Options for a MODIS Document Distribution and Information Management System

(Draft version 5)
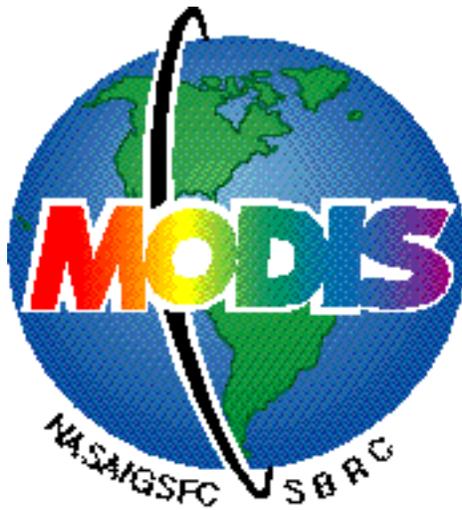
by

David Herring

Science Systems and Applications, Inc.

November 15, 1995

1.0 INTRODUCTION

Exchanging electronic files in a manner that is both platform- and software-independent continues to be a problem for the MODIS Team. Here the term "files" includes text (frequently with formulae embedded), presentations, images, and graphics—often some or all of the aforementioned file types are integrated into a single document. Sometimes, however, team members either do not, or for various reasons cannot, integrate graphics and images into the text of their documents, so they distribute multiple files. For instance, some ATBDs were submitted as text files with a dozen or so image and graphic files included as attachments in other formats. It seems that almost no two MODIS Team members go about creating documents the same way. There appears to be an almost even distribution of team members using Macintosh, PC, and UNIX platforms; and even those operating on the same platform often use different software tools. So, naturally, problems arise when some team members try to distribute their documents electronically to other team members.

1.1 Scope and Purpose

The purpose of this report is to provide a review of some noteworthy archive and document distribution systems already in existence--whether within the MODIS Team, within EOS, within NASA at large, or at other agencies in general--and to recommend ways in which MAST may modify or add to MODARCH to enhance its document archive and distribution capabilities. Discussion is provided on what software tools are available (whether freely or commercially) and how implementing them will help streamline MODIS document development and distribution. As cost and compatibility with MODARCH are major drivers, the scope of the review of new technologies contained in this report is somewhat limited.

Concurrent to, and completely independent of, the writing of this document, the NASA Center Information Officers (CIOs) established guidelines that stipulate the minimum office automation software suite required to support interoperability. The CIOs objective is set interface and product standards for components of the software suite operating on PC, Macintosh, or UNIX platforms used by all NASA employees--both civil servants and on-site contractors. Ultimately, it is the responsibility of the CIO

at each NASA campus to ensure that all NASA employees at that site have access to an interoperable workstation that is equipped with a minimum software suite for performing the basic essential functions of word processing, building spreadsheets, creating presentations, sending electronic mail, calendar scheduling, and basic Internet activities (FTP and browsing capabilities). The CIOs are tasked with establishing reporting metrics for determining overall NASA computer interoperability. The CIOs will be required to report the number of NASA employees at their respective sites that do not have access to interoperable workstations. Chapter 2 addresses the CIO guidelines further.

Chapter 3 is a brief survey of archive and distribution systems available at GSFC, as well as at other NASA centers. Chapter 4 offers recommendations as to what applications the MODIS Team should use to create documents. Chapter 5 of this report provides a brief discussion of commercial and public domain software applications available for document distribution. Chapter 6 describes how documents may best be shared among team members, and/or submitted to MODARCH, if the recommendations given in Chapter 5 are followed. Chapter 7 offers insight into how both MAST and MODIS Team members may work together to evolve MODARCH from an archive and distribution system, to a more robust information management system. Chapter 8 provides brief tutorials or background information on new subjects introduced in this document.

2.0 NASA MINIMUM OFFICE AUTOMATION SOFTWARE SUITE

The following are standards for the components of what NASA considers to be the *minimum* office automation suite. (UNIX has not yet been addressed by the CIOs).

Table 1 - **Interface and Product Standards**

Component	Interface Standard	Product Standard
Word Processing	WIN Word 2.0 File	MS Word, WordPerfect
Spreadsheet	Excel 4.0 File	MS Excel, Lotus 123
Presentation	PowerPoint 3.0 File	MS PowerPoint
Electronic Mail	TBD	TBD
Calendar/Scheduling	TBD	TBD
Basic Internet Suite	*See text below	*See text below

Table 1 lists the software considered "minimum" standards; there are, however, target standards which will likely result in an upgrade of the word processing interface standard to WIN Word 6.0; the spreadsheet interface standard will become Excel 5.0; and the presentation interface standard will become PowerPoint 4.0. Netscape version 1.2 and Adobe Acrobat Reader version 2.1 are considered mandatory software as part of the basic Internet suite for PC and Macintosh users.

For electronic mail, the MODIS Administrative Support Team (MAST) recommends using any application that supports MIME encoding of attachments. MIME encoding enables attachments to be sent across platforms. Goddard Space Flight Center has a site license for Eudora, which supports MIME encoding, for PC and Mac platforms.

For calendar/scheduling, MAST recommends using On Call's Meeting Maker, which is available for PC, Mac, and UNIX platforms.

3.0 DOCUMENT CREATION

The MODIS Team uses a variety of word processors to create documents. Microsoft Word (for Macintosh and PC-windows) is perhaps the most commonly used; some others are WordPerfect for Mac, PC, and UNIX; and T_EX, L_AT_EX, and troff for UNIX.

Recent versions of MS Word and WordPerfect have built-in filters enabling users to open, view, and print documents between MS Word and WordPerfect, regardless of the platform. Because MS Word is already used by the majority of team members, and because it contains filters for opening a wide range of document formats, it was selected by MAST as the tool of choice for creating word processed documents. It is also now the NASA standard format for distribution. Yet, some team members still prefer T_EX, L_AT_EX, or troff because they are already familiar with them, and historically these were among the few applications that allowed composition of complex formulae along with text. Also, the team members using these applications typically operate in a UNIX environment; MS Word is not available for UNIX. However, WordPerfect *is* available for UNIX, so it is a viable substitute.

The newer versions of MS Word and WordPerfect contain tools for writing and editing formulae. Additionally, MS Word and WordPerfect allow users to create or import images and graphics. Moreover, there are commercially available filters--both built into Word and WordPerfect, and third party applications--that allow platform-independent conversion of documents between Word and WordPerfect. So, MAST recommends that MODIS Team members adopt MS Word or WordPerfect as "standards" for creating documents, regardless of their respective desktop computer platforms. This solution affords team members the capacity to create documents in which they can incorporate files from a wide variety of formats; and it allows other team members to easily open, review, co-author, edit, and print these documents in their original format.

Please note, however, that the conversion filters between MS Word and WordPerfect are not without flaws. Someone using MS Word for Macintosh to open a document created in WordPerfect for Windows may notice little "quirks" in some of the formatting, especially if there are complex formulae included. For instance, some symbols may not parse exactly as they did in the original version. The best way to avoid these "quirks" when exchanging

documents between platforms and different word processors appears to be saving documents in Interchange Format (RTF, also known as Rich Text Format). RTF is an option whenever you select the "Save As ..." command. Another option is to print PDF copies of the document.

3.1 Creating Presentations

The MODIS Team's most widely-used software for creating presentations is probably Microsoft PowerPoint. PowerPoint is a flexible, well-developed application: it allows for the use of multiple typefaces, offers a wide variety of color and shading tools, and enables importation of graphics (in TIFF, PICT, or EPS formats). Moreover, PowerPoint allows users to save documents in "RTF Outline" format, which is compatible with both MS Word and WordPerfect in addition to the PowerPoint native format. Therefore, even users who don't use PowerPoint can still be given electronic access to information in PowerPoint documents.

There are other applications similar to PowerPoint (e.g., Aldus Persuasion). However, PowerPoint was adopted as the NASA standard format for presentation distribution. MAST recommends that MODIS Team Members using presentation software other than PowerPoint should ensure that their software tool can save documents in RTF. Or, to better preserve the appearance of the presentation, print from PowerPoint to PDF and distribute the latter copy.

3.1.1 Meeting Handouts and Presentations

For all future MODIS Meetings, MAST recommends that presenters, or anyone wishing to distribute handouts, bring electronic copies of their material to the meeting--preferably in PDF. In this way, these materials can be electronically submitted to MAST--quickly and easily--both for immediate electronic distribution to meeting attendees and for inclusion in MODARCH.

To facilitate electronic distribution of presentations, MAST is also prepared to support fully electronic display of presentations at MODIS Team Meetings. MAST procured a liquid crystal display (LCD) overhead projection panel which fits on overhead projectors and connects to either PC or Mac computers, so that whatever is displayed on the computer screen is also projected for everyone else in the room to see. From now on, at MODIS Meetings, MAST will have both PC and Mac laptops available, with Microsoft

PowerPoint (as well as MS Word, WordPerfect, and Adobe's Acrobat Exchange) installed on each, for electronically displaying any presentations.

If you plan to use presentation software other than one of the aforementioned applications (e.g., Aldus Persuasion), please submit an electronic copy of your presentation to the MODARCH system administrator well in advance so that it may be converted to a more accessible format (such as PDF) in time for your presentation.

3.2 Creating Images and Graphics

MAST uses a wide variety of graphics and illustration software tools, including Adobe Photoshop, Adobe Illustrator, Aldus Freehand, Aldus PageMaker, GraphicConverter, MacDraw Pro, and Microsoft Excel. Thus, we are capable of opening, editing, and/or converting graphics and illustrations in a wide variety of formats.

Encapsulated PostScript (EPS) seems to be one of the easiest formats for a graphic to be embedded into MS Word, WordPerfect, *and* PowerPoint files. Whenever creating a graphic for inclusion in a document or presentation, be sure the software tool you use has the ability to save the file as an EPS.

However, whenever creating a graphic specifically for viewing on the WWW, save the file as either GIF or JPEG, as these files are compatible with WWW protocol and EPS is not.

Adobe Acrobat Exchange or Distiller both provide the means for compiling text and graphics into a single PDF document. These are discussed further in chapter 5.

4.0 NASA ARCHIVE AND DISTRIBUTION SYSTEMS

4.1 MODIS Document Archive (MODARCH)

MAST is responsible for maintaining an archive of all MODIS documents. MAST operates MODARCH, which enables full electronic, platform-independent search and retrieval of all information stored there. For simplicity, "MODARCH" refers to the entire assemblage of hardware and software that make up the system.

The MODARCH system is comprised of a Sun SparcStation 10, about 20 Gbytes of hard disk storage space, a Fujitsu scanner, and a PC to run the Optical Character Recognition (OCR) process for scanned documents. MODARCH uses Excalibur Technologies' Electronic File System (EFS) version 3.6 software in carrying out its archive function--to ingest, index, store, and retrieve all MODIS documentation. Using EFS Client software (**note:** currently available only for Mac and PC; until the UNIX client is released with the release of EFS v4.0, UNIX users must access MODARCH via X-windows), users may search for and retrieve documents from MODARCH via the Internet. For a host of reasons, EFS is the best commercially available electronic imaging system to meet the team's diverse *archive* needs.

Refer to the tutorial at the end of this document for detailed instructions on how to print or launch other applications from the EFS Client. For a more detailed explanation of how EFS works, please refer to the *MODARCH User's Guide*, available electronically on the World Wide Web (WWW) at `http://modarch.gsfc.nasa.gov/MODARCH/modarch.html`; or contact Michael Heney at (301) 286-4044 to receive a hardcopy version.

4.1.1 MODARCH as FTP Site

To facilitate document distribution and complement MAST's archive function, space on the MODARCH system was allocated to serve as an anonymous File Transfer Protocol (FTP) site for the entire MODIS Team. Team members may freely upload or download files to or from the MODARCH FTP site in the "/pub/incoming" directory.

4.1.2 MODARCH <-> World Wide Web Interface

Michael Heney, MODARCH System Administrator, wrote an interface utility enabling users to browse the entire contents of the EFS archive via the

WWW (using Mosaic or Netscape). Currently, users cannot do more than navigate through the EFS Fileroom hierarchy down to the individual document level to see what's in each File Cabinet, File Drawer, and Folder. However, this is an important first step toward a more robust interface that will eventually allow users to access documents stored in EFS via the WWW.

Please review the MODARCH < - > WWW Interface Test Page at this URL--`http://modarch.gsfc.nasa.gov/MODARCH/EOS_FILEROOM/EOS_fileroom.html`.

4.2 EOS Project Science Office's Role in Document Distribution

The EOS Project Science Office (PSO) established a public WWW site enabling anyone to access certain documents about the Earth Observing System via the Internet. As Algorithm Theoretical Basis Documents (ATBDs) are being developed for every EOS instrument data product, the PSO took on the responsibility of collecting and distributing those documents. ATBDs provide the theoretical basis--both the physical theory and the mathematical procedures, as well as possible assumptions being applied--for the calculations that have to be made to convert the radiances received by the satellite sensors to geophysical quantities.

Guidelines were established for the formatting of ATBDs, as well as submission protocols and procedures. According to the PSO guidelines, ATBDs are to be submitted in Microsoft Word (for MS-DOS, MS Windows, or Macintosh), WordPerfect, ASCII, Interchange Format (RTF), LaTeX, FrameMaker, or PDF (Portable Document Format). Any graphics that accompany an ATBD, but that were not integrated into a single document in one of the aforementioned formats, are to be formatted as GIF, TIFF, EPS, or PICT files. Additionally, PostScript versions of all ATBD text and graphic files are to be made available in an anonymous FTP account on the submitter's computer.

The PSO established an anonymous FTP account on its own UNIX server--spso.gsfc.nasa.gov. PostScript and PDF copies of all ATBDs are available there too. Anyone can navigate to the PSO server via Mosaic or Netscape and, following the hypertext links, retrieve the ATBDs. MODIS ATBDs are available there, (as well as in MODARCH) and may be retrieved in either PostScript or PDF formats. The WWW URL is

`http://spso.gsfc.nasa.gov/atbd/pg1.html`.

4.3 ECS Data Handling System

The EOS Data and Information System (EOSDIS) Project established a document distribution accessible via the WWW, called the ECS Data Handling System (EDHS). The EDHS is the on-line distribution and storage system for documents about the EOSDIS Core System (ECS). It stores files in four different formats--plain ASCII, PostScript, RTF, and PDF--on a UNIX server. The EDHS URL is `http://edhs1.gsfc.nasa.gov/`. Additionally, EDHS provides three methods for conducting document searches: (1) a **Quick Search** that searches on document numbers, titles, or keywords; (2) **Full Text** searches; and (3) a **Document Catalog** that provides a numeric listing of all documents available there.

The EDHS is easy to use, provides several search mechanisms, as well as four different formats in which to view and/or print documents, and is platform independent. Within Mosaic or Netscape, freely available software, anyone may view or print the ASCII versions of documents within the EDHS. Or, users may download the PostScript, PDF, or RTF versions. This system gives users the best of all worlds.

However, from the perspective of a system administrator, the EDHS is a costly endeavor, relative to the operation of MODARCH. Currently, depending upon document load, the EDHS requires up to six people to operate, whereas MODARCH employs one-and-a-half people.

4.4 CERES On-line Documentation

The CERES Team has established a WWW site at NASA Langley Research Center (LaRC) for distributing CERES documents. The URL is `http://asd-www.larc.nasa.gov/ceres/docs.html`. Documents are available from this site in PostScript format only.

5.0 VIEWING DOCUMENTS

Documents (or files) created according to the prescriptions given in Chapter 3 of this report may be viewed onscreen or printed, independent of platform, provided that the same tools (MS Word, WordPerfect, or PowerPoint) also discussed in those sections are also used to do the viewing or printing. Third-party applications need not be used.

Moreover, documents created in MS Word and WordPerfect may be electronically entered into MODARCH'S EFS Archive in those formats, with any formulae or graphics intact, and then retrieved again later from the archive and viewed in their native formats (MS Word or WordPerfect). Once the document is opened into MS Word or WordPerfect, users may then read, edit, or print as they normally would. **[note:** any changes made affect only the file(s) opened on the user's computer, not the version stored in the archive.]** Users submitting documents to MODARCH electronically in either MS Word or WordPerfect help eliminate any viewing or printing problems they may have experienced with the EFS Client software in the past. (Section 8.2 of this report gives details on how you may view EFS text files in either MS Word or WordPerfect.) Once the document is in MODARCH, the submitter may notify the concerned persons via e-mail, and then those persons may retrieve copies of the document for themselves, in either MS Word or WordPerfect--whichever they prefer.

MAST recognizes that some MODIS Team Members may still opt to use T_EX, L_AT_EX, Troff or another word processor other than MS Word or WordPerfect. Or, they may interact with other "outside" persons who do not use MS Word or WordPerfect. For those cases, MAST has evaluated and procured another solution--Adobe's Acrobat modules. In the long run, it is expected that the dividing lines between EFS, the WWW, and Acrobat will continue to blur, until they disappear altogether.

5.1 Platform-independent Solutions for Document Distribution

Solutions exist that allow documents to be electronically distributed to, and viewed or printed, independent of the computer platform used. Two such solutions, developed by Adobe Systems, Inc., use printer languages to render copies of documents in one of two other formats: PostScript or PDF.

PostScript is a computer language that "tells" your printer how to reproduce a hardcopy of an electronic file. Currently, viewing PostScript files onscreen requires a third-party application, such as Aladdin's GhostScript Viewer. GhostScript (GS) is a public domain PostScript interpreter; and GS Viewer is a virtual printer that prints "electronic sheets of paper" on your screen. GS Viewer is available for Mac, PC, or UNIX; and it is WWW compatible. Users may reset Mosaic's or Netscape's preferences to launch GS Viewer whenever they wish to view a PostScript file via the WWW. However, GS Viewer is slow and still rather clunky relative to Adobe Acrobat. (For instance, the current version of GS Viewer doesn't allow you to print--on a Mac, you need Drop•PS to print--nor does it allow you to browse backward through a document.)

Adobe's family of Acrobat software applications is another platform-independent document distribution solution. Acrobat is a printer language that "distills" copies of documents into a Portable Document Format (PDF) while maintaining the "look" of the original document as best it can in the PDF version. Then, for onscreen viewing or printing, Acrobat PDF "Readers" are freely available for Mac, PC, and UNIX platforms. As stated earlier, NASA has adopted Acrobat Reader as a part of its standard Internet software suite. Anyone may obtain a free copy of the Acrobat Reader from Adobe's WWW location--`http://www.adobe.com/Acrobat/`.

MAST procured copies of a number of Adobe's Acrobat software modules to further enhance its document distribution capabilities. Copies of Acrobat Exchange were distributed to the entire MODIS Science Team. Additionally, SDST has procured copies of Exchange. [Note: Exchange has more features than Reader, so if you have Exchange you do not need Reader.] MAST also has a copy of Acrobat Distiller, which can convert PostScript files into PDF. (Adobe's PDF is discussed in more detail in section 6.1 of this report.)

6.0 DOCUMENT DISTRIBUTION & INFORMATION MANAGEMENT

Over the last 2 years, two uniquely different platform-independent applications for document distribution have emerged almost simultaneously and taken root firmly within the realm of electronic information management. One application is Adobe's Acrobat modules, the other is CERN's World Wide Web (WWW), which uses HyperText Markup Language (HTML) to render information publicly available via the Internet using CERN's hypertext transfer protocol (HTTP). These two applications are becoming public standards and, perhaps as a result, are also increasingly becoming compatible with one another. Moreover, they offer two different means for adding a new dimension to MODIS' archive and document distribution functions--information management.

During the ATBD development exercise, MAST recognized that team members would benefit by having a tool that enables multiple, platform-independent inputs into document development. The idea is that MODIS Team members separated by barriers of geographic distance, as well as differing computer platforms and differing software tools, should still be able to jointly author or peer review documents. Moreover, these same Team members may also wish to have some degree of control over how the information contained in their documents fits within the larger body of information in MODARCH. The system administrator/archivist will always retain the greatest degree of hands-on control of the input and maintenance processes, yet, in the future, Acrobat and WWW tools may be employed by Team members to determine how certain documents, or specific information within a document, can be rendered more "obvious" and therefore more accessible and more useful to the rest of the community.

6.1 Acrobat Links

Acrobat Exchange allows users to create hyperlinks within PDF documents. These links may connect related information within a single document, or they may connect related information between different documents, or they may even link to URLs available on the WWW. When creating a document for distribution, authors may embed useful links to other information sources--whether in the EFS Archive, at an FTP site, or on the WWW--within their document prior to distribution or submission to

MODARCH. Consequently, the information in that document is enriched by immediate access to a larger body of information. By analogy, imagine the set of documents within MODARCH as a set of Lego™ toys. Each document has its own structure and its own purpose, yet it can also be interconnected to other documents within MODARCH to build other, more useful information structures. By taking advantage of the new features available in EFS, Acrobat Exchange, and the WWW, it is now possible to begin connecting centralized and distributed information resources. Noteworthy is the fact that a tool is now in place to facilitate interconnection of documents within MODARCH, which could ultimately provide the means for users to trace common threads of information they contain.

From this point on, the focus of discussion in this document shifts from document distribution to information management. Though these topics are closely related, they also differ. Hereafter, use of the word "archive" refers to the function of electronically preserving a group of documents, at one site, while also providing metadata or some other means of easily searching for and retrieving desired information from that site. "Document distribution" refers to the process of electronically delivering copies of certain documents to specifically-chosen people, at different sites, in a robustly usable format. "Information management" refers to the methods of enhancing the archive and distribution functions by overlaying a procedural structure that renders the body of information contained in the archive more accessible, more relational, and more usable. Information management combines the archive and distribution functions so that information resources need not be kept at one site.

6.2 How the Team May Contribute to MODIS Information Management

When creating a MODIS document, authors should keep in mind how they want that document to be archived, as well as how they want it to distributed to other Team, and whether they want it accessible to persons not on the Team.

6.2.1 Submission to the Archive

If distribution is of no concern and the author simply wishes the document to be placed into the EFS archive for historical preservation, then the author need only submit the document to the MODARCH system

administrator in its native format (as described in Section 3.0 of this report), preferably integrated into a single file. Additionally, PDF is an acceptable format for submission to the archive. Subsequently, users may retrieve the document via EFS and view it, by way of EFS' Launch Association utility (discussed in Section 8.2), in MS Word, WordPerfect, or eventually one of the Acrobat viewers.

6.2.2 Joint Authoring/Peer Review

     If the author wishes to make a document robustly available for solicitation of feedback or input into the development of that document, the author should follow one of two courses of action. (1) Have the document marked up in HTML for placement on the WWW at a non-published URL (This method has been dubbed "stealth web" in that anyone not knowing the URL will have difficulty locating the document. Generally, no one even knows to look for it.). Once the document instance is located at a nonpublished URL, the author may e-mail its location to only those persons he or she wishes to review the document, or contribute to its development.

     MAST has established a "MODIS Test Page" for precisely this purpose and we have placed a number of document instances there for review via the "stealth web". If any MODIS Team member would like a document instance that he/she is working on to be placed on the MODIS Test Page for review by the Team, simply submit it to the MODARCH System Administrator and the documents can be made available almost immediately. A URL will be assigned and this information will be forwarded back to you so that you may decide whom you wish to have access to your document.

     (2) Convert the document to PDF, and then e-mail it to the desired reviewers/joint authors. Or, submit the PDF file to MODARCH and the system administrator will render it available via the FTP site. (The author may even set security limits on access to that file--Acrobat Exchange allows password protection.) Acrobat Exchange allows reviewers to affix textual "post-it notes" within the text of a PDF document. Then, after review, the document may be returned to the original author or chief editor of a document so that all comments and/or corrections may be integrated into the final version. Or, if the reviewer wishes, he/she may export the post-it notes to a separate file and then e-mail this file to the original author. The author may then import the notes into the original PDF file and these notes will

"remember" where (exact page and placement) they are supposed to go in the document.

6.3.3 Information Management

Creating links between documents available on MODARCH will render them more accessible and more usable. Excalibur Technologies is adding many new applications with which its *Launch Association...* utility will be compatible. Adobe's PDF will be among these newly-compatible formats. Ultimately, PDF copies of all archived documents will be stored on MODARCH. So, whenever a document is submitted to MODARCH, the submitter may request that it be linked to other PDF documents within MODARCH, or even other resources on the WWW. The document creator can create the links he/she desires before submitting the document. If PDF versions of the documents to which the author wants the submitted document to be linked do not exist, then the MODARCH Administrator may create PDF files from the files already residing in the EFS archive, whether they are text files or bit-mapped images of scanned hardcopy. The MODARCH system administrator will restrict control of creating new links in archived documents to the author. In other words, no one but the original author may replace an archived document with a PDF copy with links in place. Otherwise, quality control could become a nightmare for the system administrator.

However, upon completing a topic search in the EFS archive, a user may retrieve a dozen or so related or highly complementary documents pertaining to a subject. That user could then e-mail a PDF reference list to the MODARCH System Administrator listing the documents he/she would like to see linked. That same user would take the time to create the new links in his/her PDF reference list. In this way, over time, a number of reference lists on various hot topics could be made available in MODARCH that will allow other users to quickly trace threads of information through any number of archived documents.

7.0 CONCLUSION

Clearly, MODARCH has plenty of room to grow in its utility to the MODIS Team. This text merely presents a small subset of the available options. This was not an attempt to present an exhaustive, A-to-Z discussion of all of the possibilities; rather, this report briefly explains what currently appears to be the most cost-effective and cost-efficient means to expand MODARCH's capabilities.

With a relatively modest investment of time, primarily on the part of the MODIS Team Members, MODARCH can take advantage of the current state-of-the-art in commercial and public domain software, and evolve to become much more than a document archive. Now, the tools are available for jointly authored documents, as is the framework in which the entire Team may participate in the management of the body of literature surrounding MODIS.

8.0 TUTORIALS

8.1 Printing From the EFS Client

Some MODIS Team members were initially "turned off" to MODARCH when they encountered printing problems--EFS client users couldn't print specified ranges of pages, and print speed was unacceptably slow. A short-term solution to these problems--a Client Print API (Application Programmable Interface)--has been implemented and is freely available to the Team as part of the EFS Client system. Step-by-step instructions are provided below:

1. Once you have located the document you wish to print, double-click on it to open the Document window;
2. On your desktop, in your MODARCH folder, double-click the Print.API icon to launch it;
3. Click "Connect" and, in the next window that appears, confirm that your desktop computer is selected, that your correct network zone is selected (if applicable), and that you are using the EFS Client software. Then click "OK".

4. Click "Refresh". The path to the document currently active in the EFS
   Client Document window will appear in the "Doc Path" field.
5. You may now select a range of pages to be printed, or leave it set on the
   default "All" setting, and click "Print".

Whereas it previously took 5 to 6 minutes per page to print a document, it
now typically takes only 1 to 1.5 minutes per page.

8.2 EFS Client's "Launch Association ..."

   To enable your EFS client to launch MS Word or WordPerfect as your
text viewer, open the application on your computer and then open the
document window by clicking on the "book" icon in your toolbox window.
Next, under the "Document" menu at the top of your screen, select "Launch
Association ..." When the "Launch Association" window opens, click "New".
In the field beside "Enter New Extension:", enter the letters "doc" (The
filenames of all native format MS Word documents entered into MODARCH
will end in ".doc". So, by entering ".doc" into this field, you are telling EFS
that you wish to have the option to view any file selected ending in ".doc" in
an application other than EFS.), and click "OK". A window will appear that
enables you to navigate on your hard drive to where your MS Word or
WordPerfect is stored. Do so, then highlight the application you wish to use
as your viewer (MS Word or WordPerfect) by clicking on it once, and then
click the "Open" button. You may now use EFS to launch MS Word or
WordPerfect for viewing/printing documents entered electronically into
MODARCH.

   From now on, EFS will keep this information stored in its Preferences
folder, so it will "remember" which application you wish to use as your text
viewer. Anytime you retrieve a document so that it is displayed in the
Document window, EFS will give you the option of viewing it in ASCII, by
clicking the "Text" button, or in the application (MS Word or WordPerfect)
you selected by clicking on that application's logo, which is now also
displayed.

   By following this same sequence of commands, you may also select a
viewer for tiff images stored in MODARCH. For example, you may use
GraphicConverter to view color remote sensing images stored in MODARCH.
Moreover, Excalibur Technologies is planning to expand upon this
functionality in future versions of EFS--they are expanding both the number

of input formats with which it is compatible, as well as the number of viewers users may choose (See Appendix 1 for a complete listing of all applications for which EFS has filters.). For example, the next version of EFS will accept and display documents formatted in Adobe's PDF. This development is particularly exciting in that Adobe is currently modifying Acrobat so that it will be WWW-compatible.

In the short run, by allowing users to view documents in whatever application they choose, EFS solves both the archive problem, as well as the distribution problem--users may search for and retrieve documents electronically, view them onscreen, or print them. To become viable, this solution depends greatly upon the help of the entire MODIS Team. Each member should adhere to the recommendations given throughout Chapter 3 for creating documents and presentations.

8.3 Adobe's Line of Acrobat Products

As stated previously, MAST procured Adobe's Acrobat for Workgroups software. Acrobat for Workgroups contains the following modules: (1) Acrobat Reader, for viewing, navigating through, and printing any PDF document; (2) Acrobat Exchange, for adding value to electronic documents--users may view or print PDF documents, as well as annotate, build navigational links (similar to hyperlinks used on the WWW), and add security controls to PDF files; (3) Acrobat Distiller, for converting any PostScript language file into PDF format (Additionally, Distiller should be used for files that contain high-resolution or complex EPS images.); (4) Acrobat Catalog, for creating full-text indices for collections of PDF files shared over a network; and (5) Acrobat Search, for conducting full-text searches of indexed PDF files. A soon-to-be-released sixth Acrobat module rounds out the family nicely: (6) Acrobat Capture converts hardcopy documents into PDF format so that when displayed onscreen or printed, the PDF version looks like the original hardcopy version. Capture employs OCR/ICR (Optical Character Recognition/ Intelligent Character Recognition) software to convert scanned and bit-mapped images into textual characters. Or, if Capture cannot recognize a certain character to assign it a typeface, then it simply uses the bit-mapped character and leaves its interpretation to the reader. Capture also allows users to edit electronic PDF files after the

OCR/ICR process, so post-OCR clean up, if desired, is relatively quick and easy.

Once a document is in PDF, Acrobat Exchange users have a variety of tools for manipulating the information within the document. Briefly, aside from those capabilities aforementioned, Exchange allows users to add "post-it" notes, electronically "staple" pages, or insert hypertext links for navigating through information in a collection of PDF files or documents. These features enable the joint authoring and information management objectives mentioned in chapter 6.

As stated previously, Adobe's PDF is well on its way to becoming a government and commercial industry standard. Adobe is adding in the capability to render HTML versions of PDF files, which may also eventually lead to an easy way to interface MODARCH and the WWW.

8.4 Hypertext Markup Language (HTML)

There are some differing concepts among MODIS Team members as to how to make a document available on the WWW. As discussed previously, a number of NASA organizations accept documents in their native formats, convert them into other formats (PostScript, ASCII, RTF, or PDF), and then make them available via file transfer protocol (seen as a hyperlink on a Mosaic or Netscape page). This differs from actually rendering a document available in HTML.

Whenever you enter the WWW, you do so via a common graphical user interface viewer, such as Mosaic, Netscape, SpyGlass, Hot Java, Viola, etc. The information, or "document instance," you see onscreen was prepared in HTML and is brought to your machine by way of CERN's Hypertext Transfer Protocol (HTTP) software (which is why addresses for resources available on the WWW begin with "HTTP://"). HTML was adapted from and is defined in terms of the ISO Standard Generalized Markup Language (SGML) (Berners-Lee and Connolly 1993). SGML is a system for defining structured document types and markup languages to represent instances of those document types. Every SGML document has three parts:
- An SGML declaration, which binds SGML processing quantities and syntax token names to specific values;

- A prologue, including one or more document type declarations, which specify the element types, element relationships and attributes, and references that can be represented by markup; and
- An instance, which contains the data and markup of the document (Berners-Lee and Connolly 1993).

So, the term "HTML" refers to both the document type and the markup language for representing instances of that document type (Berners-Lee and Connolly 1993). In short, marking up a document in HTML involves structuring the text by embedding certain "tags" around particular "elements" of the document to give it the desired appearance. For example, consider the following HTML code:

```
<H1>
An Example of a Level 1 Header
</H1>
```

The tags (enclosed in <> symbols) are commands indicating that the element (text, in this case) should appear in a viewer (such as Mosaic or Netscape) as a level 1 header, as follows:

An Example of a Level 1 Header

APPENDIX 1

The following are applications and/or formats for which EFS version 3.6 has filters, and therefore are acceptable as submission formats to the EFS archive on MODARCH*:

1. ASCII (Standard, Smart, Stripped)
2. Ami Professional 1.1, 1.2, 2.0, 3.0
3. CEO Write
4. dBase II, III, III+, IV (1.0)
5. DEC DX (PS Plus)
6. DisplayWrite 2, 3, 4, 5, 5.2
7. DCA/FFT, DCA/RFT
8. DEF/Document Designer
9. EBCDIC
10. Enable (wp) 1.0, 2.0, 2.15
11. Excel/Win 2.1, 2.2, 3.0, 4.5
12. Excel Mac 3.0, 4.0
13. FoxBase
14. FrameMaker/MIF 3.0, 4.0
15. Framework III (wp) 1.0, 1.1, IV
16. HP AdvanceWrite Plus 1.0, 2.0
17. HP PCL (export)
18. IBM Writing Assistant 1.0
19. Interleaf Publisher 1.1 (ASCII markup)
20. Interleaf 5.2, 6.0
21. Lotus 1-2-3, 1A, 2.0, 2.2, 3.0, 4.0
22. Lotus Manuscript 2.0, 2.1
23. MacWrite 4.5, 4.6, 5.0, Pro 1.0
24. Mass-11 Ver. 8.0, 8.3
25. Microsoft RTF (Rich Text Format)
26. Microsoft Word 3, 3.1, 4.0, 5.0, 5.5, 6.0
27. Microsoft Word for Windows 1.x, 2.0, 6.0
28. Microsoft Works 1.0 DOS (export)
29. Microsoft Write 3.x (Windows)
30. MultiMate 3.3, Adv. 3.6, Adv. II 3.7, 4.0
33. OfficePower 6, 7
34. OfficeWriter 4.0, 5.0, 6.0, 6.1, 6.2
35. OnGo
36. Paradox 2, 3, 4 (import)
37. PeachText 5000, Version 2.12
38. PFS: First Choice 1, 2, 3, 4/database 3, 4
39. PFS: WRITE Version C
40. PostScript (export)
41. Professional Write 1, 2, 2.1, 2.2, WritePlus
42. Q&A Write 1.0, Q&A (wp) 3.0, 4.0
43. Quattro Pro (DOS) 3.5, 4.0 (import)
44. Quattro for Windows 1.0 (import)
45. RapidFile (memo) 1.x
46. Samna Word IV, IV Plus, 1.0, 2.0
47. Uniplex V7
48. Total Word 1.2, 1.3
49. Volkswriter Deluxe 2.2, 3 Ver. 1, 4
50. Wang PC Ver. 3
51. Wang WP (IIs), WP Plus
52. WordPerfect DOS 4.1, 4.2, 5.0, 5.1, 6.0
53. WordPerfect Windows 5.x, 6.0
54. WordPerfect Mac 1.0, 2.0, 2.1, 3.0, 3.0A
55. WordStar 3.3, 3.45, 4.5, 5.5, 6.7
56. WordStar 2000, Rel. 3.0, 3.5
57. WordStar for Windows 1.x
58. WPS Plus (ALL-IN-ONE wp)
59. Wang WITA
60. WIZIWORLD
61. Write Now (Mac) 3.0
62. Xerox Globalview (XIF)

31. Navy DIF
32. NBI Legacy 1.x, 2.0

63. XyWrite III, III Plus, IV, Signature
64. XyWrite for Windows

*EFS version 3.7 will also contain filters for PDF and Microsoft PowerPoint files.