CONFIGURATION MANAGEMENT & THE WORLD WIDE WEB: SOLVING THE DOCUMENT ACCESS PROBLEM

Daniel Hymel
Motorola North American Paging Subscriber Division
Computer Aided Software Engineering
Boynton Beach, Florida 33426
EDH006@email.mot.com

ABSTRACT

The management of engineering documents in a multi-user, multi-discipline engineering environment can be a monumental and difficult task. Document access and delivery is a time consuming activity often plagued by unsuccessful searches and untimely delivery of available version(s) of a document. The task is especially difficult when it includes multiple computer platform and network boundaries in pursuit of the specified document.

This paper describes a Document Management System (DMS) and a document "publication" process prototyped and deployed at Motorola's Paging Products Group site in Boynton Beach, Florida. DMS is a trio of software programs that utilize Atria's ClearCase Configuration Management Tool to support the publication process and enforce site configuration identification policies.

1. INTRODUCTION

The management of engineering documents in a multi-user, multi-discipline engineering environment can be a monumental and difficult task. Document access and delivery is a time consuming activity often plagued by unsuccessful searches and untimely delivery of available version(s) of a document.

Integration of the World Wide Web (Web) concept with configuration management of engineering documents [1] is a technology convergence offering promising solutions to this problem.

This paper describes the components of a Document Management System (DMS) and outlines an automated publication process to register, release and publish formal software engineering documents in a multi-user, multi-platform, multi-network software engineering environment.

1.1 ACRONYMS

The following acronyms are referenced in this paper:

CASE	Computer Aided Software Engineering
DMS	Document Management System
WWW	World Wide Web
CI	Configuration Identification
HTML	HyperText Markup Language
SCMP	Software Configuration Management Plan
SCM	Software Configuration Management
SQA	Software Quality Assurance

2. CONFIGURATION IDENTIFICATION

DMS is based on configuration identification (CI), a principle configuration management activity. Engineers and technical writers must be able to identify the documents they are developing or maintaining. Each item must have a unique identifier which must be qualified with the version of the item.

CI provides a standard by which documents and related meta-data are named and labeled. CI addresses additional issues: (i.e. how document numbers will be obtained, applied and controlled.

The development of a document management system can be a very perplexing task if CI were not firmly entrenched within a software engineering environment.

3. MOTOROLA PPG SOFTWARE ENGINEERING ENVIRONMENT

The software and document development environment consists principally of several hundred SUN SPARC client workstations hosting either SUNOS 4.1.3 or Solaris 2.3. Four SUN SPARC CENTER 2000 servers hosting Solaris 2.4 support the client machines. ClearCase 2.0.2, FrameMaker 5.0 and DDTS 3.1.12 comprise the principle document management tool suite. Most projects maintain their source and document files in ClearCase Virtual Object Bases.

4. THE DOCUMENT MANAGEMENT SYSTEM

DMS integrates several engineering and non-engineering domains to manage and deliver documents. See

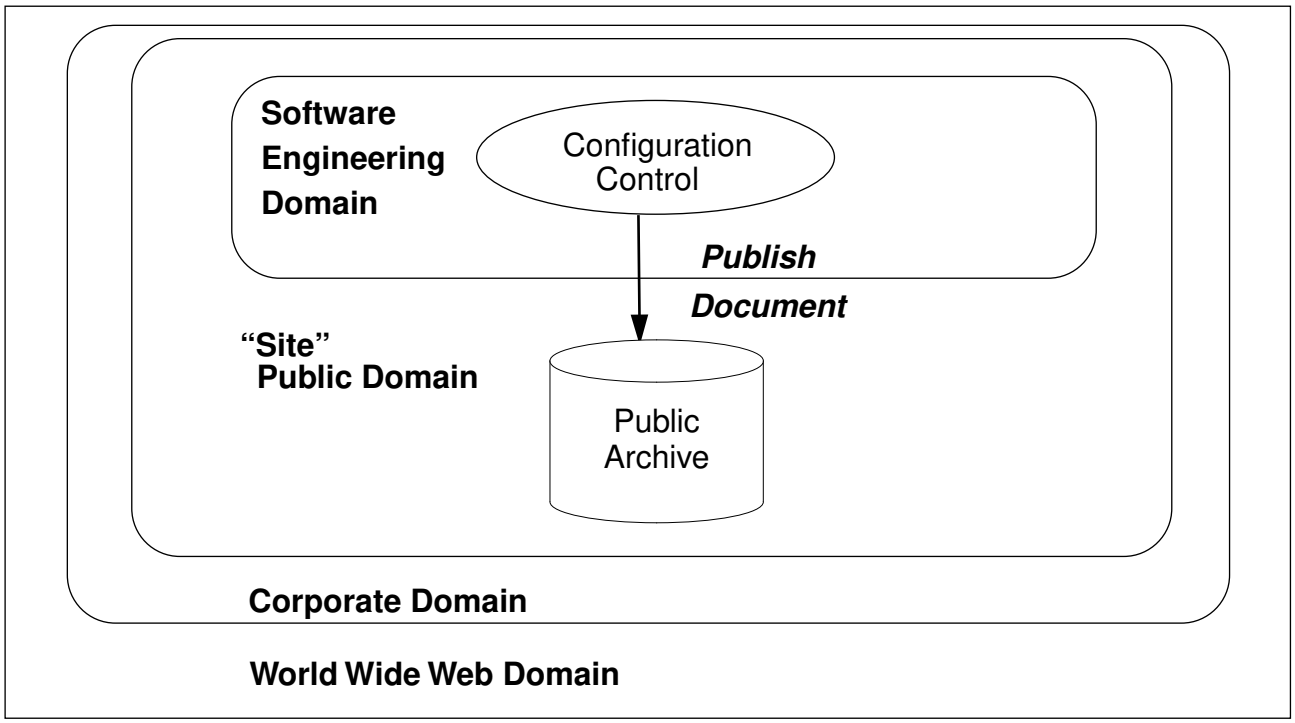

Figure 1 - DMS Domains

Figure 1. Document development and configuration control occurs strictly within the software engineering domain. Documents can be published to the public archive resident in the "Site" Public Domain. Documents are delivered or accessed through World Wide Web Domain based activities.

4.1 DMS PUBLICATION PROCESS

The general document publication process automated by the DMS is shown in Figure 2. Note that as a document approaches the "Publish Document" activity the domain accessibility increases. A multi-file document will be used as an example to describe the document publication process.

4.1.1 DEVELOP DOCUMENT

Documents are developed by members of a project team using the desktop publishing tools sanctioned by the project team and supported by the engineering development environment. The publishing tools generate document file types which are managed against changes in a structured configuration management environment. The CM environment relies upon predefined file naming conventions defined in a projects' SCMP. Table 1 provides an example of the document file types approved for use by a software project team. The document author will baseline the document to allow read access by members of the project team.

4.1.2 BASELINE DOCUMENT

A document is baselined after it has been fully approved. It serves as a basis for further development and can only be changed through formal change control procedures. Figure 3 is a ClearCase listing of a baselined DDTS requirements document. Note that the document configuration is composed of eight files.

```
DDTS_REQ.book@@/main/4
DDTS_REQLOF.doc@@/main/4
DDTS_REQLOT.doc@@/main/4
DDTS_REQTOC.doc@@/main/4
cover.doc@@/main/4
firstbody.doc@@/main/5
lastbody.doc@@/main/4
rev_hist.doc@@/main/4
```

Figure 3 - Example Document Configuration

4.1.3 REGISTER DOCUMENT

Document registration is an activity that associates a

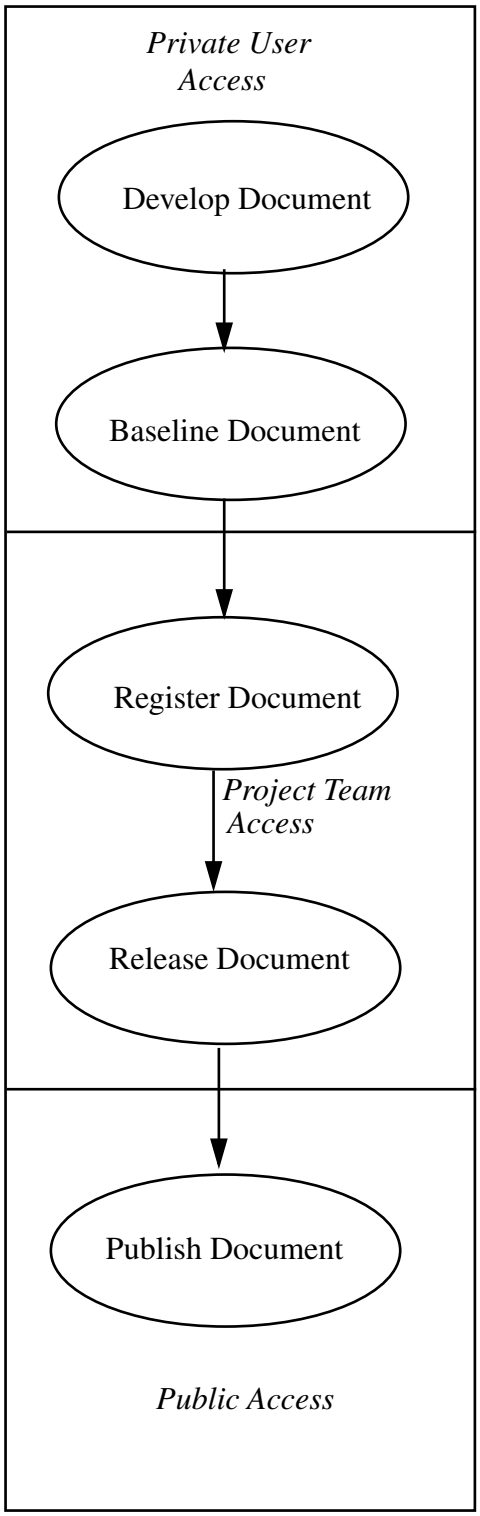

Figure 2 - Publishing Process

unique identifier with a specific configuration of a document. DMS assigned a unique document identifier to this document after the author provided input data fro the following parameters:

- PROJECT_NAME
- ORGANIZATION_NAME
- COMPUTER_TYPE
- AUTHOR
- DOCUMENT TITLE
- DOCUMENT PATH

Data validation and automated data collection are features of **register_doc**. SCM and SQA authorities receive automatic email or pager notification that a document has been registered.

The identifier assigned to this document is **FL19-CASE_SARD_94L000**. The structural definition for an identifier is shown in Figure 4.

4.1.4 RELEASE DOCUMENT

Documentation development is similar to software development in that a new document becomes less volatile as it traverses through various phases of development. The document will retain an "engineering status" during each development phase. For example, when the DDTS document was being developed its engineering status was "WORKING". After it was baselined its status changed to "REVIEW", indicating that the document was scheduled for a review. After review and approval the document attained a status of "RELEASED" and was subjected to the release activity.

The **release_doc** utility automates the following activities:

- Verify document identifier assigned to document.
- Create and apply ClearCase release label to all files and directories in path to a document configuration. [1].
- Remove write access to document files.
- Change document ownership to be that of project database owner.
- Update document mete data for the following attributes:
  - Document Version
  - Engineering Status
  - Publication Status
  - Pathname

Table 1 - Frame and Misc Documentation File Types

DOCUMENTATION FILE TYPE	FILE SUFFIX
FrameMaker Book	.book
FrameMaker Document	.doc
FrameMaker Text	.txt
FrameMaker isolatin 1	.is1
FrameMaker Interchange Format	.mif
FrameMaker Markup Language	.mml
Roman8 files	.rm8
FrameMaker Dictionary	.dict
Postscript	.ps
FrameMaker bitmap	.bm
FrameMaker Encapsulated PS	.epsi
PCX Graphic files	.pcx
Raster images	.ras
X11 window dump image	.xwd
ASCII Text files	.asc
Microsoft Word	.doc

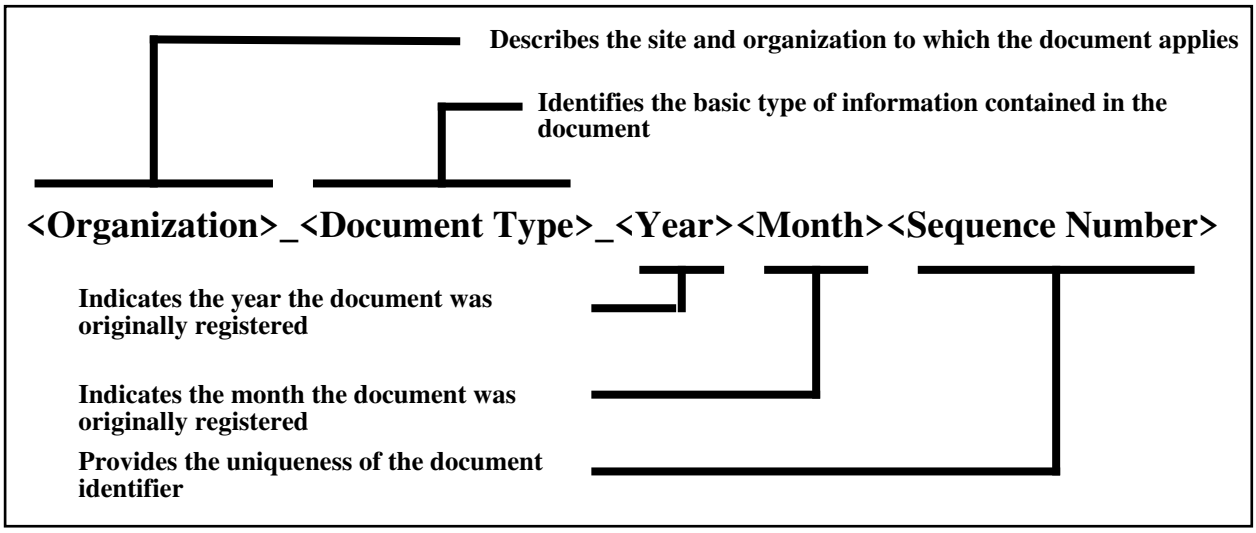

Figure 4 - Structural Definition of the Document Identifier

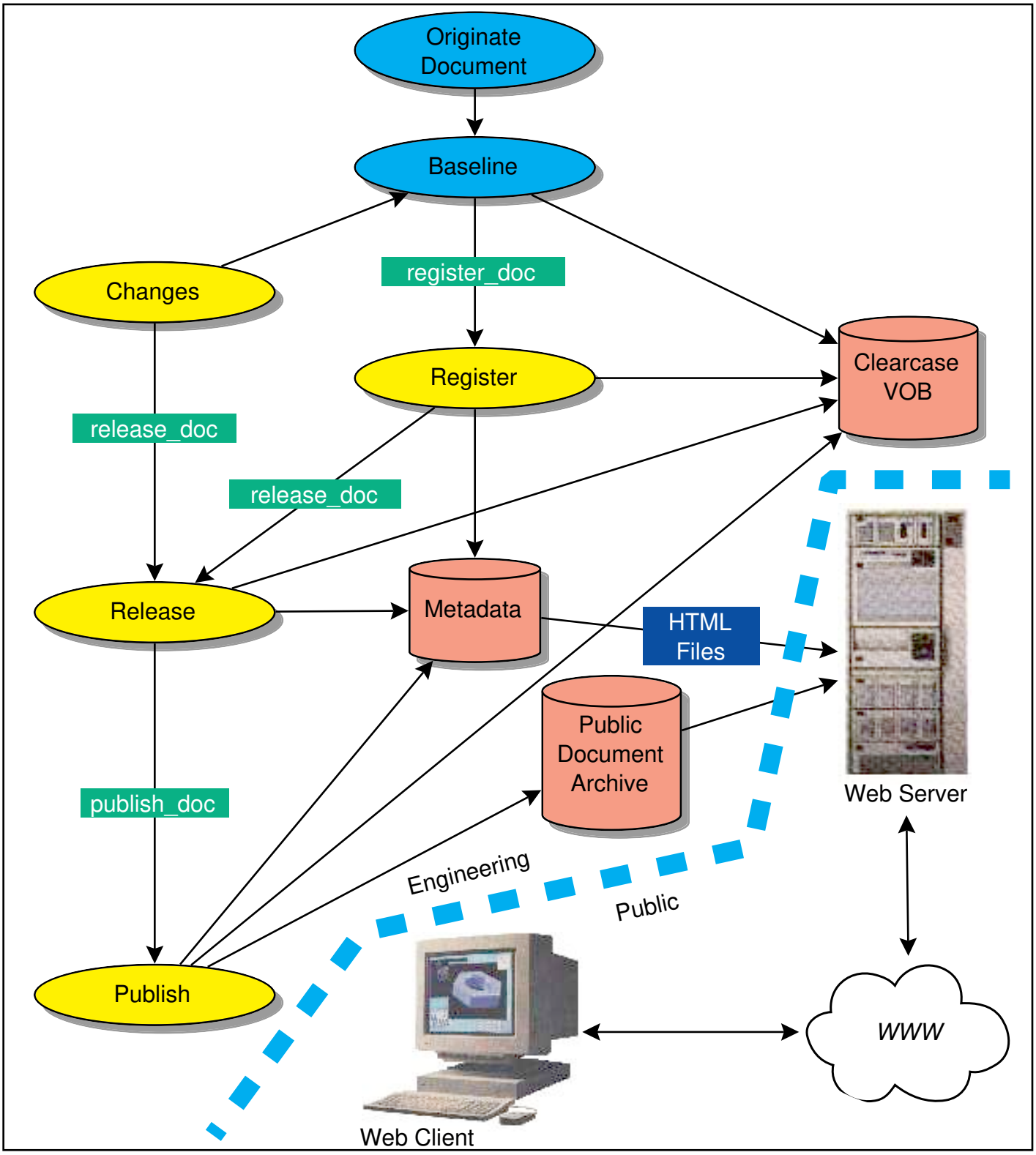

Figure 5: DMS Architecture

- Electronic mail notification
- Derive alternative file formats in preparation for document publication.
- Update list of released documents

4.1.5 PUBLISH DOCUMENT

The notion of document publishing has been enhanced with emergence of the World Wide Web and the **publish_doc** utility. A document is considered published when the following activities have been completed:

- Update document registration database by adding or updating values for the following attributes:
  - Archive Pathname
  - Engineering Status
  - Publication Status
- Copy of document source moved to "***read-only***" document archive.
- Save document in HTML format.
- Save document in MIF format.
- Update HTML formatted list of registered documents.

4.2 DMS SYSTEM ARCHITECTURE

The DMS system architecture drawing in Figure 5 depicts the publication process integrated with the DMS functional components, register_doc, release_doc and publish_doc described in Section 3.2. Other principle DMS components will be identified in this section.

4.2.1 DMS DATA STORES

DMS relies upon three functionally different data stores to manage and maintain the integrity of a document library system.

4.2.1.1 ClearCase VOBs

All documentation source files are retained in ClearCase VOBs.

4.2.1.2 Meta Data Database

This database contains meta data of documents maintained in VOBs. It is the central database of DMS and serves as the common reference for document queries and document access.

4.2.2 PUBLIC ARCHIVE

The public document archive utilizes a simple indexing system for storage of "published documents." An assigned document identifier defines the documents' archival location. Version 1.0 of, FL19-CASE_SARD_94L000, will occupy the following archive directory:

/archive/se_docs/FL19/94/L/0001/1.0

Version 2.0 will occupy**:**

/archive/se_docs/FL19/94/L/0001/2.0

This logical directory structure is accessible to all Motorola domains**. Document files can be accessed by directory traversal.**

5. JUSTIFICATION FOR THE WORLD WIDE WEB TECHNOLOGY

Within the current software engineering environment document access is difficult for project members attempting to access documents outside their project domain. The document registration process was an early attempt to address the document access problem. Knowing the location of a document did not guarantee the accessibility of the document. The matrix of 90 secure project databases among 60 projects distributed across multiple computer platforms presented even the "power-user" with a document access challenge.

Document access across the software engineering domain requires the novice user to have knowledge of the ClearCase command language, VOB access rights, and remote access to workstations having direct access to a VOB retaining the desired document.

Some of the principal reasons for exploiting the Web to solve the document access problem are:

- Web browsers are available for most in-house software development and project management computer platforms that have direct or indirect access to the Web.
- All Web client machines translate HTML document formatting codes so that a link to another document appears as a highlighted item.
- Lists of registered, released, and published documents are easily converted to HTML format supporting automatic Web-Weaving.
- HTML promises to emerge as a powerful language to make documents

readable across a variety of platforms and software.

- The Web provides an easy mechanism to find, view and make use of the huge amounts of information available on the Internet.

6. WEAVING A PUBLIC ACCESSIBLE WEB

Utilizing the Web Development Process described in *The World Wide Web Unleashed* [2], this author integrated Web technology into the Document Management System. The process for automatically weaving a web of HTML links to lists of HTML and non-HTML registered, released and published documents follows.

6.1 WEB PLANNING

Planning the Web component of DMS entailed many months of Web browsing, review and study of several volumes of Web literature, and extensive practice writing HTML documents.

Identifying the audience, purpose, and objectives for the Web accessible documents was well understood when the Web concept was first introduced as a possible solution to the document access problem.

6.2 WEB ANALYSIS

Analysis of Web technology was conducted by many members of the CASE group with the primary focus on its relevance to some pre-existing problems requiring solutions and to its ease of implementation, deployment, and maintenance in the Paging Products Group of Motorola. Web browser software availability and cost were addressed.

It is well understood that this new technology is subject to rapid growth and that applications developed using this technology **must** be designed for change.

6.2.1 WEB DESIGN

HTML specifications and information traversal techniques were some of the principal design issues considered while designing the document Web.

Since the current HTML specification does not support the sophisticated desktop publishing capabilities offered by FrameMaker it was decided that the current FrameMaker document base would be served to a Web client in Maker Interchange Format (MIF). Less sophisticated Frame documents could be saved and served in HTML format.

Information traversal techniques vary among users and differ from that of the casual browser to the seasoned Internet "surfer". The Software Configuration Management infrastructure supporting project development groups was used as a basis for designing and implementing the document Web.

6.2.2 WEB IMPLEMENTATION

Each time a document is registered, released and published, its related record of meta-data is validated. A Netscape view of this metadata is presented in a card catalogue. See Figure 6.

Each day a corresponding set of three HTML documents are automatically generated. The first document is a list of registered documents. The second is a list of released documents. The third is a list of published documents directly accessible through the Web. The list of published documents contains links to the file components of specific documents.

6.2.3 ACCESSING A DOCUMENT USING THE WEB

The DMS home page is shown in Figure 7.

The DMS Web services at this writing include:

- Viewing lists of registered, released and published software engineering documents
- Direct access to published software engineering documents from the published list on any computer hosting a Web browser that has access to the Motorola PPG Web server. This is currently limited to in-house use.
- A form (not shown) to submit comments and suggestions for improving the document Web.
- Keyword searches and group delivery of specified documents.

When the highlighted **Published Documents** link is selected, the Web server presents a linked HTML document to the Web browser running on a client computer. When processed, the HTML document reveals a bibliographical list of published software engineering documents. Figure 8 shows the bibliography for the example DDTS document. Additional bibliographical listings are revealed using the Document Library Searcher in the Netscape browser GUI.

Note the document **version** is a highlighted link. Selection of one version prompts the Web server to deliver the linked document to the client machine Since the "archived" source is a FrameMaker document then the client machine will invoke the FrameMaker activation process on the client machine allowing the user of the client machine to view the document.

6.2.4 WEB DEVELOPMENT

The success of this Document Management System is based upon its use to be measured by an increase in the number of registered documents coupled by a simultaneous

population of the public document archive. The Web development process involves publicizing the Web's availability and purpose, monitoring its usage, and making sure the planning and design processes continue for new features and information.

7. FUTURE DEVELOPMENT

The Document Management System will be enhanced to support the following features:

- Integrated software-based search engines offering word searches of archived documents
- Assembly and delivery of software process engineering document templates.
- Integration of document archive with an optical CD based archive mechanism. This moves more frequently accessed documents on-line for quicker access.
- Motorola document registration and access available through the World Wide Web.
- DMS installation at other Motorola sites.

8. CONCLUSION

Preliminary reviews and comments received from Motorola users of the Document Management System suggest that the World Wide Web concept can significantly reduce document search time and provide some relief to knowledge starvation. The Web, when integrated with a solid software configuration management environment may offer its users a 10X solution to the document access and delivery problem.

DMS significantly reduces document maintenance overhead by automating a document management process. Data validation and status accounting maintain the integrity of the local internally spun Web.

9. ACKNOWLEDGMENTS

The computer network administrative support of the CAE organization was critical to the rapid deployment of the World Wide Web.

Mark Kessler and Glen Fullmer of the Motorola CAE group provided technical assistance toward an efficient use design of the Web application.

Cathy Croll and Robert Busser encouraged and supported the definition and implementation of DMS for the Motorola PPG organization. Jeff Burke, a computer science graduate student, and summer intern, implemented the current DMS installation.

10. REFERENCES

[1] **Hymel, Daniel**, 1994. *Configuration Management Of Engineering Documents*: Proceedings Motorola Software Engineering Symposium, August 8-10, 1994, Chicago, Illinois. pp 73-83.

[2] **December, John and Randall, Neil**. *The World Wide Web Unleashed*. SAMS Publishing. Indianapolis, Indiana. 1994.

TRADEMARKS

All terms mentioned in this document that are known to be trademarks or service marks have been appropriately capitalized. Use of a term in this document should not be reagrded as affecting the validity of any trademark or service mark.

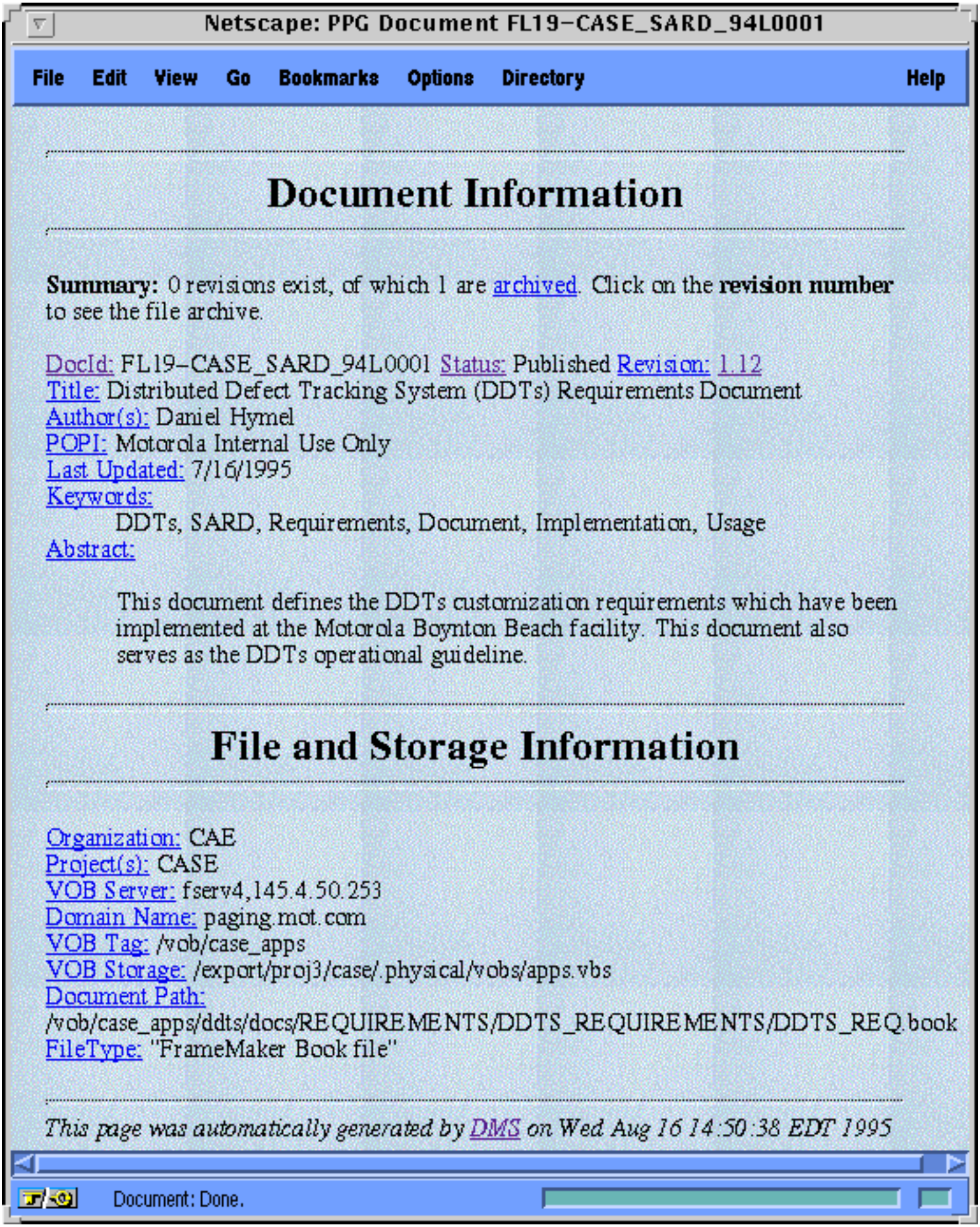

Figure 6 - DMS Card Catalogue

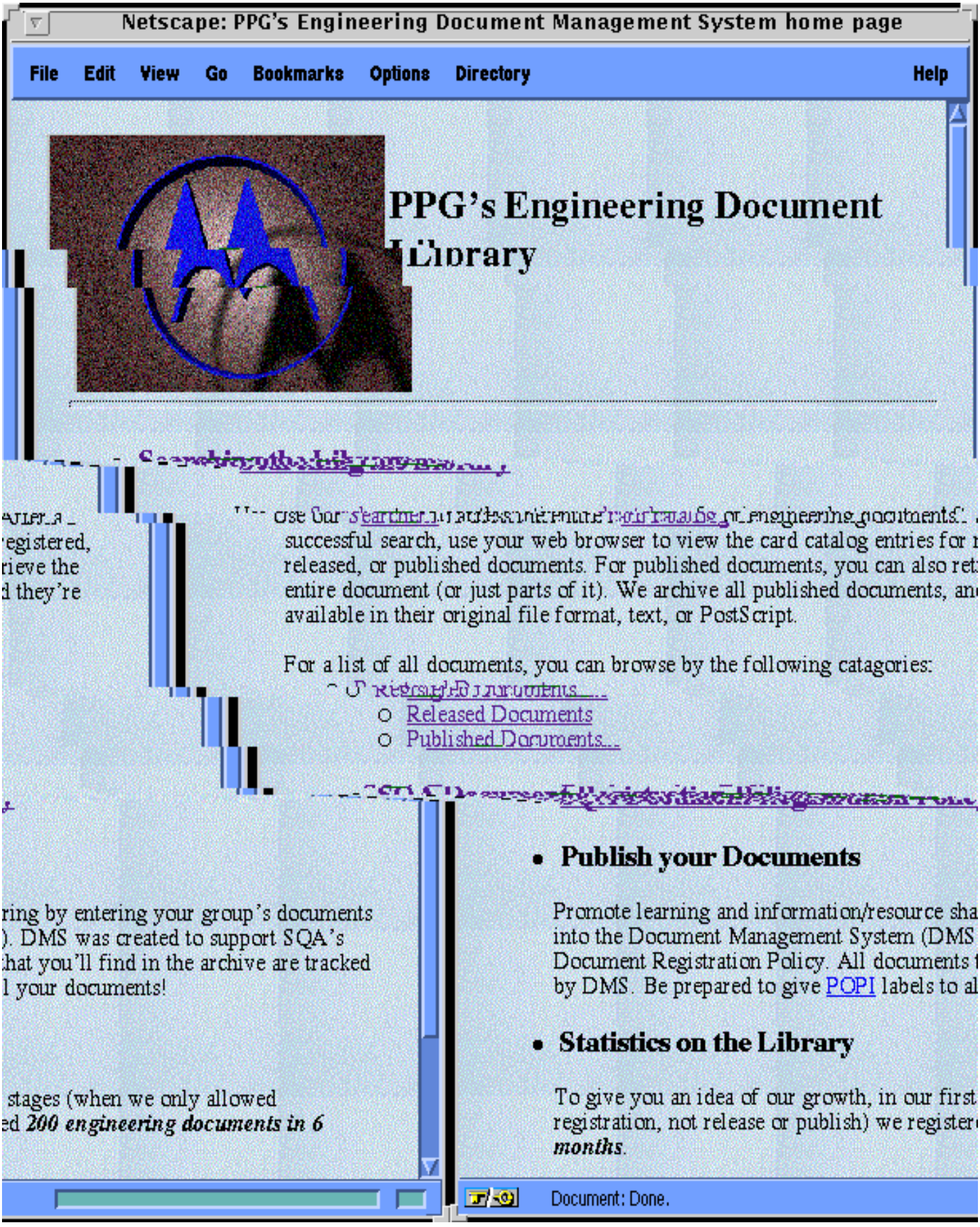

Figure 7 - DMS Web Home Page

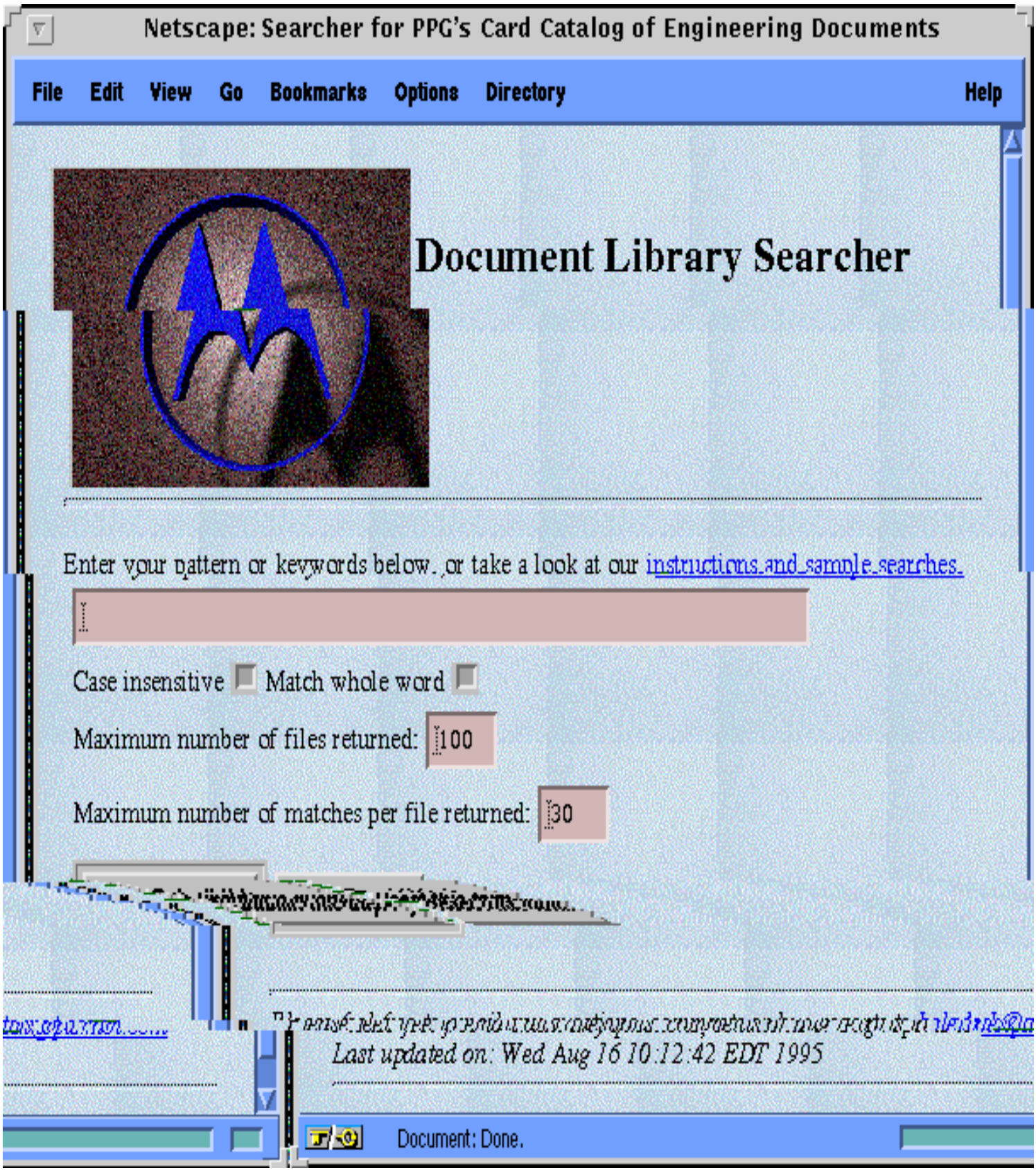

Figure 8 - DMS Library Searcher