technologies that are used to present pieces of writing.

Amongst scholars, *hypertext* is difficult-or even controversial-to define. Theorists Landow and Delany (1991), for example, offer a rather loaded definition of hypertext, describing it as "the use of the computer to transcend the linear, bounded and fixed qualities of the traditional written text" (p. 3). In keeping with this definition, many discussions of hypertext emphasize its nonlinear qualities, but these descriptions have been viewed as problematic and continue to be widely debated amongst scholars. For the purposes of this paper, the more neutral working definition offered by Crane and Mylonas (1991) is appropriate: "'hypertext' refers to the electronic linking of blocks of text" (p. 219). This definition highlights the key property of hypertext, namely its capacity to create conceptual and literal links among disparate sections of a given text or among completely separate texts.

Hypertext is highly significant for all disciplines that are concerned with the creation, dissemination, storage, and philosophy of information. Providing access to information is a core mission of all libraries, and therefore, an important line of inquiry is the effect of hypertext and its nodal organization of information on users' abilities to find and understand texts. At the most basic level, users need general reading skills along with particular equipment and the knowledge of how to use it in order to read hypertext documents. More than that, it is essential for information professionals to understand the reasons that lead clients to seek, consume, and use information in particular ways. This paper explores hypertext's affects on information seekers' understandings of reading, information seeking, resource discovery, and the bibliographic universe, particularly in relation to information professionals' abilities to provide high-quality information services.

### **Information Seeking and Use**

#### **Reading hypertext**

Scholars are divided as to whether or not the act of reading differs across hypertext and print environments. Surveys of general reading behaviors make it clear that the types of associative thinking often ascribed to reading hypertext are not at all unique to that format (McHoul & Roe, 1996). Some hypertext documents are better read in a linear manner, just as some print works, such as reference books, are designed to be read in any sequence. Genre certainly plays a role in how texts are read, regardless of whether the work is print or digital (Miall, 1999).

Scholarly articles, for example, at least in the English-speaking world, are most often presented hierarchically. A typical structure may include an abstract, statement of objectives, literature review, description of methodology, findings, analysis, and conclusions. Within each defined section, the content progresses from a general statement to more specific details to a concluding statement. Other characteristics of the scholarly formula may include a direct correspondence between sentences in the abstract or conclusion and more fully developed sections of the document (Blustein, 2000).

Of course, not all writing follows this model, and the content of a given piece is often much more complex than the formula with which it was crafted. Ideas, terms, and scholarly apparatus are connected throughout the document on a conceptual rather structural level. Hypertext, via the capabilities of markup languages, is therefore suited to literary and scholarly communication insofar as it can organize a document on many levels.

## Reading tasks

Protopsaltis and Bouki (2005) developed a model for describing the process of reading articles in a hypertext format. The model consists of eleven components in which readers set a goal, scan and read the sections or categories, select a strategy, build and incorporate knowledge, evaluate the results, and repeat the cycle when necessary. The pattern is not particularly surprising; it is similar to other established models of information seeking behavior and reading cognition. An important point, though, is that readers interact with the hypertext in both linear and random ways. The process accommodates a significant amount of backtracking, re-reading, and "monitoring" or assessing the content and structure of the text, which are all activities that rely on proper labels and linkages between document sections.

Protopsaltis and Bouki's model takes on additional significance when understood in the context of the primary navigation strategies that readers were found to use (p. 163), namely:

- serial / linear
- mixed
- mixed review

Readers who used the linear strategy chose links in the order that they were presented in the document. In the mixed method, readers chose links without any discernable pattern. Finally, readers who utilized the mixed review strategy scanned all available choices before selecting links, which sometimes resulted in a linear reading pattern and sometimes a more random order.

## Hyperlinking

The problem of reader disorientation is a concern that commonly emerges in the hypertext environment. The structure of a piece might not as easily be discerned on a computer screen as on a printed page for the simple reason that the limited amount of text that can be viewed at a given time makes it more difficult to scan the document and get a "feel" for its structure. Hypertext markup languages allow authors to organize documents' content both visually and conceptually to minimize disorientation.

As described, the most basic property of hypertext is its capacity to create links within and among texts. Within a document, elements of the scholarly apparatus, such as definitions of terms, notes, and references, can all be hyperlinked to provide readers with additional information at the point of need. Hyperlinks also can be used to draw connections among concepts that are not explicitly linked in the hierarchical document structure. For example, a "semantic link" (Blustein, 2000, p. 203) connects one passage that summarizes an argument to a more in-depth discussion of the point elsewhere in the document.

Moreover, hypertext inherently lends itself to inter-textual linking. In terms of reading tasks, inter-textual linking opens the reader to a virtually infinite body of literature that is related to the original

document in some way. However, this can again exacerbate the potential for disorientation as readers may lose focus of the primary reading task. The logical structure of the original document becomes irrelevant when another text is encountered. In the end, though, the linking properties of hypertext can take readers to additional levels of comprehension beyond the text itself. In this way, hypertext can enhance the reading experience on multilayered levels that are not possible in the print world (McGann, 1995).

## The reading experience

When reading a hypertext, the ultimate path that readers take is dependent on the choices and structural cues provided by the document's creator. It has been argued that the hypertext environment "frees" a text and gives greater control to readers to shape their own reading experiences (e.g., Bolter, 1991; Delany & Landow, 1991). This point is widely debated in academic discussions of hypertext, but if certain theorists are correct in their assertions that hypertext learning requires different cognitive skills from traditional print reading, then education is needed to enhance users' abilities in using hypertext documents.

Readers' choices in navigating a hypertext seek a coherence in the text that meshes with their information goal. Protopsaltis and Bouki (2005) and Blustein (2000) assume that the end goal of *comprehension* always guides the reading process; an aesthetic experience is not particularly desired. However, in both studies, the strategies that readers employed had more to do with preference and satisfaction rather than increased performance and comprehension. Blustein (2000) found that readers often asserted that they preferred hypertext documents, even when they did not use them particularly well in comparison with print sources. This suggests that while reading styles may be the same for both hypertext and print, there are certain fundamental differences in the ways in which people *think* about reading hypertexts. Readers' attitudes toward electronic works differ from their opinions of print resources, and therefore, they tend to approach hypertext documents with different methods and goals in mind. Information professionals should take an interaction-centered approach to understand how the hypertext format and the concerns of the reader affect each other in providing resources and education for their clients.

## Hypertext and the Organization of Information

Information literacy entails an understanding of a text itself, but more than that, it requires an understanding of the item's characteristics, such as genre, authority, and currency, and its significance in relation to other documents. While hypertext theorists remain divided as to how hypertext affects reading cognition, the format certainly alters readers' understanding of how a document is structured and where it is located in a body of literature. Links to outside content can serve to position a document in the "bibliographic universe," demonstrating how it relates to other material and how it contributes to the relevant field of knowledge.

## The Memex

No discussion of hypertext is complete without mentioning Vannevar Bush's oft-cited article "As We May Think," which appeared in *The Atlantic Monthly* in 1945. In the piece, Bush described the "memex," a container for all print material that the owner (or "user" in today's parlance) chose to input. Beyond simply being able to recall information, the primary goal of the storage system was to allow the user to visualize and access the *connections* among items, tracing related concepts from one place to another, crossing documents' boundaries regardless of type. These connections would be physically retrievable and retraceable, allowing documents to be located efficiently and knowledge to be shared.

In addition, since the links and other added notes would be created by the individual user, the information contained in the memex would be highly personalized and practical. Indeed, Bush conceived of the memex as a mechanized extension of a person's own memory. The ultimate purpose of this

mechanization would be the continued progress of society, since "[man] has built a civilization so complex that he needs to mechanize his records more fully if he is to push his experiment to its logical conclusion and not merely become bogged down part way there by overtaxing his limited memory (Bush, 1945, section 8)."

## Metadata

Bush recognized the value of linkages among documents' metadata records for the efficient retrieval of information. The theoretical memex would allow users to circumvent traditional cataloging schemes and instead search using natural language keywords and known connections between concepts. Bush pointedly argued that traditional cataloging structures were becoming increasingly inadequate for efficient information storage and retrieval. In his view, "artificial" hierarchical classification systems had limited usefulness, simply because "the human mind does not work that way" (Bush, 1945, section 6).

The use of hypertext in today's online catalogs allows library users to isolate a single subject term, name, or call number and locate records related to that particular field's value with one click of the mouse. Hyperlinking thus creates equality among the individual fields of a catalog record, in that users can find the record from any access point and search for each term individually to find related resources. Of course, the hyperlinked items are predetermined by catalogers as access points to the records, so the possible targets of a single link are finite. However, given the ease with which users can pass from one record to another, there are innumerable routes by which information can be located. Library catalogs and media retail websites alike have taken advantage of the power of hypertext by including general links to "related items" that may be of interest to clients, giving them assistance in the form of directed navigational choices.

In addition to linked catalog records, newer applications such as citation searching and tag clouds certainly harness the nodal nature of hypertext that is so celebrated among theorists. Metadata is employed in a freer fashion, including via user-generated content, to connect documents. Hypertext allows individual documents to be linked on a number of levels, thereby enhancing the hierarchical classification that has been standard in the past. When the connections between items are encoded into a system, it makes it much easier to view and trace the links among works and thereby access information with greater efficiency.

## The Bibliographic Universe

The goal of somehow bringing order to vast bodies of information certainly is not new. Sheehan (2003) discusses one organizational technique of the early modern period whereby an encyclopedia of general knowledge was constructed using the text of the Bible as "nodal points through which a variety of information could be accessed" (p. 59). The mention of a particular concept in the biblical text lead to peripheral explanations of ideas and definitions contained in works from virtually every field of knowledge. If it was brought into an electronic hypertext environment, this encyclopedia could provide hyperlinks from the biblical text to all the other mentioned documents. Instead of references and annotations on the physical page of the Bible, hypertext would allow readers to go directly to the actual document to which the author alluded.

In addition to concepts of universal, all-encompassing encyclopedias, a historical dream of librarianship has been to identify and create access to all published works through a world library or world bibliography. Earlier manifestations of this goal include the ancient Alexandrian library and Konrad von Gesner's sixteenth century *Bibliotheca Universalis*. Many believe that hypertext will be the means through which these ideals can be finally be accomplished, as reflected in initiatives like the Google Print Project and Project Gutenberg. According to Google's Library Project information website (2007), the company's "ultimate goal is to work with publishers and libraries to create a comprehensive, searchable, virtual card catalog of all books in all languages that helps users discover new books and publishers discover new

readers." (para. 3). In the Google project, hypertext provides the basis through which works are linked together to form a retrievable corpus of information.

In these examples, hypertext does not necessarily revolutionize basic principles of information management. Rather, it creates ways to implement long-standing ideas of library and information science in new ways, in new environments, and with new bodies of literature. As librarian Wills (1999) suggests, "the hypertext environment is an extension of the long-held impulse to bring together a universe of knowledge through a single encyclopedic work or a library of materials classified to show their relationships" (Print and hypertext compared, para. 11). When employed for the very old goal of linking concepts across documents, hypertext has an influence on people's understanding of collections of texts. It makes possible a broad look at bodies of literature by allowing connections between works to be literally seen.

## Conclusion

So what is the impact of hypertext theories and technologies on information services? Hypertext theory itself is certainly a less-than-homogeneous body of scholarship. Some writers argue that hypertext has brought radical change to reading, writing, and the organization of information, since its unique properties allow texts to be freed from the constraints of traditional printed books. Other writers, in contrast, criticize these theories for attributing too much of a revolutionary impetus to hypertext, questioning whether hypertext has had any effect on reading and writing cognition processes at all. Underlying all of hypertext theory is the reality that hypertext is currently being used in an enormously wide variety of settings in order to manage information via its linking properties.

Information professionals can best understand the implications of hypertext theories for their own field by integrating it with principles of information service that been developed in the past. In order for information professionals to form effective principles of information services, client needs, attitudes, and abilities must be combined with considerations of the technical capabilities of hypertext itself. Literacy is a key component of the information seeker's side of the equation, because the abilities to locate, read, and evaluate texts are the basis of successful information gathering processes. Hypertext literacy is not an isolated module within the larger goal of literacy. Because hypertext works are no more standardized than print writings, hypertext literacy is better understood in the context of larger patterns of information seeking behaviors (Moulthrop, 2005).

Just as hypertext use has been adapted to pre-existing theories of information seeking behavior, so too has hypertext been applied to fulfill organizational ideals that have long existed in librarianship. When employed in information retrieval systems and collections of information resources, hypertext is able to augment existing methods of organization by allowing individual documents to be linked to others with great efficiency. While many principles of organization have remained fundamentally the same, hypertext affords a new method of implementing these principles. In this regard, hypertext has not yet proved as revolutionary in the realm of information services as some theorists believed. But while it has not overturned existing patterns of information seeking and management, it has certainly modified day-to-day practical processes in the information professions. Perhaps more significantly, the development of hypertext has given rise to questions that profoundly consider the nature of information seeking and the resulting larger implications for understandings of the human mind.

## References

American Library Association (1989). *Presidential committee on information literacy: Final report*. Retrieved September 3, 2007, from <http://www.ala.org/acrl/legalis.html>

Bolter, J.D. (1991). *Writing space: The computer, hypertext, and the history of writing*. Hillsdale, NJ : Lawrence Erlbaum.

- Blustein, J. (2000). Automatically generated hypertext versions of scholarly articles and their evaluation. *Proceedings of the Eleventh ACM Conference on Hypertext and Hypermedia*, 201-210. Retrieved September 3, 2007, from <http://doi.acm.org/10.1145/336296.336364>
- Bush, V. (1945). As we may think [electronic version]. *The Atlantic Monthly*, 176(1), 101-108. Retrieved September 3, 2007, from <http://www.theatlantic.com/doc/194507/bush>
- Crane, G., & Mylonas, E. (1991). Ancient materials, modern media: Shaping the study of classics with hypermedia. In Delany, P. & Landow, G.P. (Eds.), *Hypermedia and literary studies* (pp. 205-220). Cambridge, MA : MIT Press.
- Delany, P., & Landow, G.P. (Eds.). (1991). *Hypermedia and Literary Studies*. Cambridge, MA: MIT Press.
- Google (2007). *Library Project: An enhanced card catalog of the world's books*. Retrieved September 3, 2007, from <http://books.google.com/googlebooks/library.html>
- Landow, G. P. (1992). *Hypertext: The convergence of contemporary critical theory and technology*. Baltimore : John Hopkins University Press.
- McGann, J. (1995). *The rationale of hypertext*. Retrieved September 3, 2007, from <http://www.village.virginia.edu/public/jjm2f/rationale.html>
- McHoul, A., & Roe, P. (1996). *Hypertext and reading cognition*. Retrieved September 3, 2007, from <http://wwwmcc.murdoch.edu.au/ReadingRoom/VID/cognition.html>
- Miall, D. S. (1999). Trivializing or liberating? The limitations of hypertext theorizing. *Mosaic: A Journal for the Interdisciplinary Study of Literature*, 32(2), 157-172. Retrieved February 22, 2007, from Academic Search Premier database.
- Moulthrop, Stuart. (2005). What the geeks know: Hypertext and the problem of literacy. *Proceedings of the 16th ACM conference on hypertext and hypermedia*, 227-231. Retrieved September 3, 2007, from <http://doi.acm.org/10.1145/1083356.1083402>
- Protopsaltis, A., & Bouki, V. (2005). Towards a hypertext reading/comprehension model. *Proceedings of the 23rd annual international conference on design of communication: Documenting & designing for pervasive information*. Retrieved September 3, 2007, from <http://doi.acm.org/10.1145/1085313.1085349>
- Sheehan, J. (2003). From philology to fossils: The biblical encyclopedia in Early Modern Europe. *Journal of the History of Ideas*, 64(1), 41-60. Retrieved September 3, 2007, from the Project Muse database.
- Wills, D. (1999). The nature of hypertext: Background and implications for librarians. *The Journal of Academic Librarianship*, 25(2), 134-139. Retrieved September 3, 2007, from the Library Literature database.