Citations in Hypermedia: Maintaining Critical Links
Corinne Jörgensen and Peter Jörgensen

The lack of complete and accurate citations for referenced works is a persistent and significant problem in scholarly writing. The advent of new forms of electronic publishing, especially hypermedia, has the potential for either aggravating or alleviating this problem. Incorrect and missing citations will become more prevalent (and, indeed, have become so within large bibliographic databases) without a concerted and cooperative effort on the part of systems developers to provide adequate references and the means for easily accessing and downloading them. This article examines some of the basic issues involved in the problem of maintaining the critical authorship links between source and expression in one specific form of electronic publishing—hypermedia—and proposes some solutions.

A

n outgrowth of Vannevar Bush's Memex concept and Ted Nelson's hypertext, "hypermedia" refers to computerized systems that incorporate multiple linkages between items of information within a variety of media. Hypermedia systems currently include video, still video, sound, text, computer-generated animation, and facsimile. Hypertext (which is limited to textual information) is a subset of hypermedia and will be included in the terms "systems" and "hypermedia" in this article. This paper does not address self-contained hypermedia applications developed for project management, for single users, or for authoring technical in-house documents.

Hypermedia makes use of extensive links and paths. Links are a means of connecting fragments of information so that users can jump from one fragment or document to another by activating the link. The exact means of activation is system dependent, but typical methods are iconographic buttons, menus, or the selection of text with a mouse. A path is created when a user activates a sequence of links. Such a sequence may be predetermined by the author or may be created ad hoc by the user. A path may or may not be stored by the system for later recall or replay. Links and paths provide the ability to navigate these systems in a flexible manner and give users the opportunity to explore many related ideas in a nontraditional, nonlinear format.

The introduction of personal computer-based programs such as HyperCard™ and Guide™ has prompted an increased interest in and access to hypermedia development within the academic community. Information professionals and scholars recognize hypermedia's potential to add a new dimension to information storage and retrieval by "bring[ing] order to information chaos." However, there has been little or no consideration of the
need for and difficulty of maintaining authorship linkages so that the user can always see complete and accurate citations to referenced works. Unfortunately, information chaos will reign unless these systems include explicit attribution of sources.

**SOME CURRENT SYSTEMS**

To date, faculty and computer center staff at universities and colleges are the primary developers of many hypermedia systems, which serve as instructional vehicles, electronic textbooks, or scholars’ workstations. Project Jefferson, developed at the University of Southern California (USC), is an example of an instructional system. The project’s aim is to “develop an effective model of computer-based instruction that includes the use of research tools and a standard interface that can act as a front end for courseware across a variety of different disciplines.”

The system includes citations to external materials, encyclopedic background material, an electronic notebook containing writing assignments, and outlining tools for writing papers, all in a hypertext package. Project Jefferson currently provides access to a number of databases loaded on the USC’s library mainframe computer and can be used for researching a variety of questions.

The Visual Courseware Machine (VCM) developed at MIT through Project Athena is one example of a scholar’s workstation. VCM multimedia stations provide students with access to information stored as full-motion color video (on disc and cable television), digital audio, high-resolution graphics, and CD-ROMs for use in their coursework. Thus, “the student is interfaced with a multimedia networked station that is very much like a gateway into a new world of learning.”

Intermedia is yet another hypermedia system designed at a major university to support the learning process. Originally developed at Brown on IBM PC RTs and DEC MicroVAX systems, it is now commercially available under AU/X on Apple Macintosh computers. Intermedia features several applications that can be used to create and maintain hypermedia documents and their links, including a text editor, a graphic editor, a scanned-image viewer, a three-dimensional object viewer, and a time line editor. With these tools users can create and edit documents of textual or graphic content and make and follow links between documents. Intermedia has been used in courses as diverse as English literature and cell biology.

Yet other examples of a hypermedia project include sports and medical information systems. On the commercial front, BRS Information Technologies has implemented a form of hypertext by linking citations in Medline and Medline references on AIDS to source documents in BRS’s Comprehensive Core Medical Library (CCML), a full-text database containing a selection of medical journals and books.

---

Information chaos will reign unless these systems include explicit attribution of sources.

---

The hypermedia systems described above demonstrate a variety of techniques for presenting citations to users. These techniques range from appending the author’s initials to the information, with a full citation given elsewhere (Intermedia), to displaying the full citation at all times (Project Jefferson). Citation practices in hypermedia systems are, therefore, inconsistent, paralleling, to some extent, those in print media.

**CITATION PROBLEMS IN PRINT MEDIA**

As mentioned earlier, citations are critical links to source materials, and incorrect citations have a substantial impact on the research community by rendering sources inaccessible. Yet citation problems remain a persistent and significant difficulty in scholarly writing. Recent articles in *Library Quarterly* and the *Chronicle of Higher Education* examining the problem of incorrect citations list citation error rates ranging from 7 percent to 50 percent (depending on various cri-
criteria used to define error rates, such as the ability to obtain materials with problem citations through interlibrary loan), with an average error rate within the scholarly literature hovering between 20 percent and 30 percent.\textsuperscript{13}

It must be emphasized that these are errors made by experienced, educated researchers, not by the student population. Students, who are the primary users of instructional hypermedia, make even more mistakes. The reasons underlying citation problems are many, ranging from simple errors of transcription or omission to intentional falsification and deception. The numerous formats used for citations also contribute to the proliferation of citation errors, with requirements varying from journal to journal and style book to style book.

**CITATION PROBLEMS IN HYPERMEDIA SYSTEMS**

Several features of hypermedia systems as they are currently implemented also contribute to a loss of the critical links that bind information to author in the reader's mind. As mentioned earlier, a variety of techniques exists for presenting citations. Because these multiple citation formats confuse students in the print media, one can expect similar problems with online information.

Another aspect that may lead to citation problems in hypermedia is the lack of visual cues differentiating one source from another on the computer screen. Text within a hypermedia system lacks the physicality of a paper document and its attendant anchoring cues, such as page numbering (hypermedia documents present information on a screen lacking any sequential clues) and even such subconsciously registered cues as varying weight, cover color, or page size.\textsuperscript{14}

Gary Marchionini and Ben Shneiderman list the following typical access points in printed books: "table of contents, indexes (author, subject, permuted, etc.), glossary, chapter, article (section), physical page, paragraph, footnotes, reference notes, lists, and appendices" and state that "hypertext databases can support all these access points except physical page."\textsuperscript{15} Nonetheless, these access points are often missing from hypermedia works. While it may not be necessary or desirable to transfer directly all of the components of print media to a hypermedia format, system developers must consider what is lost when readers do not have access to these components, and they must provide some means for anchoring readers within a particular document.

**Multiple Sources in One "Seamless" Document**

User disorientation within what is referred to as the "seamless environment" of the information space is also quite common.\textsuperscript{16} Because a variety of source data with multiple links or multiple media representations of the information is a key feature of hypermedia, these systems tend to be quite large, often bewilderingly so. Unfortunately, the flexible presentation capabilities of hypermedia can obscure the original sources. The diverse information sources have various origins, dates, and reliabilities; therefore, it is essential that users be able to check the sources of material easily.\textsuperscript{17}

Developers of hypermedia systems explored several mechanisms, such as maps, webs, and outlines, for dealing with the user disorientation experienced in "hypermedia space." However, none of these methods has been demonstrated to be fully effective in preventing disorientation, and further research is needed in the area of hypermedia navigational aids. Because hypermedia systems provide a flexible approach for exploring information, perhaps users also should be able to choose the mechanism (e.g., map, outline, or other instrument) that they find most useful.

**The Computer as Author**

Students may accept the hypermedia system as being the primary source of information. For many students, the computer replaces the author. This may be less of a problem when the system contains information from faculty members' personal lecture notes or is produced cooperatively by multiple authors, but it is too easy for these perceptions of the com-
puter as author to carry over to source information for which students should provide citations. Related to the concept of computer as author is a pervading sense among students of the computer as authority. There is a tendency to believe that “if it comes from a computer it must be true.” Even experienced researchers may too easily accept computer simulations as accurate representations of the real world. Just as readers grant printed matter a high degree of ethos, they accord information presented on a computer screen a measure of authority over information presented in other formats.

Another aspect that may lead to citation problems in hypermedia is the lack of visual cues differentiating one source from another on the computer screen.

Indeed, comments from students using early versions of Intermedia at Brown University illustrate the tendency to accept the computer as author and authority. Stimulated by Intermedia, students began to participate more in classroom discussions. However, they frequently cited the source of their information as the system itself (“Intermedia says”) when no indication of personal authorship appeared within the texts. The authors of Intermedia modules have since decided to identify themselves, stating about the earlier implementation: “This attempt to avoid emphasizing personal ownership of the essays and graphic documents made the predecessor of the ‘Dickens Web’ appear a monolithic statement of official truth.”

Lack of Document Permanency

Another characteristic of documents within an electronic information system is their lack of permanency. Online information, unlike the printed page, can be edited or added to without a visible indication that a change has been made. Some systems allow student authors as well as faculty to contribute to the information pool. This multiple authoring can add to a sense of confusion as one returns to a document only to find that it is not the same as the version previously accessed, or that newly added links have appeared, making it difficult to recall an earlier path.

Casual Attitudes of Developers

Some developers of these systems seem to take a casual attitude toward the importance of citations, thus contributing not only to the problem of inadvertent plagiarism on the part of students but to the difficulty in resolving issues of fair use. At Hypertext ’87, held at the University of North Carolina-Chapel Hill, conferees discussed the issues of tracing intellectual ownership in a hypermedia environment. At a more recent conference, the developers of a hypermedia system were questioned about the lack of citations in their system. They replied that they had decided not to deal with the questions of copyright. Discussing this same conference, Kay E. Vandergrift commented:

A number of people involved with the development of hypermedia programs described using scanners to lift diagrams, photographs or segments of text from other sources for inclusion in their stackware. Most did not mention any concern with the rights of the original authors of those materials or the legality of what they were doing under copyright law. When the question was raised by a member of the audience, one speaker indicated that he made some modifications of the scanned material; therefore, it became his own. This practice is undoubtedly not unique to the particular speaker and is going to become a more serious problem as many hypermedia products move from in-house prototypes claiming “fair use” to those shared with other professional colleagues through informal networks to full-blown commercial products.

Electronic Information and Copyright

New forms of electronic publishing also raise many complex copyright issues. Hypermedia systems that are distributed over networks are one example.
A 1986 report by the Office of Technology Assessment (OTA) addressed the potential copyright problems of electronically stored information. The OTA found that, when accessing information over a network, three important issues emerge: the proprietor's right to reproduce the copyrighted work, the proprietor's exclusive right to make derivative works, and the proprietor's exclusive rights to distribute, perform, or display the copyrighted work.

Developers of hypermedia have decided not to deal with the questions of copyright.

The report cited one example of a potentially illegal activity: the sending of copyrighted information from one user to another over an electronic network, thereby infringing on the proprietor's right to distribute the work. The doctrine of fair use depends on the circumstances surrounding the use. Thus, each possible violation of fair use must be argued in court. However, if hypermedia developers fail to ensure proper crediting of original sources in their systems, it will certainly weaken the case for fair use and could lead to a more restrictive environment for development of these systems.

SOLUTIONS

Several possible solutions exist to the problem of maintaining the critical links between source and information in hypermedia systems. Three will be discussed here: human resources solutions, technical solutions, and the development of standards for hypermedia systems.

Human Resources Solutions

More and more library functions and holdings are becoming computerized, and librarians are taking an active interest in these systems and their development by providing feedback to vendors of the systems. Likewise librarians—with their highly developed bibliographic and information skills—should become involved in hypermedia development. They are often expected to retrieve cited materials and, thus, have a vested interest in ensuring that complete and accurate citations are easily accessible in hypermedia systems.

Librarians possess skills that are applicable to the development of expert systems and, by extension, to hypermedia systems. These skills are discussed in a 1988 article by Anne Morris and Margaret O'Neill. Foremost among the skills is an understanding of the process of knowledge representation by means of such long-established techniques as indexing, classification, and thesaurus building, all of which deal with the meaning and structure of words, especially in their relational and hierarchical contexts. In particular, "LIS professionals could play an important role in helping to structure the knowledge base by creating links between terms, particularly cross-disciplinary ones, organizing hierarchical structures, and making relationships between elements clear." Librarians' skills are critical to creating meaningful links in hypermedia systems. According to Carl Franklin: "Hypermedia are part of a frontier in which the leadership and expertise of this profession [librarianship] can make a critical difference. No other profession is more familiar with the organization of knowledge." A second important area in which librarians could participate is the creation of the intellectual content of hypermedia systems. While many in the computer and knowledge engineering industry and academia think that subject experts are the most qualified to recommend texts in their fields, there can be problems with this approach. Morris and O'Neill list several. Frequently, experts forget or overlook the basic literature that is the foundation for understanding in their fields. Books that are recommended as key texts may, in fact, have been superseded by newer ones, and experts are often too busy to keep up with every new development. Subject experts also may lack a broader viewpoint that forms the foundation for creative linking within hypertext systems. Li-
brarians, as “custodians of the world’s literature ... have access to the necessary sources to give context and flesh to expert systems.”

A third important area in which librarians could apply their skills is the domain of user understanding. Librarians have been helping patrons navigate the original and confusing hypermedia system—the library—for a long time. They understand the issues involved in negotiating complex information spaces with multiple sources, and many have been creative in providing new tools and services to overcome the disorientation that is common to library users as well as to users of hypermedia systems. Librarians could apply their experience in this area to solving the problems of user orientation and way-finding within hypermedia systems.

Librarians can seize the initiative and widen the scope of the library to include a variety of new systems, including hypermedia.

Librarians, by and large, also are acutely aware of copyright issues. And they are becoming even further involved in these issues with the addition of more online resources and the expansion of collections to include such services as circulating software. By being included in the early design stages of hypermedia systems, they can help ensure that copyright issues in electronic publishing are adequately addressed and that a source will not be separated from its citation.

The potential benefits are not all one way; librarians and libraries also would benefit from inclusion in the development process, which would provide new inspirations for providing services and instruction to patrons. Hypermedia systems are being expanded to include interfaces to the computerized resources of the library, and librarians can expect to see more hypermedia systems being developed that provide links to bibliographic and full-text data. With the rapid development of new nontraditional information systems, a great number of which are external to libraries and may be commercial ventures, many libraries are suffering from identity crises. By becoming involved in the development processes of appropriate systems, librarians can seize the initiative and widen the scope of the library to include a variety of new systems, including hypermedia.

Technical Solutions

Automatic Citations. The powerful technology and the integrated nature of hypermedia also can encourage the development of tools that facilitate or enforce rigorous citation standards by automating the process. This idea is not new: Ted Nelson’s original system specifies complete protocols for automatically crediting author “bank” accounts for each use of published material. In fact, as the research/writing/publishing process becomes more centered around the computer, the automatic inclusion of complete and accurate citations becomes trivial. Commercial word processing programs already include the automatic transfer of citations (in footnotes) with text that is electronically transferred from one document to another. The Project Jefferson Notebook also includes such a feature. When source text is copied by clicking on a camera icon, the citation automatically is included. When that text is inserted into another document, the citation automatically is included. Instructors using the Project Jefferson Notebook report that submitted papers are well-supported; ease of copying citations is undoubtedly partly responsible.

A German literature seminar at Colgate University uses a system incorporating automatic transfer of citations. This system’s primary function is to facilitate information sharing between students. Students enter notes from their readings into a HyperCard stack. Then the students periodically log onto a network and initiate a transfer process that adds the notes to a master collection and, at the same time, updates their personal copies. Again, the program automatically
appends the author and title of the source document to the texts being transferred.

These two systems contrast with other extant systems that simply list authors' names or initials at the end of a document, with a corresponding bibliographic entry elsewhere in hyperspace. Without an easily accomplished method for both accessing and downloading citations as well as source material, inadvertent and deliberate plagiarism will occur. It is unlikely that plagiarism would happen immediately within the context of a course employing a hypertext database, but hypertext systems distributed over networks can be used in the creation of personal databases that may be drawn on at a later time, in another course, school, or job.

Standard Citation Icon. In addition to employing automatic downloading of citations, developers should agree on a standard citation icon as recognizable as footnote numbering. Activating this icon would provide the complete citation. A brief citation consisting of author and title (abbreviated if necessary) should always appear on the screen—again, in a standardized format and location, such as within the title bar across the top of a window. Because windowing has become a standard presentation format, the addition of a small amount of text should not pose a problem. This provision of consistent icons and windows for reference information will help to orient readers within a particular document and prevent them from being lost in hyperspace.

Developing Standards

Multiple interfaces and data models exist in today's hypermedia systems. Market constraints and the needs of particular applications have created a situation in which incompatible and competing systems exist. As is often the case in the early development of new systems, guidelines and standards are not widely used. This unregulated and flexible atmosphere promotes the maximum freedom of creativity, innovation, and growth of new ideas during the initial stages of development. However, de facto standards established by pioneering developers, and adopted unconsciously by others, have often become firmly entrenched, even when they are recognized as being less than satisfactory.

Domestic organizations, such as the American National Standards Institute's accredited standards committee, X3, and international standards organizations are currently developing standards for hypermedia systems, especially in the areas of user requirements and interconnectivity. The Dexter Workshops have recognized the need for a formal model and vocabulary for hypertext and the creation of data interchange standards. While there is recognition that a potential citation problem exists, little work is being done by any of these groups on the problems of maintaining the critical links to source documentation.

CONCLUSION

To date, developers have not addressed the complex issues of citation-source linkages and related issues of copyright in hypermedia systems. While several solutions have been proposed and, undoubtedly, other creative solutions exist, the purpose of this article is to raise an awareness of the potential problems that could arise without sufficient bibliographic control within hypermedia systems. There must be a commitment to correct bibliographic and citation deficiencies in these new systems if the library community is to avoid the further separation of source from citation. An awareness of the problem, however, is the essential first step in this endeavor.

REFERENCES AND NOTES


7. Ibid., p.353.


9. Ibid., p.44, 81.


14. Every librarian has been asked to help a patron find the "large, heavy, dark red book somewhere in this area" that was used earlier!


18. Many librarians have encountered this attitude among students who are having trouble finding a citation from an online search printout. The possibility that there may, in fact, be a mistake in the computerized bibliographic file never occurs to students.


This has happened in command-line vs. graphical interfaces.

"Documents on Hypermedia Standardization," internal working drafts produced by X3V1 and SC18, accredited standards committees of ANSI and ISO.

...begin with world-class information. It's the kind of information you'll find in Biological Abstracts® (BA) and Biological Abstracts/RRM® (Reports, Reviews, Meetings) (BA/RRM), vital resources for references to the world's life science literature.

**Comprehensive Coverage...**

...sets BA and BA/RRM apart from other reference materials. BA contains full bibliographic information and abstracts from the latest biological and biomedical literature, carefully selected from approximately 9,000 journals worldwide. BA/RRM complements this global service with exhaustive coverage of papers from international meetings and symposia, as well as books, reviews and other items not found in most life science reference tools.

**A Solid Information Foundation...**

...for life science research is yours with BA, BA/RRM and their Cumulative Indexes. Together, this powerful information package provides thorough, accurate coverage of the research that can lead to world-class discoveries!

---

**Find Out More**

Yes! Please send more information on Biological Abstracts and Biological Abstracts/RRM!

<table>
<thead>
<tr>
<th>Name</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Organization</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Address</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>City</td>
<td>State</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Postal Code</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Return this coupon to the BIOSIS Marketing Department, 2100 Arch Street, Philadelphia, PA 19103-1399 USA or the Official Representative in your area. Or call toll free 1-800-523-4806 (USA except PA); (215) 587-4800 (worldwide); Fax (215) 587-2016.

BIOSIS®
Information for Today's Decisions and Discoveries

BIOSIS is a registered trademark of Biological Abstracts, Inc.