

# Anthropology: Implications for Form and Content of Web-Based Scholarship

Glenn Davis Stone  
Washington University

© Sage Publications

*Social Science Computer Review*  
1998 16:4-15

**ABSTRACT** : Although it is rapidly emerging as a forum for social science scholarship, the Worldwide Web is now used mainly to propagate conventional scholarship. This is certain to change, given the simultaneous expansion in the capabilities of web-based communication and decline in libraries' ability to keep abreast of print scholarship. This article outlines opportunities for changing and enhancing the nature of the scholarly (peer-reviewed) articles on the web. Three mechanisms are discussed by which the form and content of the scholarly article can be improved: (1) the use of hypertext structuring, (2) the integration of multimedia components into articles, and (3) the use of differentiated pointers. Examples are on an accompanying web page at <http://arts.wustl.edu/~anthro/demo/>.

---

Invent the motion picture, Marshall McLuhan wrote, and the first thing thought to do is to put on a play and film it. It is only much later that the possibilities of motion picture technology become clear beyond a single theater with a single batch of sets and actors. This is an apt metaphor for social science and the Worldwide Web (hereafter, "the web"). Social science scholarship is well represented on the web; most journals have web "shopfronts" with tables of contents and information for authors, some now have full text of articles, and some journals exist only online.<sup>(1)</sup> In the overwhelming majority of cases, this new medium is used to propagate scholarship differing little or not at all in form from conventional printed work. Yet web-based scholarship provides opportunities for qualitatively different kinds of scholarly products, embodying enhancements that range from the convenient to the revolutionary.

The arrival of this new medium happens to coincide with a dramatic drop in the ability of research libraries to keep abreast of print publications. The experience of the Washington University library is typical for a research university: over the past decade, annual expenditures on print based journals more than doubled, while the actual number of titles in the collection fell by almost five percent. The number of scholarly books published by university presses is dropping sharply in large part due to academic libraries' ongoing inability to pay for their existing subscriptions. While existing print collections will continue to be important to researchers, and while such collections will continue to grow for the indefinite future, web based scholarship represents the future for both scholars and libraries.

This paper examines some of the implications of this for the form and content of peer-reviewed scholarly publications. I include examples of [html](#) (Hypertext Markup Language) for implementing some of what I describe. Although this is not a how-to guide, it at least indicates what is involved, and provides a starting point for those interested in developing

web-based papers. The particular vantage point is that of an anthropologist, but the discussion is mostly applicable to other social sciences as well.

## Hypertext

Hypertext literally means "text above." It is used for various aspects of information management, but its signal function is to refer, or "link," one document to another. Hypertext linkage of documents is not new; the seminal formulation was Vannevar Bush's 1945 article, "As We May Think." In this visionary description of an online text and graphics system, Bush anticipated the information explosion of decades later. He did not, however, anticipate the computer revolution, relying instead on microfilm and photocells (the only "computer" in 1945 was the Army's "ENIAC," a baroque colossus of 20,000 tubes incapable of storing a program). Hypertext has been an issue of continuing interest to information theorists (Englebart, 1963; Nelson, 1980) but the past 5 years has seen a virtual explosion of hypertext-based information. This is because hypertext is the principal concept of the web; the protocol used in displaying pages is Hypertext Transfer Protocol (http) and the set of tags used in designing pages is Hypertext Markup Language (html). Html controls appearance and behavior of the web document, but more importantly it controls the structure of a scholarly display by controlling linkages among files. Where a "link" is indicated, a mouse click connects the reader to an address containing text, a program, an image (still or moving), a sound recording, or a separate web site. The reader may then retreat to the same place in the original text and continue reading. Although hypertext may be designed to lead the reader along a linear path, its distinctive strength is the capacity to offer multiple readings. As Bush anticipated, such link-based structure has enormous potential for scholarship.

It will be useful to give a cursory description of what is involved in creating web documents, and to clarify web-related terminology. A server is a computer that provides some service for other computers connected to it via a network. Our concern here is with web servers, computers that can be reached via the internet. A web page, or web site, consists of one or (usually) more computer files, residing on a server, and capable of being displayed by browser programs such as Netscape and Internet Explorer. Each file has a unique address consisting of its domain name, directory, and filename; when preceded by the name of a protocol, this address becomes a "URL" or Uniform Resource Locator. (The protocol for web documents is the familiar "http://"; others include "ftp://", "telnet://", and "Mailto:".)

Some of these files contain programs or digitized sounds or images, but the key to the web are files containing plain text. Within a plain text file there may be content (words to be displayed on the reader's screen) and html. Html consists of "tags" enclosed in <brackets>. A complete tag generally includes a start and stop command; for instance, the word "hello" is italicized by

```
<I>hello</I>
```

or centered by

```
<CENTER>hello</CENTER>.
```

Tags may contain parameters as well. For instance, text color could be changed to red by

```
<FONT COLOR=RED>hello</FONT>
```

While these tags give the browser information on formatting and on how to insert images and other elements, their structural function is to encode the links to other documents. The essential tag in constructing these links is the <a>...</a> anchor tag, which designates part of

the document -- typically an image, word or phrase -- as a pointer to a target. The pointer is displayed in a distinctive color, and mouse-clicking on it leads the browser to the target.

For example, let us say we have a description of a research protocol in a file named "PROTOCOL.HTML". The following use of html makes the word "here" a pointer to that target:

Click <a href="PROTOCOL.HTML">here</a> to see the research protocol.

Numerous primers and manuals on html are available on the web; a net search on "Introduction to HTML" returns over 5,000 hits.

### Document Structure

Inherently linear in form, the typical social science publication introduces an issue, summarizes current thinking, argues for a particular perspective, presents a study implementing that perspective, and draws conclusions. Rudimentary use is made of hypertext in such documents, principally in notes -- footnotes which are commonly read but expensive for publishers to produce, and endnotes which are easily produced but less commonly read. Embedded bibliographic citations are rudimentary hypertext as well. This highly restricted use of hypertext protects the linear character of the presentation.

The linear development of a thesis has an obvious place in scholarship, but the intellectual world in which social scientists operate and teach is anything but linear. In the main, social science articles are not so much "findings" as arguments that derive meaning and significance from how they relate to a network of scholarship. An article may rely on basic assumptions from a, disagree with the paradigm of b because of points made by c, borrow and adjust the methods of d, be tacitly influenced by e, quote f, and claim to refute the findings of g. Articles are parts of larger currents in thinking, developed in response to other currents. A knowledgeable reading may require awareness of these relationships beyond what the article can actually discuss. The limitation is not simply space; linear writing itself is poorly suited to presentation of these simultaneous networked relationships. This is hypertext's particular strength.

An article made up of a network of documents also lends itself to multiple readings. Most scholarly publications are packaged for a much narrower audience than their potential readership. An article in a professional journal may contain good material for an undergraduate course except that it is too long, too detailed, or too reliant on the reader's familiarity with professional literature; at the same time it may lack the theoretical exposition, detailed methodology, or raw data that specialists want. By allowing structuring of texts so as to give readers interactive control, web-based scholarship can be packaged for several audiences simultaneously. Structuring here refers to partition of a text into discrete files that are networked with html. (The allusion is to "structured programming" in computer science, in which a program is organized into discrete, named, procedures and functions rather than as an undifferentiated list of commands.) Structuring may be used to remove technical discussion, leaving a simpler document for students. At the same time, the ability to remove material to linked documents can allow publication of more technical detail of interest to more specialized audiences than may have been possible otherwise. A highly structured web article might have a front page containing only an overview, with links to segments that elaborate and illustrate. Good design allows the document to serve as an interactive table of contents and also as an outline to the work; such a document may be described as self-outlining. The accompanying demonstration page provides an example (<http://artsci.wustl.edu/~anthro/demo/>).

### Multimedia Considerations

The ability to incorporate color graphics is one of the most obvious features of web-based scholarship. I will not belabor the value of this except to note that it is convention, regrettable but necessary, for most social science articles to contain no imagery despite its high heuristic value and its ability to raise the interest level of scholarship. What is less obvious is the wide flexibility in how imagery may be incorporated into web-based scholarship. A document may include an embedded image ranging from thumbnail to full size, or a link to a single image or an extensive array. The flexibility is important not only in designing content but in giving the reader control over the downloading process (a small image linked to a larger version makes downloading of the larger one optional; this is helpful to users with slow connections).

The two widely supported graphics file formats are GIF and JPG. JPG (jpeg) allows compression of photographs into very small file sizes; GIF files are good for images with few colors. GIF files also can be used to create dynamic graphics (animations) which can depict complex, sequential, or spatial patterns that are often a topic of social science. These animated GIF's are relatively easy to construct, and are viewable by web browsers. I have described the use of animations in anthropology (Stone, 1997).

Web documents can also incorporate audio. Several different audio file formats can be played by browser add-in programs. Those with the greatest potential for social science are those that contain recorded sounds. WAV and AIF files are easily created on Windows and Apple systems respectively using programs normally included with the operating system (they do require a sound card and microphone). Brief recordings may be contained in small files but file size is quick to swell. Such files must be loaded (transferred from the web server to the user's computer) in their entirety before playing. In contrast, RAM-format files contain recorded sounds that are played as they are loaded. The data is read into a buffer, from where it is played and then discarded. This allows use of long recordings. An exemplary sound archive is Northwestern's Oyez Oyez Oyez (with recordings of arguments before the Supreme Court) at <http://oyez.nwu.edu/>.

An important use of audio in web scholarship is for pronunciations of foreign words, especially useful in disciplines such as anthropology which regularly traffics in exotic languages. Other uses include publishing of speech acts, interviews, animal calls, and music.

### **Data**

Since published scholarship necessarily summarizes patterns in data, it often omits information readers want: correlation coefficients appear without scatterplots, means are given without histograms, and crosstabulations lump categories the reader wants to see separated. Publishing detailed data is a generally unsatisfactory solution because of problems in transferring data from print to computer. Many questions about published analyses remain therefore unanswered, even those that could be quickly answered by reanalysis. In this way, unfortunate research conventions grow around technological limitations.

It is simple in web-based scholarship to include ASCII (plain text) datasets, in either columnar or delimited formats that can be read directly into a spreadsheet or other analysis program. Data files may be linked to articles downloaded with the "Save As" menu option on all browsers.

### **Durability and Dynamic Documents**

Conventional scholarly products are highly "durable"; the text appears in hard copy in thousands of libraries and private collections. There may be debate on what an author meant and whether the author was right, but rarely on what the author said. A web document has no

such durability. It can be easily sent to one's printer, but the document itself can be changed by anyone with access to the account under which it is stored. This presents both opportunities and problems.

The ability to revise, correct, and augment work after it has been published is something most scholars long for. Errors in wording and analysis creep into published scholarship regularly, sometimes with disastrous results. The usual corrective -- a published note in a later issue of the journal -- is ineffective because the original article obviously cannot refer to the later correction. Works can also often be improved by inclusion of new examples, data, or references to subsequent discussion and criticism. These changes are easily accomplished on the web. The visitor to a journal web site can read an article that has, since its original appearance, had an error corrected, a passage clarified, and a reference inserted to a later work.

The dynamic nature of web scholarship is also its weakness. Scholars can always change their opinion, but they should not be able to eradicate the publication of the original opinion. A solution to the problem of retaining originally published texts while also allowing subsequent changes is to make both available on a journal web site, with amended text linked back to the original (see the demonstration page). Journals may wish to allow, even encourage, authors one opportunity to submit a "second edition" of a web publication.

Durability also requires that publications be archived in multiple locations, and it seems likely that libraries will be the archivists of web scholarship stored on media such as cd-rom. The problem inherent in the storage of web documents is that they are often linked to documents residing at other addresses, which are in turn linked to documents at other addresses; since linked targets are, in a sense, part of a web document, it will be impossible to completely archive many documents. Such is the nature of a network. Journals can mitigate the problem somewhat by posting the files comprising an article within a single directory and using relative addresses rather than complete URL's. [\(2\)](#) This allows all those files to be easily copied and archived.


### **Differentiated Links in Web Scholarship**


As noted above, hyperlinks are achieved through the `<a>...</a>` anchor tag which designates part of the document as a pointer. By default, browsers use black for normal text and bright blue or underlined bright blue as the pointer color (images used as pointers are outlined in the pointer color). The pointer color may be adjusted in the web document or in the browser itself, but browsers have no way to indicate what the link is to. The reader of a web page in which the word "world" is highlighted does not know if the target is a photograph of the planet, a file that is part of the present article, a sound file of "We Are The World," or to the United Nations web site. The unpredictability (and occasional serendipity) may be a boon to web "surfing," it is a crucial problem in web scholarship. The reader of scholarship needs to know where a link is pointed before interrupting the linear reading of an article.

One solution shown on the demonstration page is use of a vocabulary of icons for classifying targets. Rather than using elements of the document as a pointer, small icons are inserted into the text to serve as differentiated pointers. Such icons need not be completely standardized -- journals have different styles and different needs -- but they should be intuitively clear, and articles should be accompanied by a key. The demonstration page uses the following differentiated pointers:

**STRUCTURAL ICONS.**

••• Ellipses: further discussion of a point.

 Foot: footnote. The foot and ellipses represent points on a continuum. The ellipses links to material more directly relevant to the argument that is removed to a separate page as part of "self-outlining"; the foot links to more peripheral material.

 Editorial delete: previous version of a publication.


## MULTIMEDIA ICONS.

 Mouth: pronunciation in a sound file.


 Eye: photograph or other image.



Globe: map.


 Spider Web: related web site.

## BIBLIOGRAPHIC ICONS.

 Book: bibliographic reference. Multiple citations will normally be collected into one page.

Dictionary: definition of a term.

## DATA ANALYSIS

 Chart: analytic graphic, table or chart.



Notebook: data. Tables and data are points on a continuum; tables in scientific publications often list relatively raw information. The normal use of the notebook icon would be for data in the "codebook and cases" format.

An alternative to embedded icons is to insert explicit labels as pointers (also shown on the demonstration page).

## Practical Aspects of Web Scholarship

An immediate practical consideration is what the advent of web-based scholarship means for print. Printed matter has several vital advantages over the computer screen, and will continue to play a major role in social science research even as the nature of professional journals change. A basic feature of web browsers is the ability to send a web page to one's printer; images are printed as well, and the reader can even adjust the font. The printed web page has all the advantages of an offprint (e.g., it is easily filed under author or topic). However, what the browser prints is an individual file, and what I have described here is a scholarly product comprising multiple linked files. The structure -- i.e. the relationship among the files as encoded in hyperlinks -- is difficult or impossible to capture to one's printer. This is the distinctive feature of hypertext documents: encoding of network relationships to which print documents are poorly suited.

The second practical consideration is the amount of work needed to produce web-based scholarship. A conventional texts can simply be posted on the web with almost no alterations, but capitalizing on the capabilities I have described demands at least some html coding. To see an example of the amount of html alterations involved in this article's demonstration page, one can open the browser's View>DocumentSource window while visiting the demonstration site. Creation of a web article is particularly time-consuming when the author is learning html and experimenting with style, and before shortcuts have been devised. However, the software market has responded quickly to the mushrooming demand for html authoring tools, and most recent word processors now include html conversion features that are relatively easy to use. The popular Netscape browser now includes an editor; the Explorer browser edits by loading the document into Word (in both cases, the document is automatically downloaded to the user's computer for editing, after which it must be uploaded again).

But in considering the work requirements of web scholarship, we should not view the coding and uploading simply as added work; it is actually a replacement for the work of typesetting, printing, and mailing for which journals are now responsible. It must also be remembered that a structured web text can in effect serve as a multiple articles, suited for multiple audiences. Journals will develop procedures for working with authors in development of web-based scholarly products; most major journals will likely have an "internet editor" on the editorial staff. The skills required for such work are commonplace on campuses today, and they will be moreso in the future.

In the end, the most dramatic changes in the logistics of producing scholarship will likely be the direct result of freeing journals from the time demands and costs of print publication. Freed from the physical confines of producing individual issues, journals will have new latitude in publication policies. The possibility of very speedy publication will improve the timeliness of scholarship and provide opportunities for followup discussion and debate. The selectivity of journals, the sina qua non that separates scholarship from soapbox, is now maintained in part by page limitations resulting from publication costs. There is no such limitation on web scholarship; once server connection costs are paid, the costs associated with posting articles are negligible. The web-based journal's level of exclusivity therefore becomes entirely a matter of editorial policy, editors' time, and the traffic in scholarship.

My concern here is to highlight the coming change and to explore how the new medium can be used to create different kinds of scholarship, not to address the myriad larger questions raised by this transformation. However, it is worth noting that the advent of web-based scholarship portends changes in the public profile of social science (serious research may become more accessible and widely read), in the mechanics of teaching (course web pages with URL's of assigned articles can replace the piles of photocopies on reserve), in the finances of journals (production costs will be dramatically reduced), and in the procedures for evaluating scholars' output.

---

## Notes

I am grateful to B.J. Johnston of the Washington Univ. Libraries and to M.P. Stone for comments and discussion.

## References

Bush, V. (1945) "As we may think." Atlantic Monthly July 1945: 101-108.

Englebart, T. (1963) A conceptual framework for the augmentation of man's intellect. In Vistas in information handling, Vol I, ed P.W. Howerton and D.C. Weeks. London:Sparten Books.

Nelson, T.H. (1980) Replacing the printed word: A complete literary system. IFIP Proceedings Oct 1980, pp. 1013-1023.

Stone, G. 1997 Animated images: A new tool for web-based anthropology. Cultural Anthropology Methods 9(1):15-16 and

<http://artsci.wustl.edu/~anthro/research/CAMPAPER.HTM>.

### **Biographical Sketch**

Glenn Davis Stone is an ecological anthropologist. His research in recent years has focused on indigenous agricultural systems in Africa, especially their social and spatial aspects. He is Associate Professor in the Dept. of Anthropology at Washington University, St. Louis MO 63130; [stone@artsci.wustl.edu](mailto:stone@artsci.wustl.edu), 314-935-5239 voice, 314-935-8535 fax.

### **Footnotes**

1. For example, the Johns Hopkins Univ. Press offers 43 full-text journals at <http://muse.jhu.edu/journals/>; the Committee on Institutional Cooperation's electronic journal collection offers 146 online journals at <http://ejournals.cic.net/index.html>. The "NewJour" mailing list, devoted entirely to announcements of new online journals, lists several dozen each week.

2. A relative URL refers to a file on the same server as itself, making it unnecessary to include the entire URL.