

*The Manufacturing
Information Technology
Transfer Project
Electronic Library: An
Implementation
Description*

**Michelle D. Potts
Peter F. Hart**

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards and Technology
Manufacturing Engineering Laboratory
Manufacturing Systems Integration Division
Gaithersburg, MD 20899

Table of Contents

Introduction	3
Client/Server Model	4
Server Capabilities	5
Email Archive Server	5
Gopher	5
FTP	6
WWW	7
WAIS	7
Majordomo	8
Related Tools	8
Report Generators	8
HTML Editors	8
Restricted Login Services	9
Kermit	9
xfront	9
Express Server	9
Clients	10
Mosaic	10
Gopher	10
WAIS	10
Data Repositories	11
Elib's Manufacturing Information	11
STEP Information	12
WWW Documents	13
Access and System Security	13
Elib Service Metrics	14
Glossary	15
Appendices:	
A: Software Specifications	17
B: Hardware Specifications	18
C: Information Architecture and Access Mechanisms	19

D: xfront Applications	21
Bibliography	22

The Manufacturing Information Technology Transfer Project Electronic Library: An Implementation Description

Introduction

The Manufacturing Information Technology Transfer (MITT) project is funded under the auspices of the Systems Integration for Manufacturing Applications (SIMA) program, which is part of NIST's (National Institute of Standards and Technology) HPCC (High Performance Computing and Communications) Initiative program. The project is being implemented in the Manufacturing Systems Integration Division (MSID) which is part of NIST's Manufacturing Engineering Laboratory (MEL).

The MITT project's objective is to provide access for the manufacturing community to manufacturing research and automation technology through the use of emerging tools, technologies and growing electronic networks, resulting from the information exchange revolution. The project's first year approach has centered on building a coordinated and focused information infrastructure using the Internet and freely available tools to provide access for the manufacturing community to MEL's manufacturing information repositories. This document is a description of those services provided and the MSID information structure housing the data repositories served. These dissemination mechanisms and practices are a pillar in the MITT's proposed infrastructure to support on-line collaboration among researchers in the manufacturing community.

This document is intended to be a snapshot of the MITT Electronic Library (Elib), its services¹ and information structure near the end of the MITT

1. Certain commercial software packages, hardware systems and trade names are identified in this paper to specify the project adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the items identified are necessarily the best available for the purpose.

project's first year. The services available on Elib allow users to search, browse, and retrieve manufacturing information using a wide range of access methods freely available on the Internet. The Elib system is designed to allow access for users with only basic Internet capabilities, such as electronic mail (email), as well as to those users with quite sophisticated tools, such as WWW (World Wide Web) browsers. The information stored in Elib is currently composed of a manufacturing research information repository including NIST research papers, programmatic descriptions and pointers to other pertinent manufacturing research information repositories.

The intended audience of this document is anyone interested in:

- what tools and server capabilities the MITT project has deployed, which are detailed in the Server Capabilities, Restricted Login Services, and Clients sections, Appendices A and D,
- how the MITT project's information bases are organized to minimize data redundancy while maximizing allowed access mechanisms which is detailed in the Data Repositories section and Appendix C,
- access and security considerations are discussed in the Access and System Security section,
- usage statistics (use and generation) is discussed in the Elib Service Metrics section, and
- server (machine) configuration to support Elib – server details are in Appendix B.

Client/Server Model

Many of Elib's services conform with the client/server model. In the client/server model, there are conceptually two parts, a front-end and a back-end. The front-end interfaces with the user, providing a means to communicate with the back-end, which does the work based on the commands sent through the front-end.

A front-end is called a *client*; it is usually responsible for:

- providing the user interface,
- taking user commands and conveying them to the appropriate server,
- opening and maintaining connection(s) to the appropriate server programs,
- receiving messages and information from the server and appropriately displaying it, storing it, etc.

A back-end is called a *server*. A server carries out the corresponding activities specified by the client's commands, sending the output of any of these commands back to the client program.

The next several sections detail Elib's server capabilities, related tools and client services. Particulars about each software package's version at the time this was written are in Appendix A.

Server Capabilities

The information servers chosen and configured for the Elib range in sophistication from simple electronic mail access to remote execution of application software. This suite of services provides a variety of access capabilities for users with older interfaces, as well as, state of the art access for high tech Internet surfers.

The currently available servers and related tools on Elib are detailed below.

Email Archive Server

The Email (electronic mail) Archive Server allows users to request and retrieve directory listings and copies of files using electronic mail. Users send email messages containing keywords to request information and the server returns email messages containing the requested information. The Email Archive Server uses the keywords in the request message to determine what information, and possibly its format, to return to the requestor. Directory and file listings are returned to the requestor in the response email message(s).

The Email Archive Server also provides encoding for emailing binary files to a requestor. Binary files are specially formatted files or files which contain special characters; they must be encoded to preserve their contents within an email message. They may be encoded in one of several ways. The server provides encoding for MCI mail, binary to ascii (BTOA), and unix to unix (UUENCODE). Users can specify the type of encoding they require from within their email request message body.

The Email Archive Server described above is a generalized concept. Elib provides two specialized instances of this concept: the STEP email archive server and the general Elib email archive server. The specializations are a result of customization of the available command set for the targeted applications and audience.

Gopher

The Internet Gopher server provides hierarchical menu access to a variety

of Internet information resources. Menu items can be documents, directories, searchable index queries, telnet connections, or links to other Gopher servers. Menus are normally generated automatically by the server or manually by the system administrator.

Gopher has a close relationship to WWW, FTP, and WAIS.

Elib's functionality requirements:

- Search local Gopher indices of documents.
- Search local WAIS indices of documents.
- Query remote WAIS servers and return the results to gopher clients.
- Query remote ftp sites and return the results to gopher clients.
- Be queried by WWW (World Wide Web) clients either using built-in gopher querying.
- Invoke telnet sessions to remote hosts.
- Directory level access control.
- Link to other Gopher servers in gopherspace.
- Menus may be generated automatically by the server, or by using a script. When scripts are used, they must be executable by the staff responsible for maintaining those menus.

FTP

The file transfer protocol (ftp) is the primary mechanism by which files are copied to and from Internet hosts. The two types of ftp are full service and anonymous ftp.

Full service ftp permits the two way transfer of files between two computers connected to the Internet. It requires that the user have an account on the machine being accessed, as well as the appropriate passwords and access privileges necessary to read and write to that system.

Anonymous ftp allows anyone to connect to and download copies of files that are publicly available on the anonymous ftp server. Users do not need an account on the server machine and will typically login with the user name "anonymous" and use their email address as the password.

Elib's functionality requirements:

- Serve text and non-text files.
- Provide access security capabilities.
- Provide anonymous ftp logins.

- Provide limited login capability for uploads.

WWW

The World Wide Web (WWW) uses the client/server model and provides a hypertext interface to Internet resources. The hypertext, or hypermedia, document concept is one where there are references or links within a document which are jumping-off points to other documents (which may contain multi-media), lists, resources, or actions. These resources include information that can be accessed using any of the following communications protocols: gopher, WAIS, telnet, ftp, NNTP (Usenet News groups), and HTTP.

WWW servers use the Hypertext Transfer Protocol (HTTP) for communication with clients. Hypertext documents must conform to the HyperText Markup Language (HTML) specification. WWW clients handle all of the necessary protocols that allow users to navigate the vast world of Internet resources with a single hypertext user interface. Users are able to follow hypertext links in HTML documents to browse and retrieve information that may be located anywhere on the Internet. These links, called Universal Resource Locators (URLs), allow for transparent communication with various Internet protocols.

Elib's functionality requirements:

- Ability to serve hypertext documents to clients.
- Ability to serve multi-media documents to clients.
- Provide directory level access security.
- Execution of scripts.
- Support for forms.
- Support for access reporting.
- Database index searching.

WAIS

Wide Area Information Servers, or WAIS, is an architecture for the search and retrieval of information in distributed database systems. WAIS provides a full text searching capability based on the Z39.50 standard query protocol. It uses the client/server model over a TCP/IP connection. Clients perform natural language keyword searches of distributed databases and are returned a list of the files which contain the specified keywords. From this list, a user may view or download a file, or formulate a new query to further filter the list of files. The list of files may contain documents from different sources located on the Internet. The identity of these sources is hidden from the user. However, the user may specify a set of sources to be used for the query to

limit the search domain.

Elib's functionality requirements:

- Serve text and multimedia documents.
- Full text indexing.
- Name only indexing for non-text files.
- Generation of local and remote indices using source description.
- Incremental index generation.
- Database source access security.
- Domain/host security features are required for index source access.
- Access of WAIS indices for gopher and World Wide Web servers.
- Generate of indices for use in local as well as remote searches.
- Provide incremental index generation and the ability to recursively index subdirectories.

Elib's text documents will be fully indexed, while other file formats will be indexed by name only. The non-text file formats will include postscript, graphics, movie, and sound.

Majordomo

Majordomo is an email based service that automates the management of mailing lists. Users send email messages that contain commands to majordomo requesting actions such as: subscription to or unsubscription from majordomo supported mailing lists, help/information requests, and transfers of files and available (for transfer) file listings.

Related Tools

Report Generators

Report generating software provides the ability to filter log file entries for specific server access information and statistics. Selected servers, currently: kermit, ftp, http, wais, and gopher, each have an utility which processes its server's log file(s) to automatically generate usage statistics on a regular basis.

HTML Editors

Various HTML editors and software packages which supplement existing editors are used to facilitate writing html source.

Restricted Login Services

Kermit

A Kermit server application is available for performing file transfers using the Kermit protocol. Most modem communications applications have kermit capabilities that provide commands for file uploading and downloading. The Kermit server is invoked by logging into the Elib server and specifying the user name “kermit”. Users are then prompted for their first and last names, followed by their phone number. After this has been completed, a series of information screens are displayed showing general help and information. The kermit server then becomes ready for interactive use.

This server is accessible via a dial-up phoneline (301/948-9720) for modem communications. For detailed information regarding access, see [1].

xfront

xfront is a utility which allows users to remotely execute selected applications on Elib while displaying the application’s output locally.

It is invoked by logging into the Elib server with the user name “xfront”. The user is prompted to enter the appropriate DISPLAY variable value, which often is “hostname:0.0”. After the display variable value has been correctly entered, a menu of available applications is displayed, and the user is prompted to enter a selection. The selected application begins execution and is remotely displayed at the user’s system.

See Appendix D for available applications.

Express Server

The Express Server [5] provides users with email access to a suite of automated services surrounding EXPRESS (ISO 10303², Part 11) and “Clear Text Encoding of the Exchange Structure” (ISO 10303, Part 21).

These services include:

- EXPRESS schema analysis for semantic and syntactic errors.
- Part 21 exchange file analysis for semantic and syntactic errors.
- Data Probe [2] creation and use for entity instance browsing and editing.
- Generation of Annotated Listings by merging multiple Integrated Resource Model schemas using a specified Application Protocol.

2. ISO 10303 is an International Standard for the computer-interpretable representation and exchange of product data.

Clients

- Retrieval of standard and draft standard schemas from a public files repository.
- User collaboration on schemas and Part 21 information.

There is also login access to the Express server. Users login to the Elib server specifying the user name “express”. The user is then prompted for his DISPLAY variable value and identifier key. This mechanism is restricted in its use.

Clients

Mosaic

Mosaic is a WWW browser developed at NCSA (National Center for Supercomputing Applications). Mosaic communicates with HTTP servers in its “native” mode, but its true strength is that it can also communicate with other widely used Internet protocols such as ftp, gopher, WAIS, NNTP, etc. It does this without intervention by the user, therefore, the user needs only know one client interface to communicate with many servers regardless of the server’s native protocol.

The MITT project is encouraging its users to use Mosaic as their primary client application at the time of this publication.

Gopher

A gopher client is a program which is specifically designed to communicate with gopher servers; it is the interface between the user and an information server. Gopher servers can be local or anywhere on the Internet. A gopher client can use pointer information from one gopher server and use it to automatically access other gopher servers anywhere on the Internet transparently to the user.

Gopher client software is available on Elib, but unsupported.

WAIS

A WAIS client is a program which is specifically designed to communicate with WAIS servers. It generally allows the user to specify the parameters of a WAIS search, which are keyword(s) and source(s), returns the result of the search to the user, and provides a viewing mechanism for selected search results. WAIS clients generally allow multiple WAIS databases to be searched with the one request; this is accomplished by selecting multiple source databases.

WAIS client software is available on Elib, but is unsupported.

Data Repositories

The data repositories on Elib are logically organized in a hierarchical directory structure generally by subject content with some further grouping by dissemination method where appropriate. The system provides general access to the manufacturing engineering information as well as specialized access when appropriate and necessary. Most files are accessible through more than one access mechanism. This was done to maximize the number of users who can access the information based on their Internet capabilities and needs.

Additionally, specialized access mechanisms have been configured for some segments of the manufacturing information for targeted user communities and selected file formats. For instance, the EXPRESS email server services a defined user community using specialized data and services.

Many file formats are used, they include: text, WordPerfect, Postscript, EXPRESS, LaTeX, graphics, HTML, audio, and video. Some files, because of their format, should only be accessed via appropriate servers.

Listed below is a categorization of Elib's data repositories by topic vs. dissemination method. See Appendix C for a graphical representation.

Elib's Manufacturing Information

Description:

Elib's entire data repository, including: programmatic descriptions, published papers repository, on-line computing services, and STEP repositories.

Files Types:

File types include: text, WordPerfect, Postscript, EXPRESS, LaTeX, graphics, HTML, audio, and video.

Storage Organization:

Files are stored in an hierarchical structure, generally organized by topic, although some files are initially grouped by dissemination method and then further grouped by topic/content.

Dissemination Methods applied to data repositories:

- Gopher – SOLIS data.
- Anonymous FTP – all data.
- WWW – selected data sets.
- WAIS – selected data sets.
- Email Archive Servers – selected applications: SOLIS and EXPRESS.
- Kermit – SOLIS data.

STEP Information

Description:

The STEP [1] information includes the latest parts of ISO 10303³ and related documents such as issue logs, schedules and software.

Files Types:

File types include: text, EXPRESS, LaTeX, WordPerfect, style guides, and postscript.

Storage Organization:

Files are stored in an hierarchical directory structure by part number and then by revision.

Dissemination Method:

- STEP Archive Mail Server
- Anonymous FTP
- Gopher
- WWW

3. ISO 10303 is an International Standard for the computer-interpretable representation and exchange of product data.

WWW Documents

Description:

The documents which comprise our local externally accessible Web; the primary access mechanism is through Elib's HTTP server. These documents include descriptions of programs, MSID's published papers repository, pointers to other related repositories, and staff information.

Files Types:

File types include: HTML, text, graphics, audio, video, Postscript and portable document format (pdf).

Storage Organization:

Files are stored in an hierarchical directory structure based on information content.

Dissemination Method:

- HTTP
- Anonymous FTP

Access and System Security

The access mechanisms on Elib were chosen to allow access for users with only basic Internet capabilities, such as electronic mail, as well as to those users with quite sophisticated tools, such as WWW browsers to access its data repositories. Additionally, some data have file format requirements which also must be considered when choosing access mechanisms. These considerations enable small manufacturing companies, that typically have low-end communication capabilities, as well as, other members of the manufacturing community with more sophisticated Internet-based tools, to access Elib's information bases and get the information they need in useful formats.

The Elib is currently primarily accessible via the Internet. Its address references on the Internet are in part determined by access method, see Appendix C for more information. One data repository, SOLIS (Step On-Line Information System) is also accessible via a dial-up phone line: 301/948-9720 and Kermit, see Data Repositories for more information.

Security is a high priority for any system at NIST. Elib is configured to allow users to have controlled access to its files and applications. Software mechanisms and network isolation provide service security, directory level access control, and network file system security.

Appropriately configured server software enables both service and directory level access control. This allows specification of domain and host level access using standard Internet protocol addressing. Additionally, log files are kept and reviewed for requested services access and the corresponding requesting machine.

The Gopher server has an additional low level security mechanism which has been configured to limit the “view” given to a user when accessing menu interfaces. Top level menus, such as the SOLIS Gopher service, were specifically generated to show only selected directory structures.

In addition to software security, Elib has been separated from other NIST computers by a firewall, in that it is physically isolated from other NIST computers. This prevents Elib from mounting any file systems other than those physically attached to Elib itself. Therefore, Elib itself is prohibited from accessing other NIST file systems in any way.

The Elib provides two access modes: publicly accessible and internal-only accessible. The publicly accessible information resides on the firewalled machine. The publicly accessible file system is readable by anyone and has very restricted write access. The internal-only accessible information is served from a non-isolated machine, and the inclusion of additional information sets is not restricted; this was done to encourage collaboration and internal information sharing. The primary advantage of this division is that it is more secure, while not restricting internal use.

Elib Service Metrics

Metrics have been generated for all Elib services starting in January 1994 with the exceptions of the WWW and Gopher servers, which were deployed in March and June respectively. These statistics were generated for the purpose of analyzing the number of users and the number of requests for files made to Elib. User statistics are subdivided to show local and remote user access as well as total number of users. Local users are considered to be anyone with a NIST IP node or email address. The number of users listed in monthly and total statistics indicate the unique users. File request statistics are also subdivided into local, remote, and total requests, as with the user metrics.

Metrics are generated by Perl scripts and C programs customized to meet our needs and system configuration. Most statistical reports contain additional details beyond the basic user and file request statistics. Time of day and day of week statistics are generated for the WWW and Gopher services. The email archive service script gives statistics on the number of indices and directory listings sent to users. Additional metrics are easily implemented due to the use of Perl scripting and standard log file formats.

Glossary

Archie – A network service that searches FTP sites for files.

browser – Software that provides an interface to the World-Wide Web.

CERN – The European collective of high-energy physics researchers (European Center for Nuclear Research).

client – A computer or program that requests a service of another computer or program.

client-server architecture – An information system design in which programs use and provide distributed services.

File Transfer Protocol (FTP) – A common method of transferring files across networks.

Gopher – A versatile menu-driven information service.

Graphics Interchange Format (GIF) – A commonly used graphics format which compresses the image and stores color information within the file.

homepage – (1) the default document a user sees when first invoking a Web browser, also (2) the default document one gets from a Web server site, if no document is requested in the URL. Note: homepage should not be confused with a personal page, a Web document about a person that the individual puts on the Web as self-advertisement.

hyperlink – A connection between hypermedia or hypertext documents and other media.

hypermedia – Files that contain more than one format (text, graphics, audio, and/or video) and contain hyperlinks.

hypertext – Text documents with hyperlinks to other documents.

Hypertext Markup Language (HTML) – The standard language used for creating hypermedia documents within the World-Wide Web.

Hypertext Transmission Protocol (HTTP) – The standard method that World-Wide Web clients and servers use to communicate.

inline image – A graphic in a hypermedia document that is displayed on the same page as text.

Internet – The global collective of computer networks.

Mosaic – A graphical user interface to the World-Wide Web developed by the NCSA.

National Center for Supercomputing Applications (NCSA) – A federally-funded organization whose mission is to develop and research high-technology computing resources for the scientific community.

Network News Transfer Protocol (NNTP) – A common method by which articles over Usenet are transferred.

page – A hypermedia document as viewed through a World-Wide Web browser.

server – A program which provides a service to client programs.

telnet – A program that allows users to use computers across networks remotely.

Uniform Resource Locator (URL) – A standardized way of naming different documents, media, and network services on the World-Wide Web.

Usenet – The global news-reading network.

Veronica – A network service that allows users to search Gopher systems for documents.

webmaster – The administrator responsible for the management and often design of a World-Wide Web site.

Wide-Area Information Servers (WAIS) – A service which allows users to search for information in text repositories throughout the Internet using keywords.

World-Wide Web – (World-Wide Web project) The initiative to create a universal, hypermedia-based method of access to information. By extension, but incorrectly, used to mean the Internet itself.

Appendix A: Software Specifications

Note: At the time of this publication, these specifications are Elib's current versions.

Servers:

- Email Archive Server – Dorequest version 1.19
- Gopher – NorthWestern University Gopher (GN), version 2.03
- WWW – NCSA HTTP, version 1.3
- WAIS – Thinking Machines Corporation MCNC/CNIDR, (Clearinghouse for Networked Information Discovery and Retrieval), freeWAIS version 0.3
- Majordomo – Brent Chapman, GreatCircle.COM, version 1.92

Tools:

- HTML Assistant – Howard Harawitz, version beta 0.6
- HoTMetal – SoftQuad Inc., version 1.0 release 1.9
- Report generators – locally written and maintained scripts

Restricted Login Services:

- Kermit – C-Kermit 5A(189), Frank Da Cruz, Columbia University Center for Computing Activities
- xfront – NIST author: Scott Paisley, version 1.1
- EXPRESS Server – NIST author: Don Libes, version 1.19

Clients:

- NCSA Mosaic:
 - for X: version 2.4
 - for Mac: version 2.0 alpha 8
 - for MS Windows: version 2.0 alpha 7
- Gopher – University of Minnesota
- xwais – Thinking Machines Corporation MCNC/CNIDR, revision 1.18

Appendix B: Hardware Specifications

At the time of this publication, all applications and data are located on a SUN Microsystems workstation, that is configured as detailed below:

- SUN IPC 4/40
- 12 MB Memory
- SUN OS 4.1.3
- 6 GB online disk storage

Appendix C: Information Architecture and Access Mechanisms

Public Archive Directories

*http://elib.cme.nist.gov/
gopher elib.cme.nist.gov
ftp ftp.cme.nist.gov
library@cme.nist.gov*

SOLIS/STEP Documents

*http://elib.cme.nist.gov:70/
gopher elib.cme.nist.gov
nptserver@cme.nist.gov
kermit dialup access*

MSID WWW Documents

http://elib.cme.nist.gov/

The Public Archive directories are accessible through the four services listed above: WWW, Gopher, FTP, and Email Archive Server.

The SOLIS/STEP and MSID WWW documents are a subset of the Public Archive directories and are accessible through their respective service listings, as well as through the services listed for the general Public Archive documents. SOLIS/STEP access includes Gopher, Email Archive and

Kermit dialup service. MSID WWW documents are accessible through the HTTP and FTP protocols.

Appendix D: xfront Applications

At the time of this publication, the current xfront applications are:

- DataProbe (STEP) [2]
- DataProbe (ALPS) [3], [4]
- ShowMe – teleconferencing software

Bibliography

- [1] Gaylen R. Rinaudot, "The IGES/PDES Organization, STEP On-Line Information Service (SOLIS)", NISTIR 5511, National Institute of Standards and Technology, Gaithersburg, MD, October 1994.
- [2] David Sauder, "Data Probe User's Guide", NISTIR 5141, National Institute of Standards and Technology, Gaithersburg, MD, March 1993.
- [3] "ALPS - A Language for Process Specification", International Journal of Computer Integrated Manufacturing, special issue on Process Planning and Design for Manufacture, 4, #2, 1991. p. 105.
- [4] "Using the ALPS Process Plan Model", Proceedings of the ASME Manufacturing International Conference, Dallas, Texas, March 1992.
- [5] "The NIST EXPRESS Server", First International Workshop on Services in Distributed and Networked Environments (SDNE), June 27-28, 1994, Prague, Czech Republic.