Dissemination of University Research via the WEB

Vicki L. Willis, James K. Nelson, Jr., and Scott D. Schiff

Clemson University Department of Civil Engineering

Abstract

With little question, electronic dissemination of information has emerged as an acceptable means of information distribution. Considerable information has been and is being placed on the World Wide Web (WEB). The type of information that can be placed there could feasibly include reports and technical papers about research conducted by university faculty and graduate students. This paper discusses possible formats for document distribution, a possible framework for disseminating such information on the WEB, and the implications of such dissemination.

Introduction

Yearly, considerable state-of-the-art research is conducted at the nation's colleges and universities. Much of this research is documented in the form of graduate student theses and dissertations, and technical reports prepared by the faculty. Theses and dissertations are placed in university libraries where they are not readily accessible to the general public and the information contained in them is not necessarily abstracted for easy reference. Further, because of the delay between preparation and publication of technical reports, the information presented may no longer be state-of-the-art.

Rapid dissemination of research findings into the open literature is beneficial for both the research community and practicing engineers. In *Civil Engineering*, Najib Abboud states that "You may be looking for software that is not necessarily mainstream, or that is a bit out of the mainstream. If you can locate a 'breaking'-type of technology that is still at the university stage, you can actually talk with the professor to see if it does what you need. In this area, the Web has the potential to tremendously speed up the transfer of technology from researcher to practicing engineer." (Salwen, 1996). Michael Hough (1997) of A/E/C Systems indicates that the Internet has the potential to shrink the design world.

By presenting theses and reports on the Internet after they have been prepared, they can be available for use in a matter of days instead of months. Through relatively minor changes in the current way theses, reports, and similar scholarly publications are prepared, they can be placed on the WEB with little additional effort. The documents would then be available through virtually any computer with a WEB browser and access to the Internet. The receiving computer does not need to have the same software available as that used to prepare the original manuscript, and it does not need to use the same operating system.

Another advantage to placing research findings on the web is that better presentation of the topic through embedded audio and video clips is possible. Through video clips, and perhaps animation sequences, failures or other phenomenon can be demonstrated rather than just being described. This will lead to a more comprehensive understanding by the reader.

Currently, three methods are viable for preparing technical documents for dissemination on the WEB. These are:

1. Printing the manuscript in a portable document format,
2. Preparing the manuscript as an HTML document, and
3. Preparing the manuscript as a Windows help style document.

An overview of the methods to prepare manuscripts in these formats, as well as the advantages and disadvantages of each method, is discussed in this paper. In the context of this paper, a WEB document is any document that can be downloaded from the Internet or viewed using a WEB browser. This paper, in each

of the formats discussed, is available on the WEB at (http://tiger2.ces.clemson.edu/webdoc).

Structure of the Electronic Document

Because theses and reports tend to be very long documents, a single electronic document for a thesis or report has the potential to be a very large file. There are two disadvantages to large files on the WEB. First, they require a significant amount of time to acquire. With the increasing traffic on the WEB, the time required quickly becomes unacceptable, especially if a modem connection through an Internet Service Provider (ISP) is being used. Secondly, navigating through a large document is unwieldy. Going to, or returning to, specific locations within a document is difficult unless internal links are provided or can be provided.

For these reasons, theses and reports distributed via the WEB should be divided into smaller logical units. A suggested structure for electronic theses and reports is shown in Figure 1. It should be noted that the organizational structure of an electronic document is not different than current paper documents.

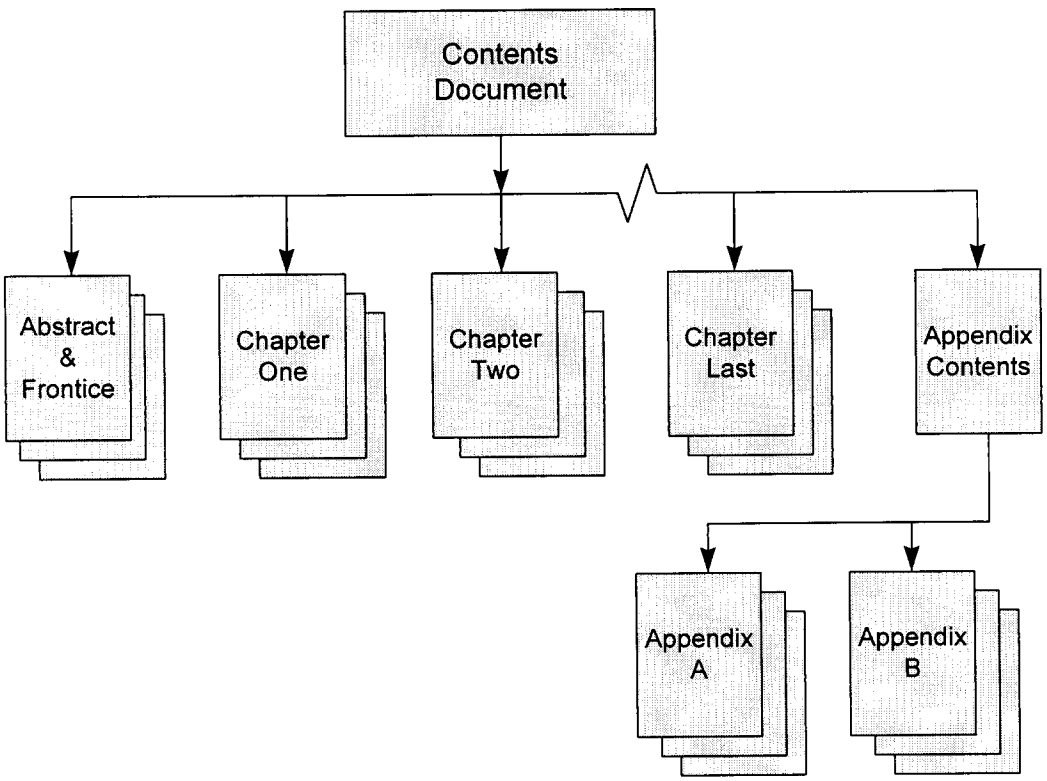

Figure 1–Hyperlink Structure of the Electronic Thesis

There are many similar variations of this structure. The contents document is a document that links to the sub documents. For distribution as an HTML or a PDF document, the contents document is a typical HTML document. When using the HELP format, this document is the HELP contents document. Each chapter or significant section is another data file to which a hyperlink is provided. As discussed later, hyperlinks may be provided within sections to further facilitate navigation.

The *Abstract & Frontice* document is a relatively small document containing the abstract and the listing of the contents. An interested reader can quickly download and view this document to decide if it is relevant to his or her pursuits. If it is deemed relevant, as many or as few of the remaining units can be downloaded as necessary. Utilizing this scheme, or a similar scheme, yields smaller files that are more easily manipulated.

Portable Document Format

The simplest method for preparing documents for distribution on the WEB is through the use of a portable document format (PDF). PDF's are prepared by printing the document from a word processor (such as Microsoft Word or Corel WordPerfect) to the portable document printer. Drivers for portable document printers are available from third party software providers such as Adobe, through its Acrobat software package, (http://www.adobe.com) and Tumbleweed, through its Envoy software package. (http://www.twcorp.com). Both companies require licensing of the print drivers, but have viewers that are freely distributed, and include "plug-ins" for common WEB browsers. The "plug-ins" enable the document to be viewed without requiring user intervention with another application—the WEB browser launches the necessary interface.

Documents prepared as PDF's appear virtually the same as the original document and are used in a similar manner — the reader must page through the document to find the needed information. The document file is transferred to the client computer and is viewed there. A paper copy can be made for future reference.

Planning and preparation for a document to be presented as a PDF document is essentially the same as a document for normal dissemination. The only difference, as indicated above, is that the final document is printed to a PDF writer for electronic distribution as opposed to a paper printer for normal distribution. This statement is based on the assumption, given the capability of current word processing and desktop publishing software, that illustrations, tables, and similar information are normally imbedded in the document and not collated afterwards.

The only additional planning required for preparing a document in the PDF format, compared with preparing documents in a normal manner, is the capabilities of the PDF writer used. Some fonts or special formats may not transfer directly. This is an insignificant drawback because there are a myriad of similar fonts available. If one font is incompatible, a similar, compatible font should be available. As such, the appearance of the document remains unchanged.

The HTML Document Format

An alternative to preparing a document in a PDF format, is preparation of a document in the HTML format. This format permits the document to be viewed directly without need for a "plug-in" or external viewer. Yet, like a PDF document, it can be downloaded and saved on disk for later viewing. Also like PDF documents, an HTML document can be prepared using standard word processors and then converted to HTML format using readily available utilities, such as the intrinsic utilities Microsoft Word and Corel WordPerfect, for example.

The structure of an HTML document would fundamentally be the same as that for a PDF report. While using the same breakdown for units as PDF, the HTML format allows for "dynamic" viewing and easy access to the sections of the report. The table of contents can become links to sections and through the use of bookmarks can be a direct link to the subsections. This latter capability is not as direct with PDF documents. It facilitates readability and use because information within the document can be reached more easily.

There is more preparation involved with this form of dissemination because the document has to be prepared in the HTML format. This requires the placing of "tags" to obtain the necessary formatting. This is not a major hurdle because of the available converters for most word processing software. These converters are not "perfect", though, and some modification after conversion is usually required.

The author of the document will have to manually provide the links to sections within the chapters through bookmarks. This step is not necessary in PDF documents because internal links will probably not be used. Although providing links is not difficult with current HTML editors, it is another step in the process that needs to be checked. The increase in ease of use, however, makes this minimal additional effort very beneficial for the end user.

Windows Help Style Document

The HELP format requires more effort, but not significantly more if planning and preparation start at the beginning. The author must be in the right "thinking mode" as the document is planned and the information to be included needs to be collected from the beginning. If the author waits until the end to begin preparation of the manuscript as a HELP document, this form could be overwhelming.

Text and illustrations to be presented as a HELP document are prepared using the rich text format. Such a format is available in most word processors. Internal formatting and "styles" are used to control the appearance of the end product. Internal links are provided through use of special footnotes that must be provided by the author. The basic text, though, is the same as with any of the other formats.

The resulting document is "compiled" using a HELP compiler to form the final help module. Because this is a compiled document, the most notable disadvantage of this type of document is platform dependency. The dependency arises because it is developed for implementation in a Windows® operating environment. It cannot be viewed through a WEB browser; it must be downloaded to a computer with Microsoft Windows® for viewing.

The HELP format, though, provides some advantages for the user. The help application in Windows® provides extensive search and linking capability that are either not available or are not as usable with the other formats. Complete word searching can be performed if the user wishes. Through use of the special footnotes, the author can provide the keywords used to perform a search and related sections and topics in the document can be grouped together in the automatic indexing process.

Dissemination Structure

Because of the potentially large number of documents that can be distributed electronically within a university, or even within a department, careful consideration must be given to the structure of the WEB site through which these documents are distributed. Equally important is the security of the site so that "infected" documents and bogus information are not proliferated. Maintenance and frequent archiving of the site is important so that data is not lost.

The logical location for control of the WEB site used to distribute theses and dissertations is the library because the library currently manages such documents. They can be electronically cataloged on a HTML based home page in a manner similar to the way current indexing is conducted. Cataloging on a home page, though, should not replace conventional cataloging, as not all individuals conducting a search will use the WEB. The WEB will likely be a secondary search source because not all holdings of the library are indexed on the WEB. Departments within a university can hyperlink to this directory from their individual home page.

Research reports likely will not be maintained by the library because they traditionally have not been maintained by the library and there is not significant reason to change. These documents can be catalogued and stored on departmental WEB servers. Currently research reports are stored in departmental and perhaps college libraries. Placing reports on departmental WEB servers, then, is consistent with current practice. By using the WEB, though, availability of the information will be much greater. Of course, propriety research not intended for the public domain would not be catalogued in this manner.

Style guidelines for preparation of the electronic document should be prepared for theses and dissertations, as well as for research reports, to provide consistency across the institution or at least across the department. Mostly such guidelines for paper documents are already in existence and review procedures could be the same.

The duration of documents in electronic form needs to be established. The duration will depend on several factors, including:

1. The number and size of documents prepared each year,

2. Availability and practical limitations of storage space, and

3. Reasonable period for ready availability.

Indefinite availability of documents in electronic form is not reasonable, just as indefinite availability in printed form is not reasonable. Unlimited electronic storage space is not likely. Further, the current relevance of documents becomes less useful and easy reference is less of a factor. As such, at some time electronic documents should be removed. After the electronic document is removed, paper and microfiche archival copies would still be maintained through current distribution methods.

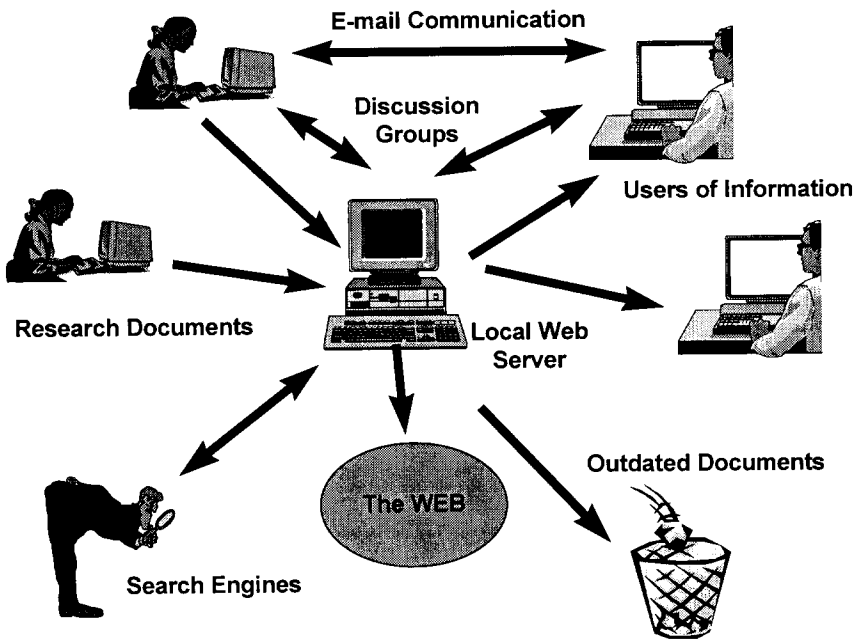

Figure 2—Interaction Between Users and Producers of WEB-Based Documents

Obstacles to Overcome

The mechanics of preparing theses and reports for electronic distribution on the WEB is relatively simple and the necessary tools are readily available. There are obstacles, however, that must be overcome before institutional use of this delivery mechanism becomes a reality.

The primary obstacle to disseminating theses and reports in the manner discussed in this paper is one of philosophy. Traditionally, theses and dissertations are prepared by the student in printed form, are reviewed for acceptability in regards to content and format, and bound copies are then archived in the library. The electronic thesis may be viewed by some as a departure from this process. In fact it is not a departure from the process of documentation and review, but only a departure from the traditional form in which the document is disseminated. The difference is that the documentation will now be prepared in a format conducive to electronic distribution and the greater availability of the research thereby afforded. For archival purposes, a bound copy, or perhaps a microfiche copy, of the prepared text should probably still be maintained by a library.

The information presented in this paper is directly applicable to the publication of technical papers on the WEB as well. By doing so, the availability of the information would be significantly enhanced. However, because the publication of technical papers in peer reviewed journals is so closely tied to faculty promotion and tenure, methods for peer review of WEB publication is essential. Such review and publication could continue to be handled by the technical societies. The only difference would be subscription access on the WEB versus a subscription to periodicals. The abstract for these publications as well as a link to the archival electronic journal could be placed in the publications section of technical society, university, and departmental WEB pages.

Faculty may be reticence to accept electronic theses and other electronic publications because they are perceived as a change from traditional rigor and publication channels, such as libraries and society journals. As indicated above, the electronic document is a departure from traditional dissemination only. After having information more readily available, faculty are likely to embrace the change because their work can be referenced more and they will be able to obtain information more easily.

Professional and User Responsibility

Professional responsibility for placing and using information on the WEB emanates from two sources. These are the user and the author of the available information. These responsibilities should be self-evident. However, they are important considerations and should be discussed.

As with any information used in engineering analysis and design, the design engineer has the responsibility to determine the accuracy of the information and the appropriateness of that information to the task at hand. This is not a new responsibility. Rather, it is one that has existed since research findings were published and computing began to be used for engineering calculations. The responsibility is being compounded now because of the quantity of information available from a proliferation of sources.

The individual placing information on the WEB, or preparing a document to be placed on the WEB, has the responsibility to provide information that is accurate and as timely as possible. With student theses and dissertations, this is not much of a problem, or at least should not be a problem, because the student's academic advisory committee provides oversight on the document. When dealing with research reports, and particularly when dealing with technical papers, such a review does not exist, except for that provided by the refereeing process.

When documents are in electronic form, as are all documents on the WEB, plagiarism is easier than it has been at any time in the past. An individual simply needs to "copy" information from the WEB browser and "paste" it into the document they are preparing. There is nothing fundamentally wrong with this, provided the original author is given credit.

Advantages and Disadvantages

Electronic dissemination of documents electronically through the WEB has advantages when publishing research findings. There are, however, some disadvantages. As stated earlier, one of the primary advantages to publishing through the WEB is the ability to rapidly disseminate research findings to the research community and perhaps even to the general public. However, one of the most significant impacts could be the ability to create interaction with other research investigators and users of the information. This interaction could be a result of ordinary exchanges of re-

search data or computer programs placed out on a WEB server. More interesting for readers could be the on-line discussion of research findings through discussion groups or "chat" rooms. These types of discussions, in time, may serve the same function as traditional technical reviews of journal papers and may result in refinement and republication of WEB documents.

One of the possible problems associated with this form of distribution could be the dissemination of poor-quality information. In the effort to quickly disseminate information, the lack of rigor during the initial review of documents to be published may result in documents being published before they are really ready to be published. In order for WEB dissemination to be effective, a rating system, perhaps based on the level of review, could be established and posted with the document.

Clearly the above disadvantage can be overcome through responsible publication of WEB documents. Once quality information is made available though the WEB, many will find the WEB to be another valuable source of information, not replacing the traditional library, but simply a better means to gather state-of-the-art information. Since most users rely upon search engines to find information out on the WEB, it will be extremely critical for authors to develop a keyword list for each published document that will allow the search engines to find the information needed by a user.

Looking to the Future

During the past 18 months Internet programming languages such as JAVA and ActiveX have been introduced and are gaining significant application for animated and interactive WEB documents. At this time, question exists whether these technologies offer benefits for dissemination of research whose end product is not computer based. In these types of research projects, the information presented is read and evaluated by the user; there usually is not an interactive element associated with the research findings.

JAVA and ActiveX have direct application when the end product of the research is an interactive training, tutorial, or similar document. For example, the purpose of a graduate research project could be development of tutorial software for use by sophomores taking a course in mechanics of materials. Such a tutorial could be developed using either JAVA or ActiveX instead of some other development language such as C++, PASCAL, or FORTRAN. If the tutorial was developed using JAVA or ActiveX, it could be easily disseminated on the WEB and thereby realize the greatest availability.

Conclusions

Electronic dissemination of information has emerged as an acceptable means of information distribution. Considerable information has been and is being placed on the World Wide Web. The type of information that can be placed there could feasibly include reports and technical papers about research conducted by university faculty and graduate students.

Use of the methods discussed in this paper provides a consistent format for a "user friendly" WEB document that can be platform independent. Publishing on the WEB greatly improves accessibility. To be a good and useful source of information, though, students and faculty need to begin publishing on the WEB in earnest. As more information is available, more information will become available.

The methods used are likely to change in the near future. These changes should only enhance the capability available today. Further, the changes are likely to be backward compatible so documents published today will still be usable when change occurs. Tools for JAVA and ActiveX applications are becoming more readily available but are still in their infancy. In the future, they may become viable forms of publication for theses and research reports.

References

Hough, Michael (1997) "Editorial: Computing Trends for Civil Engineers," *Journal of Computing in Civil Engineering*, American Society of Civil Engineers, Volume II, Number 4, p 216.

Matteson, Brenda L., Project Editor (1995) *Microsoft Windows 95 Help Authoring Kit*, Microsoft Press, Redmond, Washington.

Salwen, Peter, (1996) "Sticking with the WEB," *Civil Engineering*, American Society of Civil Engineers, Volume 66, Number 6, June, pp. 36-41.

JAMES K. NELSON, JR.

Dr. Nelson graduated from the University of Dayton in 1974 with a Bachelor of Civil Engineering degree. He received a Master of Science degree and a Ph.D. from the University of Houston in 1976 and 1983, respectively. He taught at Texas A&M University for ten years before joining the Civil Engineering Faculty at Clemson University in 1989. Currently, Dr. Nelson is a Professor of Civil Engineering and Program Director of the Clemson University Graduate Engineering Programs at The Citadel. He is a registered Professional Engineer.

Dr. Nelson has co-authored an undergraduate structural analysis textbook and has developed software for three structural analysis and design textbooks. Since joining Clemson University, Dr. Nelson has taught graduate courses in structural engineering and numerical methods. For the past five years these courses have been taught on television as part of the distance education program at Clemson University.

SCOTT D. SCHIFF

Dr. Schiff obtained a Bachelor of Science degree from the University of Cincinnati in 1982 and his Master of Science and Ph.D. degrees from the University of Illinois at Urbana-Champaign in 1984 and 1988, respectively. He joined the faculty at Clemson University as an Assistant Professor of Civil Engineering after a one-year stay as a Visiting Assistant Professor of Civil Engineering at the University of Illinois. Currently, Dr. Schiff is an Associate Professor of Civil Engineering and serves as the Graduate Program Coordinator for the Department.

Dr. Schiff has taught classes in the structural engineering area at both the undergraduate and graduate level as well as continuing engineering education courses in the area of earthquake engineering. His primary research focus is in the mitigation of damage from severe wind storms and earthquakes to low-rise buildings.

VICKI L. WILLIS

Ms. Willis received a Bachelor of Science in Civil Engineering from Clemson University in May of 1996. She is currently pursuing her Masters in Civil Engineering with a focus on structural engineering. Graduation is expected in December 1997. For her graduate study, she is working with the South Carolina Sea Grant Consortium and the City of Charleston to provide a community sustainability center which will be housed in and feature a retrofitted historic structure. Ms. Willis' thesis is entitled *A Structural Evaluation and Retrofit Methodology for Historic Houses in Charleston, SC.*