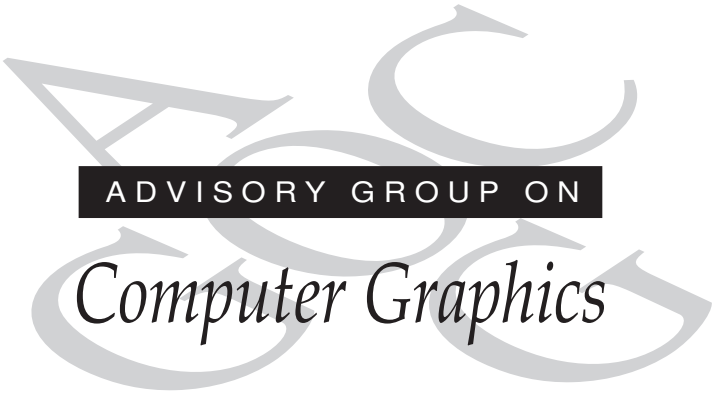
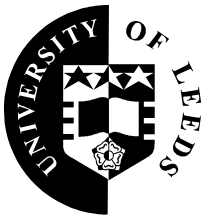


Running a World-Wide Web Service

Brian Kelly

Part of the JISC New Technologies Initiative

SIMA Report Series
Number 6 **ISSN 1356-5370** January 1995



Running A World-Wide Web Service

This handbook gives an introduction to running a World-Wide Web service.

Author Brian Kelly
University Computing Service
University of Leeds

Date February 1995

Edition 1.2

Changes Minor changes have been made to this edition.

About This Handbook

This Handbook was produced using Word For Windows version 2. The graphics were captured using Paintshop Pro. Paintshop Pro was also used to reduce the colour depth and to alter the colour of the images so that they were more suitable for inclusion in the printed version of the Handbook.

The Handbook was converted to HTML format using the RTFtohtml conversion programs.

About The Author

Brian Kelly is the Head of User Support, Computing Service, University of Leeds. He first came across the World-Wide Web (WWW) at a workshop on Internet tools organised by the Information Exchange Special Interest Group, University of Leeds on 9th December 1992. In January 1993 the Computing Service installed the CERN httpd server on its central Unix system - this was probably the first WWW service provided by a central service in the UK academic community.

Following an unannounced visit from Robert Cailliau, one of the WWW co-developers from CERN in March 1993, the Computing Service became convinced of the importance of WWW. The Computing Service contribution to the University Open Day, held in May 1993, was centered on the World-Wide Web: for example the Open Day programme was available on WWW.

Brian has given presentations about WWW at the universities of Aberdeen, Bangor, Bradford, Kent, Oxford, Sussex and Manchester Metropolitan University. He gave a poster presentation at the first WWW '94 conference in Geneva and gave a paper on *Becoming An Information Provider on the World-Wide Web* at the INET 94 / JENC 5 conference in Prague in June 1994. He ran a WWW Tutorial at the Network Service Conference in London in November 1994. He was cochair of the "*WWW A Strategic Tool For UK Higher Education*" workshop.

Acknowledgments

I would like to thank the following for their assistance and comments on this handbook:

Bruce Altner, Nigel Bruce, John D Lewis, Chris Lilley, Jim Hobbs, Ken Hensarling, Roger Horton, Jon Knight, Inke Kolb, Martijn Koster, Paul Leclerc, Neal McBurnett, Sean Martin, Eric Morgan, George Munroe, Alan Richmond, Paul Sutton, Ton Verschuren, Anne Worden, Bruce Washburn.

The author, of course, accepts responsibility for any errors in this handbook.

Feedback

The author welcomes constructive comments and feedback on this handbook, which should be sent to the email address `B.Kelly@leeds.ac.uk` The author is unable to provide individual assistance.

Copyright

Copyright (C) 1994 by Brian Kelly.

All rights reserved. This work may be copied in its entirety, without modification and with this statement attached. Redistribution in part or with modifications is not permitted without advance agreement from the copyright holder.

Copyright of WWW pages shown in this Handbook belongs to the individual or organisation which created the pages. Any copyright holder who wishes for an image to be removed from this Handbook should contact the author of the Handbook.

Obtaining Copies Of This Handbook

This handbook is part of the SIMA subscription series of reports. For subscription details contact `A.M.Mumford@lut.ac.uk` Additional copies of the handbook can be purchased from the University of Leeds. Please contact `B.Kelly@leeds.ac.uk` for pricing details.

The Handbook is available on the World-Wide Web at a number of locations including `http://www.leeds.ac.uk/ucs/WWW/handbook/handbook.html` Consult the `README` file for further information.

Print Record

This document was printed on 20-Feb-95.

Contents

1 Introduction **1**
- Aim of this Document 1
- Target Audience 1

2 About The World-Wide Web **2**
- History 2
- Client-Server Architecture 2
- Growth In Popularity 4
- Examples of Usage 5
- Terminology 9

3 World-Wide Web Browsers **10**
- Publicly Available Telnet Browsers 10
- Text-Based Browsers 10
- NCSA Browsers 11
- Cello Browser 13
- EINet Browsers 13
- Netscape Browsers 14
- Air Mosaic Browsers 14
- GWHIS Browsers 15
- Other Browsers 15
- Future Developments 16
- Further Information 16
- Conclusions 16

4 HTML **17**
- About HTML 17
- HTML Authoring Tools 18
- Word Processing Tools 19
- Browser Editing Tools 21
- HTML Document Conversion Tools 22
- HTML Quality Tools 23
- Review of HTML Tools 25
- Writing Style 26
- Finding Out More About HTML 26

5 Graphics **27**
- HTML Graphical Tags 27
- Active Maps 27
- Graphical Tools 29
- Appropriate Use Of Graphics 29
- Further Information 30

6 Searching And Indexing **31**
- Robots, Spiders and Worms 32
- Aliweb 32
- SWISH 32
- WAIS 32
- CNIDR Isite 34
- Further Information 34

7 WWW Servers **35**
- Example of Installing A Server On A PC 35
- Server Configuration Files 36
- Caching 38
- Proxy Gateways 39
- Security 40
- Summary of Server Software 40
- Server Strategies 41
- Which Server? 41
- Further Information 42

8 Extending WWW **43**
- External Viewers 43
- Running Client Applications 43
- Server-side Extensions 44
- Forms 46

9 Utilities 49

10 Legal and Ethical Issues 50
Liability 50
Computer Misuse Act 50
Pornography 50
Copyright, Designs and Patents Acts 50
Data Protection Act 51
Equality Of Access To Information 51
Advertising 51
JANET Acceptable Use Policy 51
What Is Your WWW Service For? 52
Further Information 52

11 CWISes And WWW 53
Finding Out More 53

12 Teaching And Learning On WWW 54
Examples of Teaching And Learning On WWW 54
CD ROM Facilities 55
National Resources 56
Further Information 56

13 Collaboration On WWW 57
Asynchronous Systems 57
Synchronous Systems 59
Virtual Conferences 60
Further Information 60

14 Libraries And WWW 61
Example Of A Gateway To A Library Catalogue 61
Resources 62
Finding Out More 62

15 Future Developments 63
Uniform Resource Identifiers 63
New Facilities 63

Appendices

1 Mailing Lists 64
Usenet News 64
CERN Mailing Lists 65
Other Mailing Lists 66

2 WWW Resources 70
WWW Online Resources 70
WWW Icons and Clip Art 70
WWW Conferences 71
Other Resources 71

3 National UK Services 72
Services 72
Directories 72
WAIS Resources 72

4 Conferences On WWW 74
Electronic Conferences and Workshops 74
Reviews of Electronic Conferences 74
TaTOO '95 74

5 References 75
Books 75
Magazines 75

1 Introduction

Aim of this Document

This handbook, *Running A World-Wide Web Service*, has been funded by the Advisory Group On Computer Graphics (AGOCG) through the Support Initiative for Multimedia Applications (SIMA) to provide support for UK academic institutions who wish to run a World-Wide Web service. The objectives of the document are to provide the reader with:

- An understanding of the history of the World-Wide Web (WWW).
- Examples of how WWW is currently being used in a number of areas.
- An understanding of the role of a WWW browser and examples of some of the popular browsers.
- An understanding of the Hypertext Markup Language (HTML) and examples of popular HTML authoring tools, document converters and quality tools.
- An understanding of the use of graphics on WWW and a review of some of the popular graphical tools.
- Examples of some of the popular searching tools which are available on WWW and a review of some of the popular indexing tools.
- An overview of WWW server software and a review of the main functions carried out by a server.
- Examples of extensions to the basic WWW architecture, including the use of external browsers, running client-side and server-side applications.
- An awareness of the legal and ethical aspects of providing a WWW service.

The handbook also gives a number of examples of the use of WWW in the following areas:

- Campus (or community) wide information systems (CWIS).
- Teaching and learning.
- Collaboration.
- Libraries.
- Books and magazines about WWW.

The handbook also provides pointers to a variety of sources of further information including:

- WWW mailing lists.
- WWW resources.
- National UK WWW services.

Target Audience

This document is intended primarily for the UK academic community. It should be suitable for computing service or administrative staff responsible for managing a World-Wide Web service, and for academic staff who wish to run a departmental service.

2 About The World-Wide Web

History

The *World-Wide Web* (which is often referred to as W^3, the Web or, as used in this document, *WWW*) is a distributed multimedia hypertext system. What is meant by this?

Distributed: information on WWW may be located on computer systems around the world.

Multimedia: the information held on WWW can include text, graphics, sound and even video.

Hypertext: access to the information is available using hypertext techniques, which typically involve using a mouse to select highlighted phrases or images. Once a phrase or image is selected it can result in information being retrieved from around the world.

The World-Wide Web was initially developed by Tim Berners-Lee and Robert Cailliau of CERN Laboratories, Geneva to provide an infrastructure for particle physicists throughout Europe to share information. Since the physicists were located in various organisations and used a variety of computer systems and applications software (including various word processing and text markup programs for producing reports) the World-Wide Web was developed using the *client-server architecture*, which ensured cross-platform portability.

Client-Server Architecture

The World-Wide Web is based on the *client-server architecture* which is illustrated in Figure 2-1.

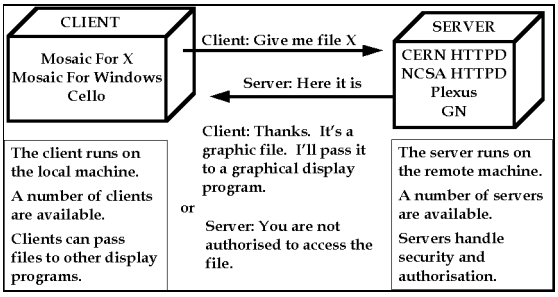


Figure 2-1 WWW Client-Server Architecture.

The end user accesses the World-Wide Web using a browser client, typically on a desktop machine such as a PC, Macintosh or Unix workstation. The client will display hypertext links in some manner, such as underlining the links. Selecting a link (by clicking a mouse button with a graphical client, typing the number following the link using a simple text-based client or using speech or foot pedals, for example, with browsers for disabled users) to send a request over the network (which could be a local network, a national network such as JANET, or over the global network which can be referred to as the Internet). The request is sent to a World-Wide Web server, which typically runs on a powerful computer system. The server will retrieve the file which has been requested and deliver it to the client.

Once the client has started to retrieve the file it can display it on the local machine. If the client cannot display the file (many clients, for example, cannot view video clips) the client can pass the file on to an external viewer which can process the file.

This is a very simple overview of the WWW client-server architecture. Many other features are available: for example the server could send a message to the client, saying that the user is not authorised to access the file. However an understanding of this model will help you to see how the WWW can develop.

Early Browsers

One of the first browsers to be developed was the CERN command line browser. This can be accessed by using the command:

```
telnet telnet.w3.org
```

from a computer system which runs the telnet software. An example of use of the CERN command line browser is illustrated below.

```
telnet telnet.w3.org

                Welcome to the World-Wide Web
THE WORLD-WIDE WEB

This is just one of many access points to the web, the universe of
information available over networks. To follow references, just type the
number then hit the return (enter) key.

The features you have by connecting to this telnet server are very
primitive compared to the features you have when you run a W3 "client"
program on your own computer. If you possibly can, please pick up a client
for your platform to reduce the load on this service and  experience the
web in its full splendor.

For more information, select by number:

A list of available W3 client programs[1]
Everything about the W3 project[2]
Places to start exploring[3]
The First International WWW Conference[4]

This telnet service is provided by the WWW team at the European Particle
Physics Laboratory known as CERN[5]
[End]
1-5, Up, Quit, or Help:
```

Figure 2-2 The CERN Command Line Browser.

Notice that in the CERN command line browser in order to select a hypertext link you need to type the number which follows the link.

The CERN command line browser is a very simple client. The first WWW browser was developed by Tim Berners-Lee, the father of the World-Wide Web, for the NeXT system. However the NeXT hardware was not a commercial success and is no longer manufactured. One of the earliest graphical browsers was the Viola client which was developed for the X windows environment. Viola is illustrated in Figure 2-3.

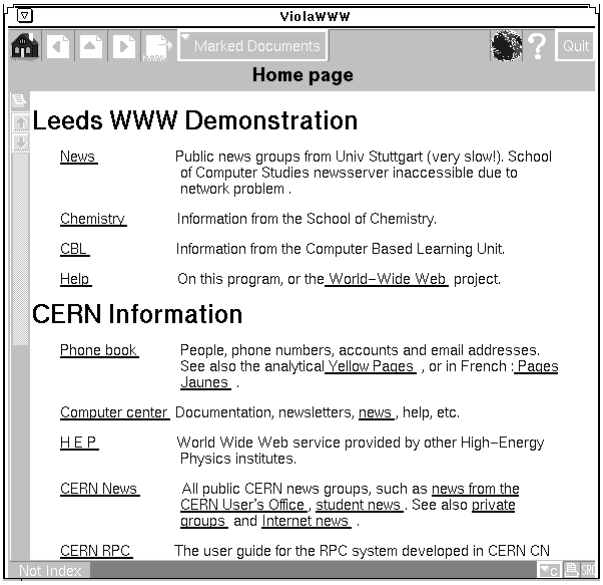


Figure 2-3 The Viola Client.

Notice that in the Viola client the hypertext links are identified by the use of underlining.

Growth In Popularity

As shown in Figure 2-4 use of the WWW has grown tremendously since 1993. This chart, which compares the growth of WWW with a simpler distributed information system known as Gopher, is available at the URL `ftp://ftp.isoc.org/isoc/charts/networks-gifs` (the term URL will be explained later in this chapter). Much of this growth in popularity was due to the release of browsers for the X Windows, PC and Macintosh environments by the National Center For Supercomputing Applications (NCSA) at the University of Illinois.

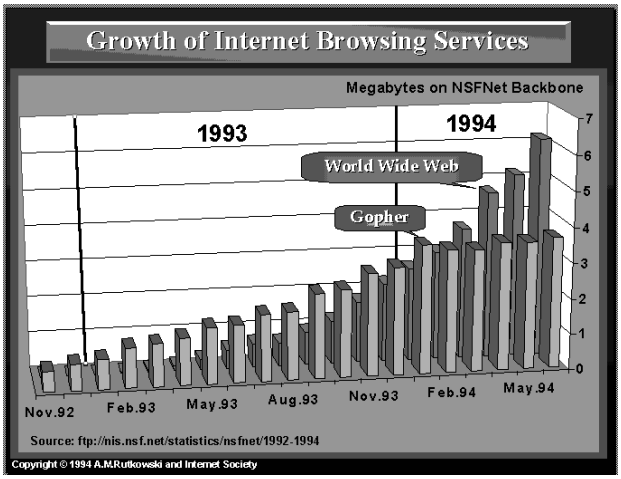


Figure 2-4 Growth In Popularity of WWW.

Since CERN's remit was research in particle physics the WWW development team realised that they needed to involve other organisations in WWW development work. The involvement of NCSA in the WWW development programme resulted in the NCSA Mosaic For X, which was released in early 1993. An illustration of a pre-release version of Mosaic For X is shown in Figure 2-5.

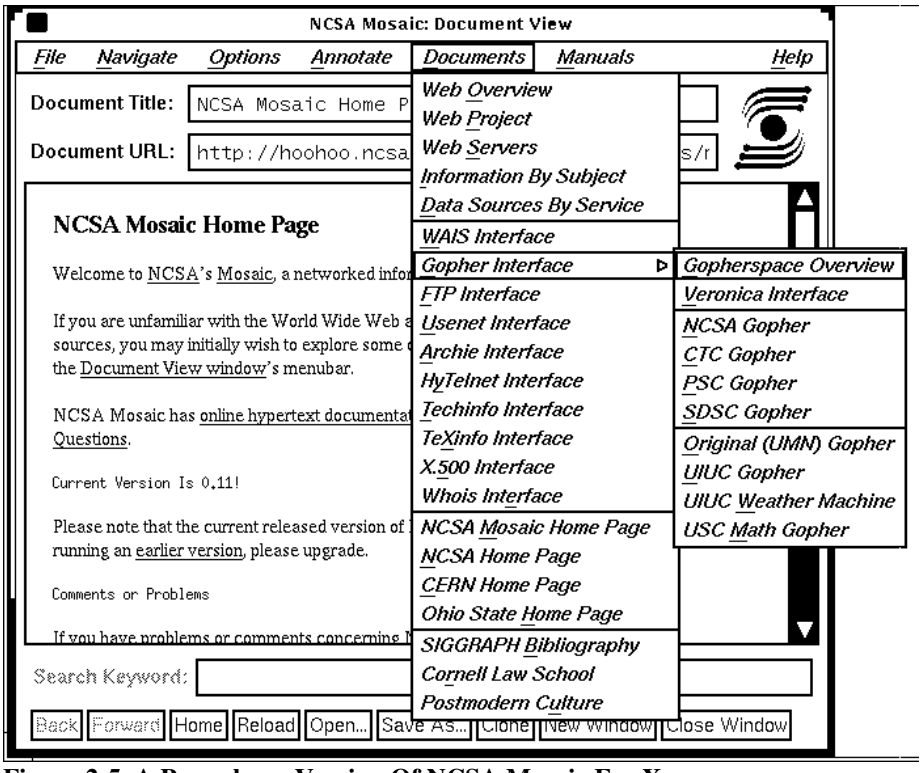


Figure 2-5 A Pre-release Version Of NCSA Mosaic For X.

As can be seen from Figure 2-5 NCSA Mosaic For X provides access to a number of types of resources, including WAIS, Gopher, FTP, Usenet, Hytelnet, TeXinfo, X.500 and Whois resources. NCSA Mosaic was developed by a group of programmers at NCSA led by Marc Andreessen. NCSA Mosaic For X became such a success because:

- It was a professional looking product from its initial release.
- It used the Motif toolkit and was available on workstations platforms widely used in academia.
- It benefited from NCSA's expertise in scientific visualisation.
- It introduced inline images.

In November 1993 NCSA released versions of Mosaic for the Microsoft Windows and Apple Macintosh environment. These browsers, which are freely available to the academic community, provided access to WWW for people who did not have access to Unix and X-Windows systems.

Examples of Usage

A number of examples of how the World-Wide Web is currently being used are given below. These are just a few examples of the many thousands of WWW serices which are currently available.

Publishing Research Information

Figure 2-6 illustrates how CERN (the European Particle Physics Laboratory) makes its technical papers available on the World-Wide Web. The URL for the paper illustrated is `http://www1.cern.ch/ALICE/ENGINEERING/engineering.html`

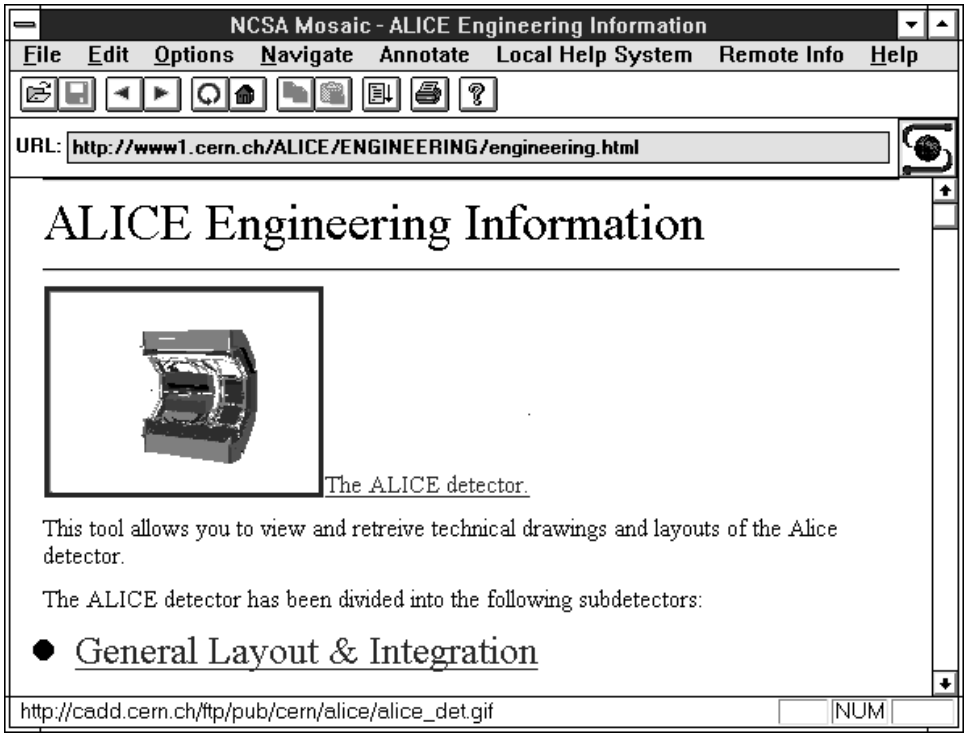


Figure 2-6 Example of Scientific Information Held At CERN.

Campus Wide Information Systems

The Honolulu Community College Campus Wide Information System (CWIS) was the first multimedia CWIS on the World-Wide Web. The URL for this CWIS is `http://www.hcc.hawaii.edu/`

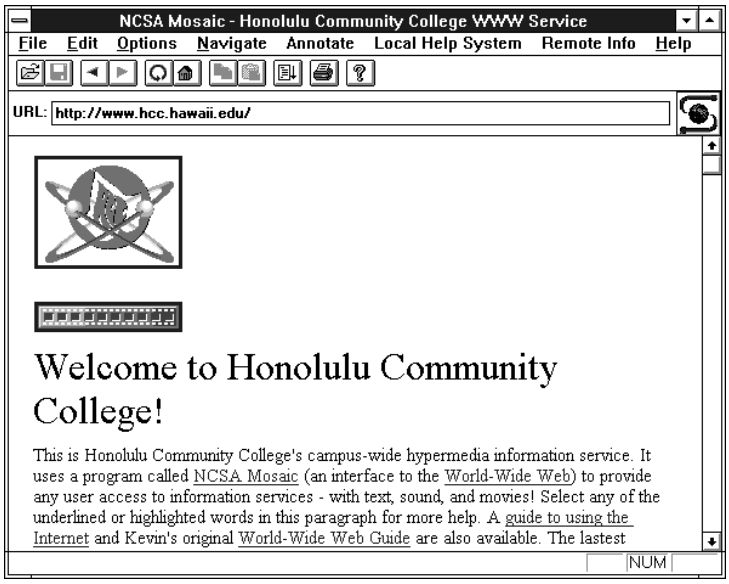


Figure 2-7 The Honolulu Community College CWIS.

Teaching Applications

The Globewide Network Academy (GNA) won a Best of the Web 1994 award for the *Introduction to Object Oriented Programming Using C++* distributed teaching application. The URL for this application is `http://uu-gna.mit.edu:8001/uu-gna/text/cc/index.html`

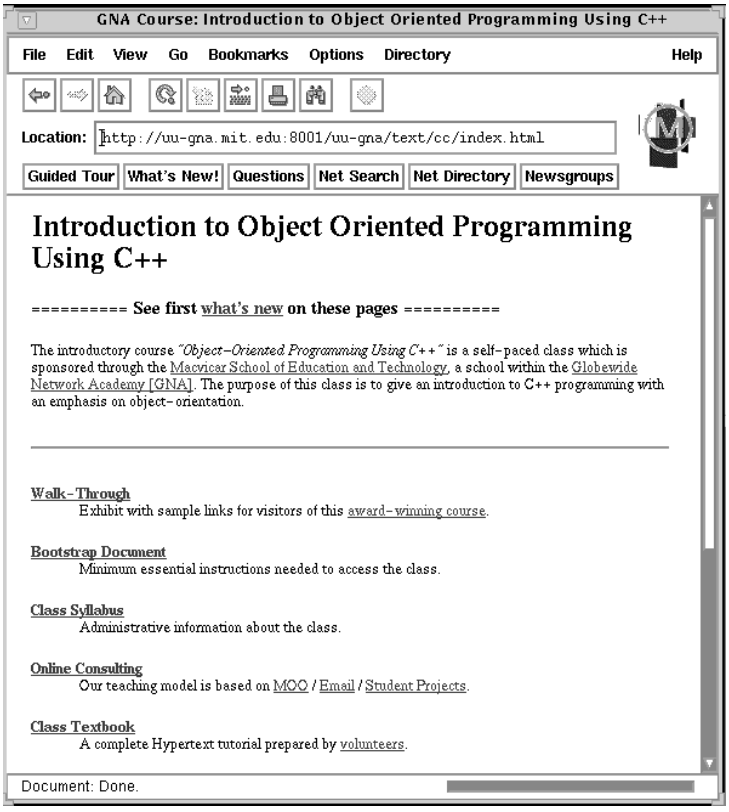


Figure 2-8 A Distributed Teaching Application.

Publicity

The School of Computer Studies at the University of Leeds was one of the first departments to use the multimedia capabilities of WWW to market its courses to potential students. The URL is `http://agora.leeds.ac.uk/WWW/MSc/MSc_text/leeds.html`

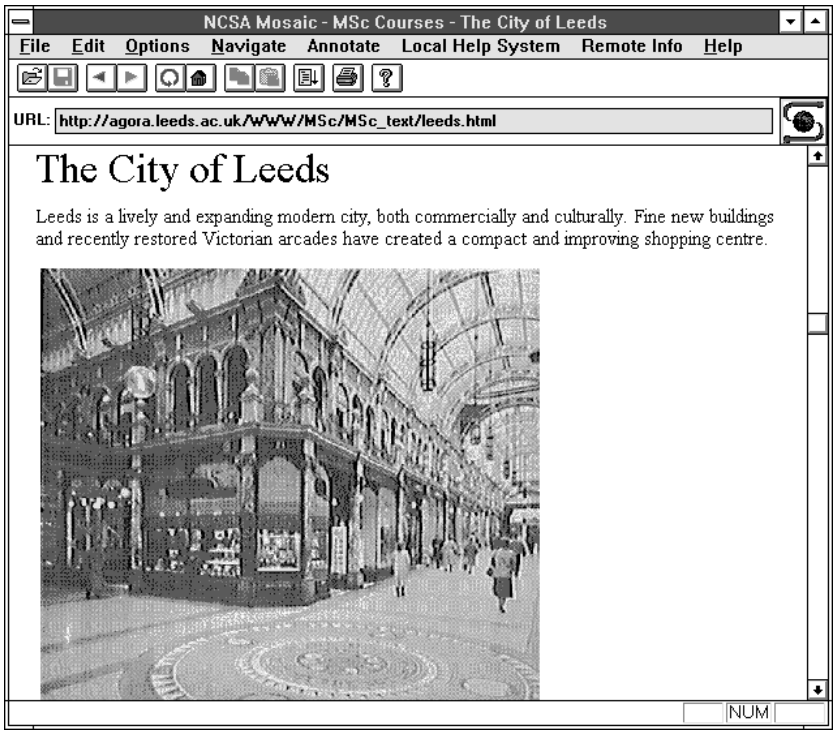


Figure 2-9 University of Leeds Prospectus Information.

Virtual Libraries

Many virtual libraries, art galleries and exhibitions are available on the World-Wide Web. One of the first was the Vatican exhibition. The URL for this virtual exhibition is `http://sunsite.unc.edu/expo/vatican.exhibit/Vatican.exhibit.html`

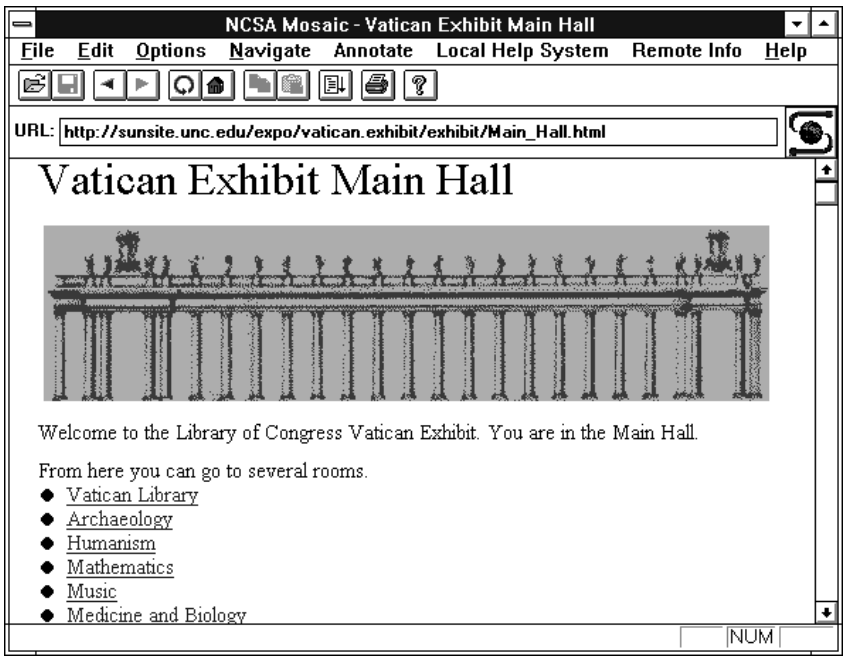


Figure 2-10 The Vatican Exhibition.

Commercialisation Of WWW

The World-Wide Web is increasingly being used by commercial companies. For example the URL for the Pizza Hut ordering service is `http://www.pizzahut.com/`

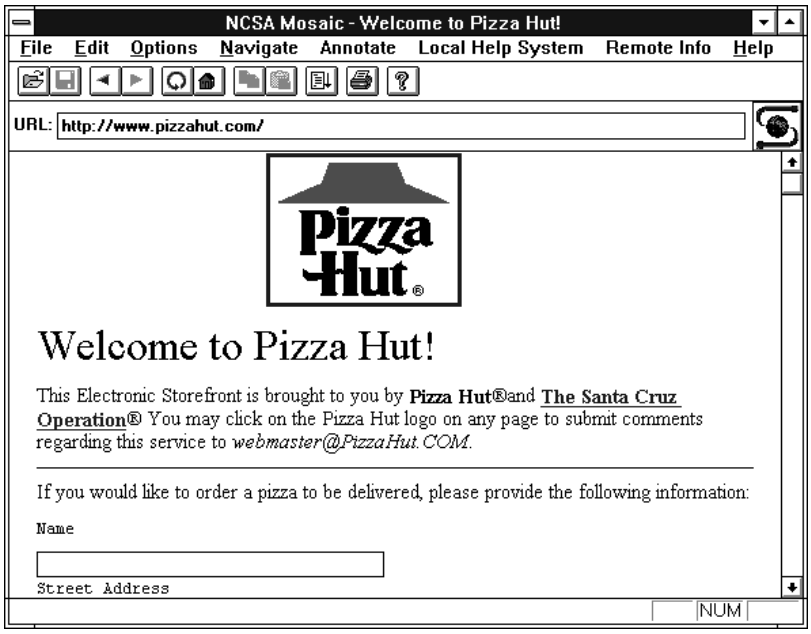


Figure 2-11 A Commercial Application On WWW.

Government Use Of WWW

The World-Wide Web is also being used by governmental agencies. For example the URL for the CCTA is `http://www.open.gov.uk/`

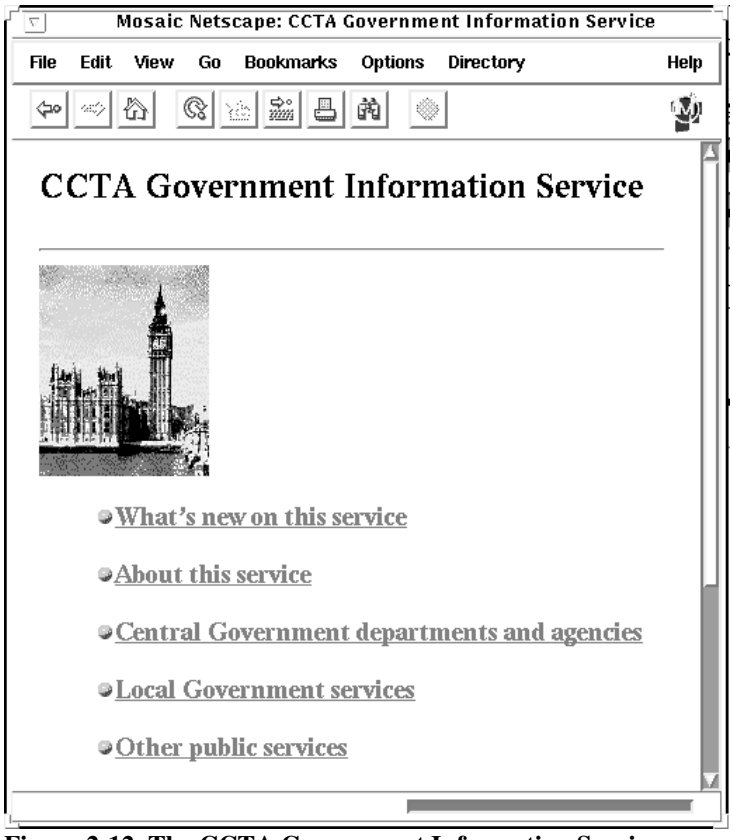


Figure 2-12 The CCTA Government Information Service.

Terminology

The following terms are used in this document:

Browser: An interactive program which is used to access information held on the World-Wide Web.

Client: Often used as a synonym for *browser*. A client is the software which normally runs on the local desktop machine (such as a PC, Apple Macintosh or Unix workstation). The client sends requests to the server software.

Server: Software which is used to deliver information to a *client*. Note that this term can also refer to the computer system on which the server software is running.

URL: Uniform Resource Locator. Can be regarded as the address of a file on the World-Wide Web. It includes the protocol (rules) for retrieving the file, the domain (name) of the computer system on which the server software runs and the file name to be retrieved. For example the URL `http://www.cern.ch/hypertext/WWW/TheProject.html` uses the `http` protocol to retrieve the file `TheProject.html` in the directory `hypertext/WWW/` from the computer called `www.cern.ch`

HTML: Hypertext Markup Language. The native language for documents held on the World-Wide Web. HTML is an SGML (Standard Generalised Markup Language) application.

HTTP: Hypertext Transport Protocol. The protocol (set of rules) used to define the communications between the client and WWW server software.

Note that these terms are, for reasons of clarity, in some cases over-simplified.

3 World-Wide Web Browsers

In order to access the World-Wide Web you will need to use a browser (or client). A wide range of clients are available for many different platforms: although the Mosaic client is very popular you should not think that Mosaic **is** the World-Wide Web.

Publicly Available Telnet Browsers

A number of browsers are publicly available which can be accessed using the *telnet* protocol. These include:

telnet.w3.org

dir.mcc.ac.uk

rsl.ox.ac.uk

These browsers can be accessed by giving the command **telnet** *address* (for example **telnet dir.mcc.ac.uk**) In some cases you will automatically be logged in, in other cases you must enter a username (which is often `lynx`).

An example of the use of the telnet browser at the Radcliffe Science Library at Oxford University is illustrated in Figure 3-1.

```
telnet rsl.ox.ac.uk

         Radcliffe Science Library & Bodleian Library WWW Server (p1 of 6)

     RADCLIFFE SCIENCE LIBRARY & BODLEIAN LIBRARY WWW SERVER

             UNIVERSITY OF OXFORD

   [IMAGE]

Welcome! At present this WWW server is still feeling its way. This
page is intended primarily as a starting point for Oxford users
wishing to explore Internet services and information sources. From
this home page you can also access some of our Local WWW applications
which are for the most part still under development. For newcomers to
the Web, one good introduction is Entering the World-Wide-Web: A Guide
to Cyberspace by Kevin Hughes. Another is CERN's WWW FAQ (list of
Frequently Asked Questions).

Apologies to our regular Lynx users. We have phased out the old Lynx
opening page and you will now commence with this one. If you would
like to voice your opinions or your feelings, please feel free to use
the comments form below.

_____________________________________________________________________
-- press space for more, use arrow keys to move, '?' for help, 'q' to quit
```

Figure 3-1 The Client At Radcliffe Science Library.

It should be noted that the organisations running these publicly available clients do not guarantee to provide the service on a long term basis.

Text-Based Browsers

The browser illustrated in Figure 3-1 is a text-based browser (which is sometimes referred to as a command-line browser). Text-based clients run on a text-based operating system environment (e.g. DOS rather than Microsoft Windows, or Unix rather than X Windows). Command line clients place less demands on the local computer system, but do not provide the ease-of-use or range of functionality provided by graphical clients.

Lynx

The most widely-used text-based browser is probably Lynx. Lynx was developed at the University of Kansas, originally for Unix. An example of the Unix implementation is illustrated in Figure 3-1.

Lynx has been ported to the MS DOS environment. DosLynx, as the implementation is known, will run on a PC with 512 K of RAM, running MS DOS 3 or later. It provides access to the World-Wide Web from an entry level PC which has the appropriate networking capability. DosLynx is illustrated in Figure 3-2.

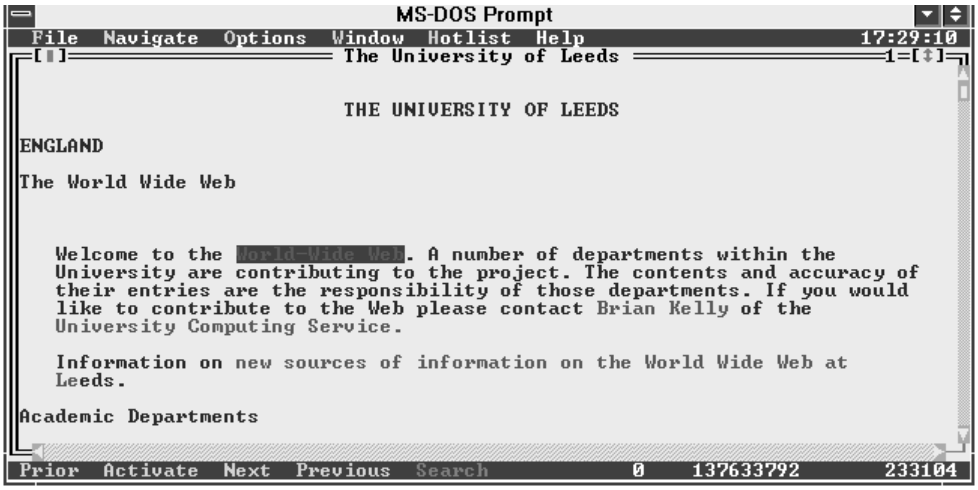


Figure 3-2 DOS Lynx.

Availability

The Lynx browser software is available at the URL `ftp://ftp2.cc.ukans.edu/pub/Web/` In the UK it is also available at the URL `ftp://src.doc.ic.ac.uk/packages/WWW/lynx`

The DosLynx software is available at the URL `ftp://ftp2.cc.ukans.edu/pub/WWW/DosLynx/`

Details of the system requirements for DosLynx are available at the URL `ftp://ftp2.cc.ukans.edu/pub/WWW/DosLynx/readme.htm`. A Listserv mailing list exists at the address `DosLynx-Dev@ukanaix.cc.ukans.edu` for the distribution of DosLynx related information, updates and development discussions. To subscribe send an email request to `listserv@ukanaix.cc.ukans.edu` to be added to the list. All new releases will be announced on this list.

NCSA Browsers

The NCSA Mosaic browser is available for the X Windows, Microsoft Windows and Apple Macintosh environments.

NCSA Mosaic For X

Although it was not the first graphics browser, NCSA Mosaic For X helped to popularise the Web. At the time of writing version 2.4 is available, although a beta version of 2.5 is also available (which includes support for a number of new features including hierarchical hotlists).

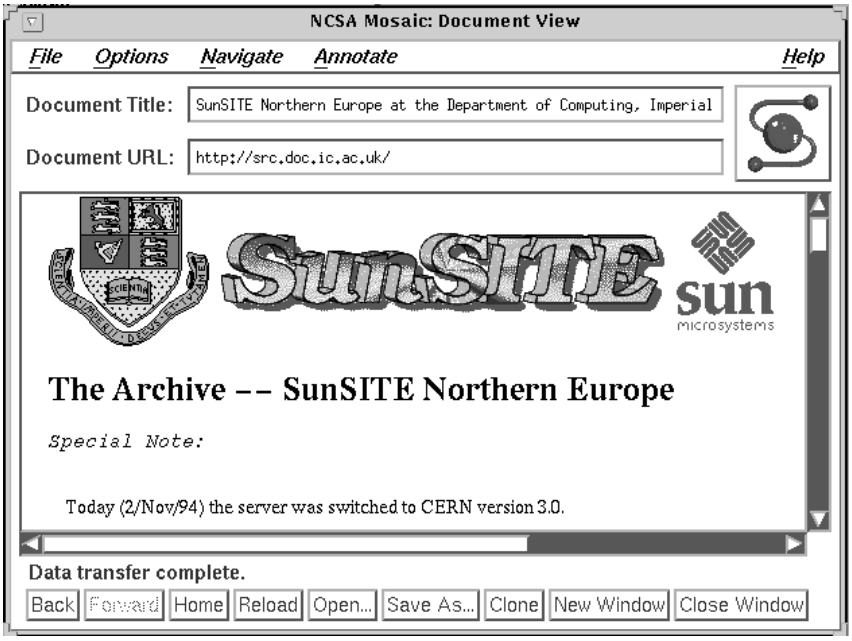


Figure 3-3 NCSA Mosaic For X.

NCSA Mosaic For Windows and the Macintosh

If NCSA Mosaic For X helped to popularise the Web, NCSA Mosaic For Windows and for the Macintosh made it available to a much larger number of people.

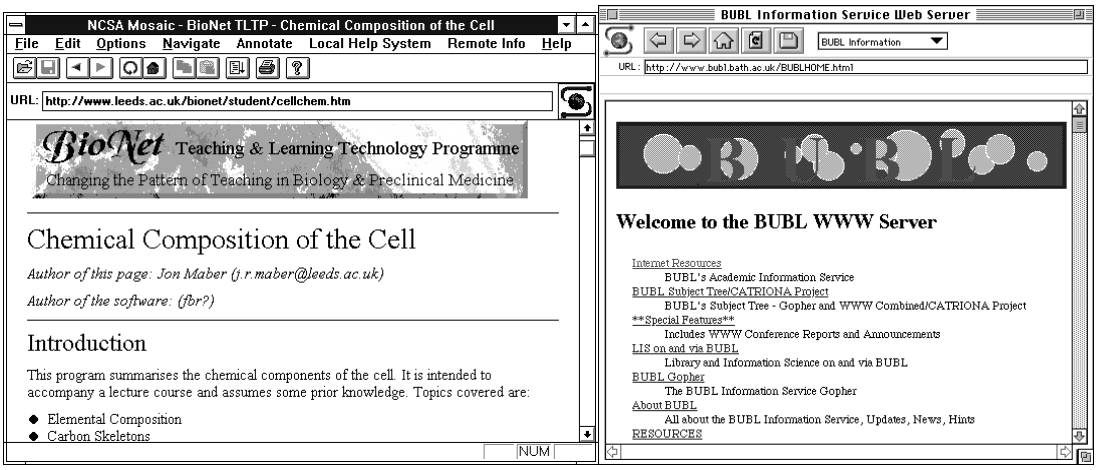


Figure 3-4 NCSA Mosaic For Windows and the Macintosh.

Availability

The NCSA Mosaic browser software for the X, Microsoft Windows and Apple Macintosh platforms is available at the URL `ftp://ftp.ncsa.uiuc.edu/pub/Web/` In the UK it is available at the URL `ftp://src.doc.ic.ac.uk/packages/WWW/Mosaic/`

Further information about NCSA Mosaic For Windows is available at the URL `http://www.ncsa.uiuc.edu/SDG/Software/WinMosaic/HomePage.html` Further information about NCSA Mosaic For the Macintosh is available at the URL `http://www.ncsa.uiuc.edu/SDG/Software/MacMosaic/` Further information about NCSA Mosaic For X is available at the URL `http://www.ncsa.uiuc.edu/SDG/Software/XMosaic/`

Cello Browser

Cello was one of the first WWW browsers to be developed for the Microsoft Windows environment. It was written by Thomas R Bruce of the Legal Information Institute, Cornell University.

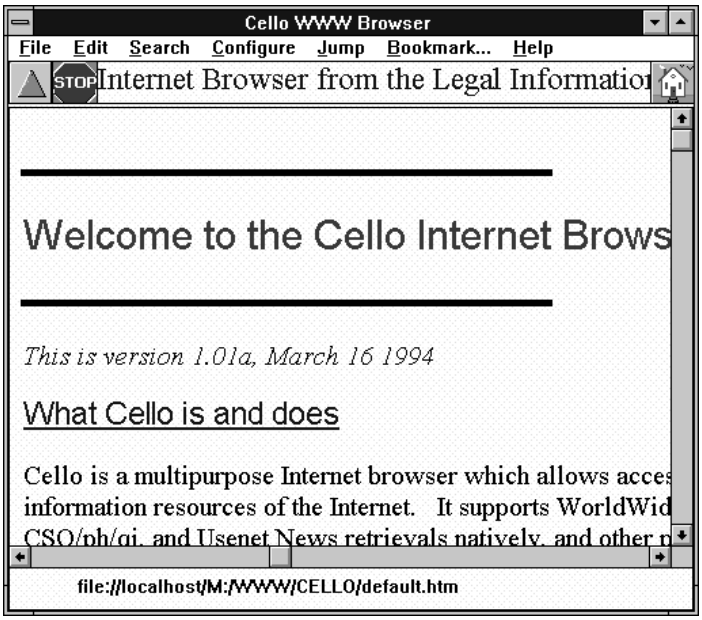


Figure 3-5 The Cello Browser.

Availability

The Cello browser software for Microsoft Windows is available at the URL `ftp://ftp.law.cornell.edu/pub/LII/Cello/`

EINet Browsers

EINet have developed the WinWeb and MacWeb browsers for the PC and Apple Macintosh platforms.

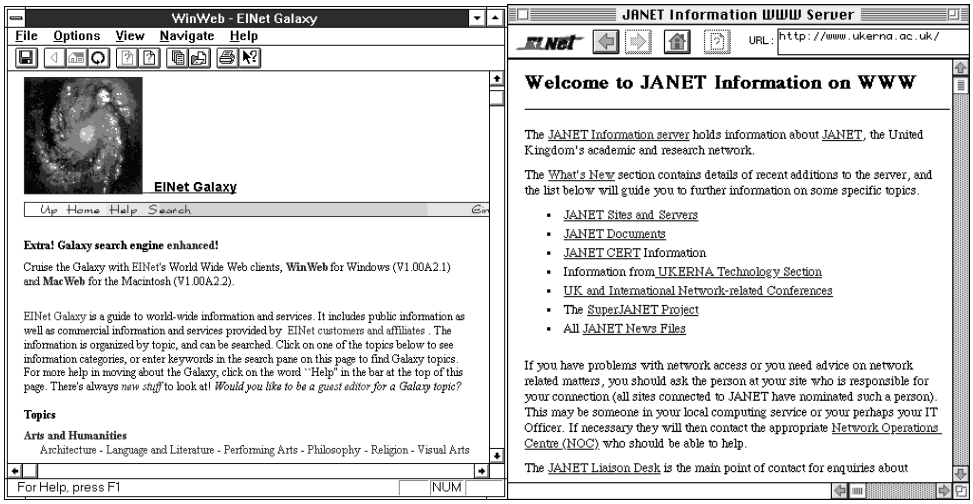


Figure 3-6 The WinWeb and MacWeb Browsers.

Availability

The EINet browsers software for the X, Microsoft Windows and Apple Macintosh environments are available at the URL `ftp://ftp.einet.net/einet/`

Netscape Browsers

Netscape Communications Corporation (MCOM) was set up by Jim Clark, founder of Silicon Graphics. MCOM recruited the developers of NCSA Mosaic to develop a WWW browser. A beta release of Netscape was released in October 1994. It generated a tremendous amount of interest, because of its speed and functionality. However it also caused concern, since it included extensions to the HTML standard which had not been part of the HTML standardisation process.

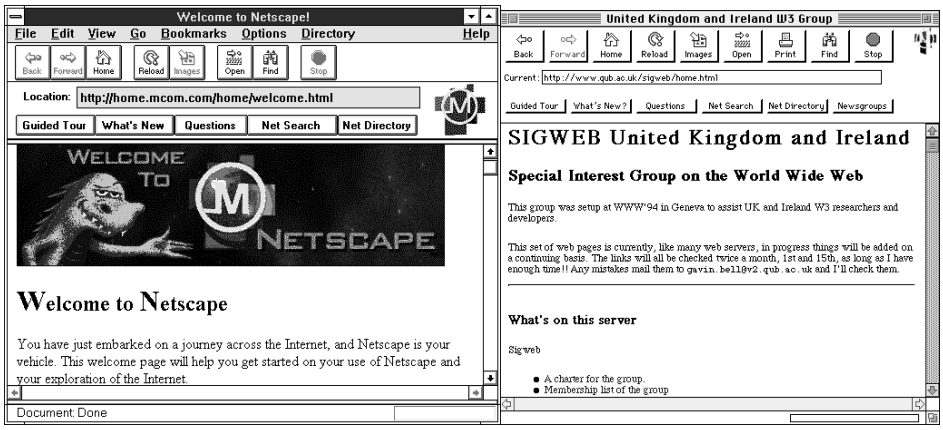


Figure 3-7 The Netscape Browser for Windows and the Macintosh.

Availability

The Netscape browser for the X, Microsoft Windows and Apple Macintosh environments is available at the URL `ftp://ftp.mcom.com/` In the UK it is available at the URL `ftp://src.doc.ic.ac.uk/packages/WWW/Netscape/` Further information is available from the URL `http://home.mcom.com/home/welcome.html`

Air Mosaic Browsers

Air Mosaic is another commercial browser which is based on the NCSA Mosaic source code.

Figure 3-8 The Air Mosaic Browser For Windows.

Availability

An evaluation copy of the Air Mosaic browser software for the X, Microsoft Windows and Apple Macintosh environments is available at the URL `ftp://ftp.spry.com/demo/` Further information is available at the URL `http://www.spry.com/`

GWHIS Browsers

GWHIS is a commercial WWW browser marketed by Quadralay. GWHIS (Global-Wide Help and Information System) consists of a WWW browser, an application program interface (API) for integrating GWHIS into applications and a search engine.

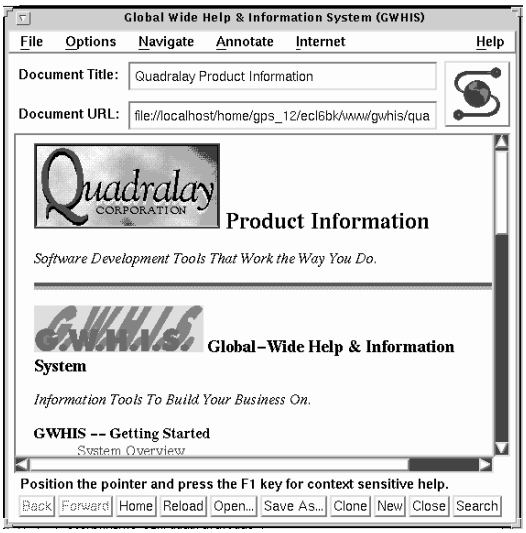


Figure 3-9 The GWIS Browser For X Windows.

Availability

An evaluation copy of the GWHIS browser software for the X, Microsoft Windows and Apple Macintosh environments is available at the URL `ftp://ftp.quadralay.com/pub/gwhis` Further information is available at the URL `http://www.quadralay.com/`

Other Browsers

Many other browsers are available or are currently being developed. Some of the browsers are aimed at the business community. Of particular interest to the academic community are the Internet browsers which are being developed by Microsoft (for inclusion with Windows 95), IBM, Apple and Novell.

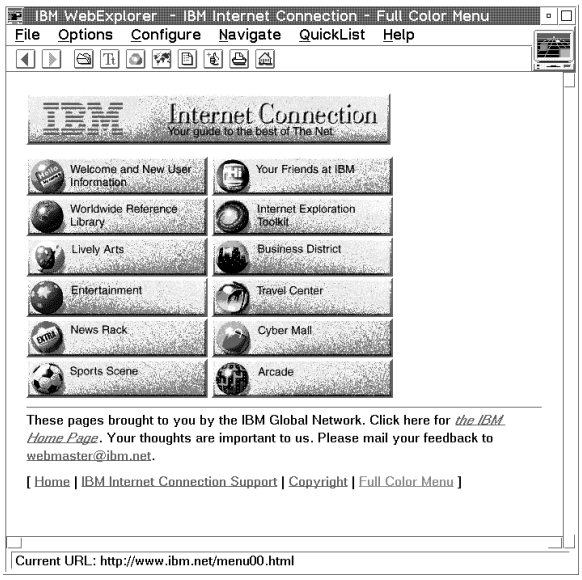


Figure 3-10 The Warp OS/2 Browser.

Future Developments

A browser known as Arena is currently being developed which will handle HTML 3. HTML 3 is a new version of HTML which contains a number of facilities which are not available in HTML 2, including table handling and mathematical formulae.

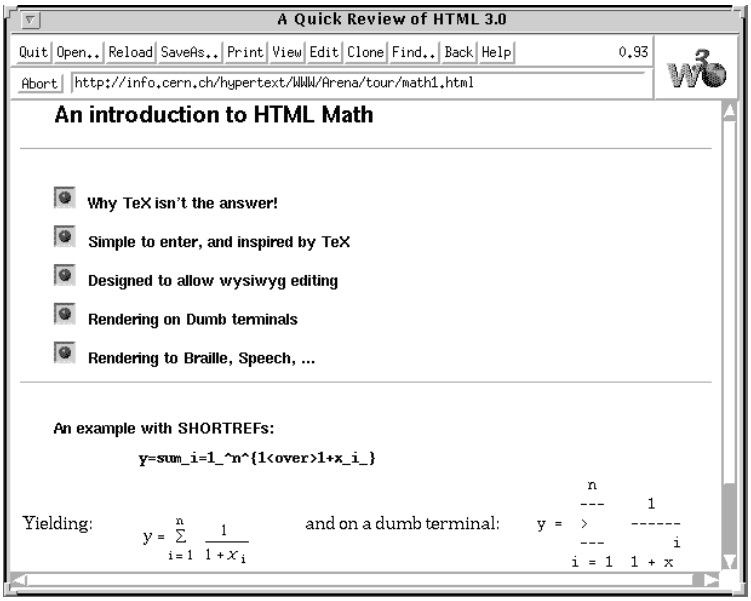


Figure 3-11 The Arena Browser.

Availability

Arena is currently a beta program. It can be obtained from the URL
`ftp://ftp.w3o.org/` In the UK it is available at the URL
`ftp://src.doc.ic.ac.uk/packages/www/arena/`

Further Information

A list of browsers is available at the URL
`http://info.cern.ch/hypertext/WWW/Clients.html`

Another list, which includes a brief summary of known bugs. is available at the URL
`http://www.hotwired.com/browsers.html`

A third list is available at the URL
`http://www.charm.net/~web/Vlib/Users/Clients.html`

Conclusions

Which is the best browser? There is no longer a simple answer to this. The growth in the number of browsers, the different licensing arrangements and different areas they address is making it difficult to adopt an institutional policy on choosing a browser. At the time of writing the Netscape browser looks very attractive. However it was developed primarily to address the needs of commercial users, many of whom requested greater control over the appearance of HTML pages in order to reflect a corporate identity. Will Netscape, however, be as quick to support mathematical equations, which will be of interest to most academic institutes? Will it be the best browser for providing control over external applications - an area which is likely to be of interest to academics who wish to develop distributed teaching materials?

Perhaps the only conclusion to be made at this point is that academic institutions should avoid being locked in to a particular browser.

4 HTML

About HTML

Native documents on the World-Wide Web are written in HTML, the *HyperText Markup Language*. HTML defines the structural elements in a document (such as headers, citations, addresses, etc.), layout information (bold and italics), the use of inline graphics together with the ability to provide hypertext links.

A simple HTML document is illustrated in Figure 4-1.

```
<TITLE>The World-Wide Web</TITLE>
<H1>About The World-Wide Web</H1>
<P>The World-Wide Web is a <EM>distributed
multimedia hypertext</EM> system.</P>
```

Figure 4-1 A Simple HTML Document.

Structural elements in the document are identified by start and end tags. For example the <TITLE> and </TITLE> tag is used to specify the title of the document, which is often displayed by a client. The <H1> and </H1> tag is used to define the first level heading. Clients will normally display headers differently from the body text: for example, a graphical client could display the header using a larger or different font, whereas a text-based client could display a header as centred text or in all capitals.

Figure 4-1 also illustrates the <EM> container. Text held in the container (which is defined by the <EM> start tag and the </EM> end tag) will be emphasised in some way. A graphical browser could render the emphased text by displaying it in italics, whereas a browser with audio capabilities for the visually impaired could render the emphasis by a change in the tone of the voice output.

Figure 4-1 also shows the paragraph container. It is important to understand that the <P> tag is part of a paragraph container and is **no longer** a paragraph separator (as many people mistakenly believe). If the </P> is not used the existence of the next <P> tag will imply a </P>. In future versions of HTML it will be possible to specify paragraph attributes: for example <P ALIGN=Centred>.

Although browsers will display the HTML document shown in Figure 4-1, for reasons of performance and upwards compatibility it is strongly recommended that HTML documents contain additional elements including the <HTML>, <HEAD> and <BODY> tags, as shown in Figure 4-2.

```
<HTML>
<HEAD>
<TITLE>The World-Wide Web</TITLE>
</HEAD>
<BODY>
<H1>About The World-Wide Web</H1>
<P>Information about the World-Wide Web is available
<A HREF="http://info.cern.ch/hypertext/WWW/TheProject.html">
at CERN</A>.</P>
</BODY>
</HTML>
```

Figure 4-2 A Simple HTML Document.

The <HTML> container is used to define the extent of the HTML document. Within the HTML document there are two other containers: <HEAD> and <BODY>. The <HEAD> container provides information about the document itself. This can include the title of the document (as illustrated) copyright information, keywords and expiry dates (for use by caching software). It is important to make use of the <HEAD> tag since, for example, an automatic indexing program which wishes to index the title of HTML documents can parse only the

information contained in the <HEAD>container. If the <HEAD>container is not present the entire document may have to be parsed, which will place unnecessary extra load on the server.

Figure 4-2 also illustrates the use of the anchor <A> container. This tag is used to provide hypertext links. In the example the text *at CERN* which is contained between the <A> and </A> tags will be highlighted in some way by the browser. Selecting this highlighted phrase will cause the client to send a request for `http://info.cern.ch/hypertext/WWW/TheProject.html` This request will use the `http` protocol and will be sent to the server running on the system at `info.cern.ch`

HTML Authoring Tools

Initially information providers on the World-Wide Web used standard editors such as vi and emacs to create HTML documents. As WWW grew in popularity authoring tools were developed to assist information providers. This section describes three authoring tools which are available for the Microsoft Windows environment: *HTML Assistant*, *HTML Hyperedit* and *HTMLEd*.

HTML Assistant

HTML Assistant is a simple authoring tool which can be used to create and edit HTML documents. Frequently Asked Questions about HTML Assistant is available at the URL `http://cs.dal.ca/ftp/htmlasst/htmlafaq.html` HTML Assistant is available at the URL `ftp://ftp.cica.indiana.edu/pub/pc/win3/misc` In the UK it is available at the URL `ftp://src.doc.ic.ac.uk/packages/WWW/tools/editing/ms-windows/html-assistant`

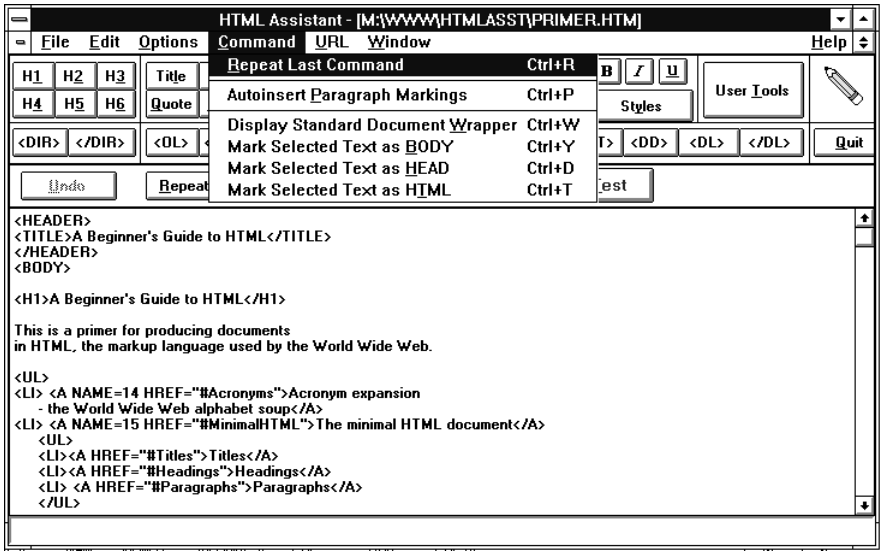


Figure 4-3 HTML Assistant.

HTML Hyperedit

HTML Hyperedit (which was developed using the Toolbook authoring system) not only provides an environment for producing HTML documents, but also contains a tutorial which gives an introduction to HTML. HTML Hyperedit is available at the URL `ftp://info.curtin.edu.au/pub/internet/mswindows/hyperedit` In the UK it is available at the URL `ftp://src.doc.ic.ac.uk/packages/WWW/tools/editing/ms-windows/win-htmledit`

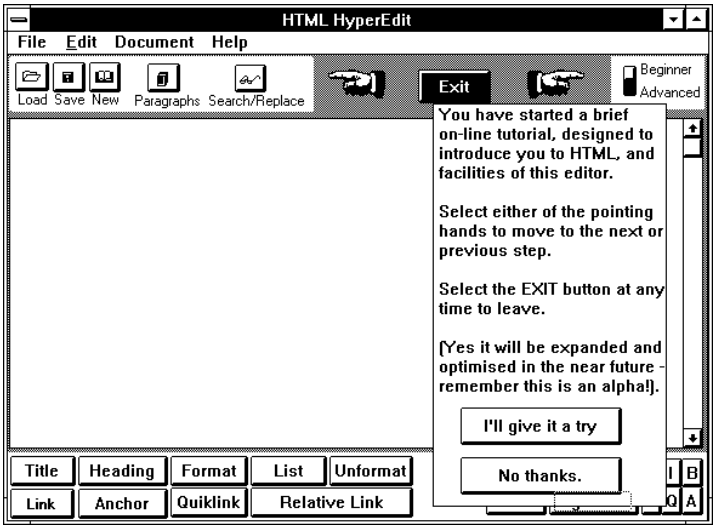


Figure 4-4 HTML HyperEdit

HTMLEd

HTMLEd is a simple authoring tool which can be used to create HTML documents. In the UK it is available at the URL
`ftp://src.doc.ic.ac.uk/packages/WWW/tools/editing/ms-windows/`

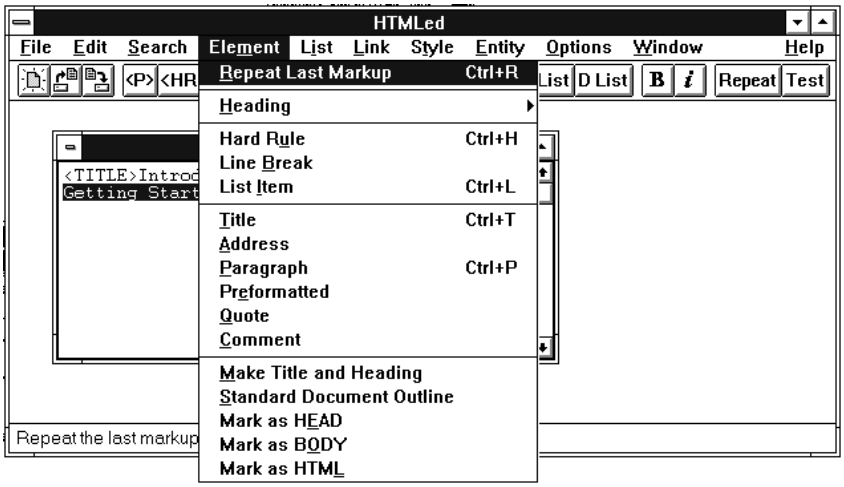


Figure 4-5 HTMLEd.

Word Processing Tools

HTML Assistant and HTML Hyperedit are self-contained authoring tools. Another approach is to develop authoring tools which work within a word processing environment. These tools are normally implemented as macros for popular word processing packages, such as Word For Windows or WordPerfect. This section describes three tools which have been developed for use within Word For Windows: the *GT_HTML*, *CU_HTML* and *ANT_HTML* macros.

Word processing tools have the advantage that they provide a consistent environment for existing users of word processors. However they do have their disadvantages. Because they are normally implemented as macros, they can be very slow, especially when used with large or complicated documents. There is also a danger that HTML markup which is embedded as hidden text could cause conflicts with other word processing tools if, for example, the word processed document was used by other users.

GT_HTML

One of the first word processing macros which could be used to create HTML documents was the GT_HTML macro. This macro, written for Word For Windows, was developed at the Georgia Technical Research Institute. In the UK the software is available at the URL `ftp://src.doc.ic.ac.uk/packages/WWW/tools/editing/macros/ms-winword`

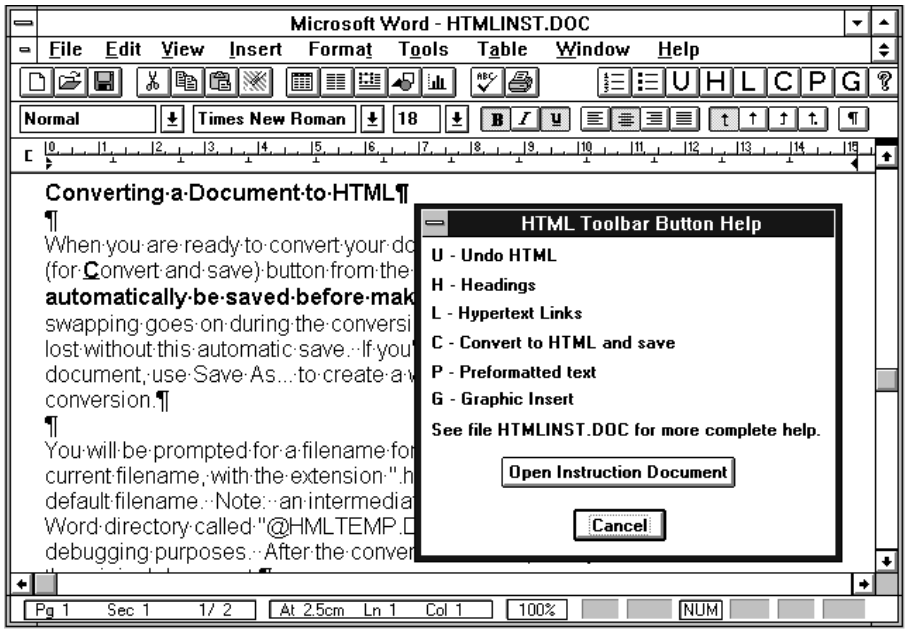


Figure 4-6 The GT_HTML Macro.

CU_HTML

CU_HTML is a template designed to work within Word For Windows. The template was written by Anton Lam (`mailto:anton-lam@cuhk.hk`) The software is available at the URL `ftp://ftp.cuhk.hk/pub/www/windows/util`

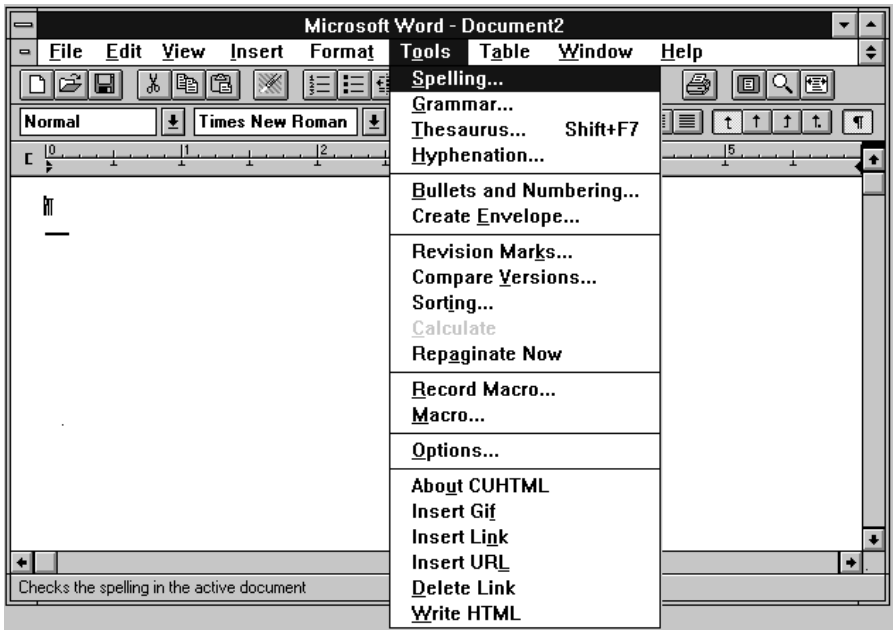


Figure 4-7 The CU_HTML Macro.

ANT_HTML

ANT_HTML is a template designed to work within Word For Windows 6.0. The template was written by Jill Swift (`mailto:jswift@freenet.fsu.edu`) The software is available at the URL `ftp://ftp.einet.net/einet/pc/ANT_HTML.ZIP`

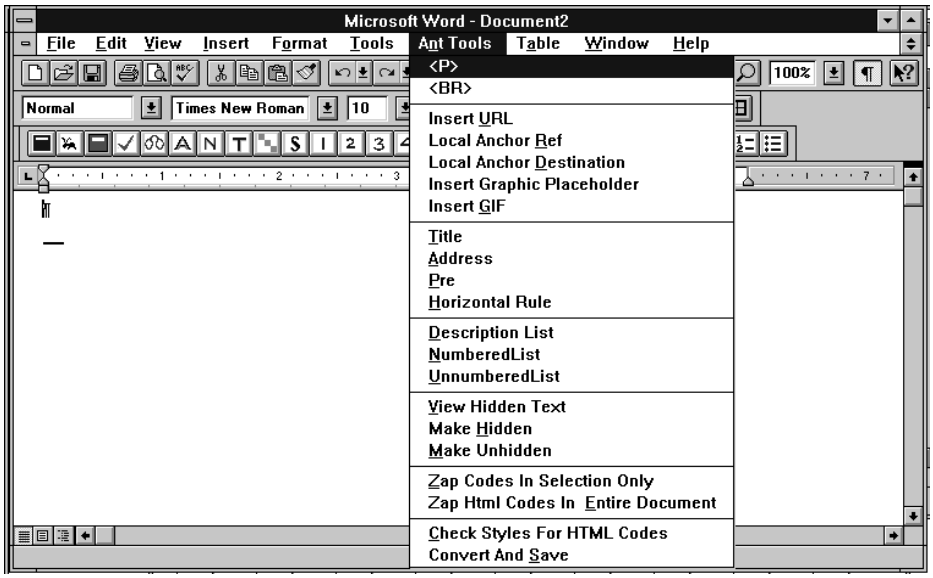


Figure 4-8 The ANT_HTML Macro.

Browser Editing Tools

Another approach to editing HTML documents is provided by browsers which are integrated with editing tools. The Arena browser enables an external editor to be invoked to edit the displayed HTML document. Figure 4-9 illustrates the Arena browser used in conjunction with the Emacs editor.

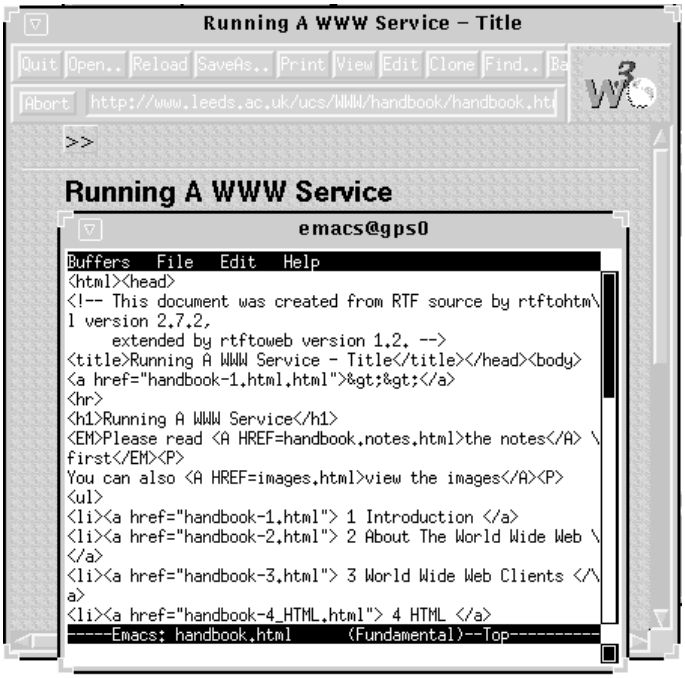


Figure 4-9 Editing A Document From Arena.

HTML Document Conversion Tools

Authoring tools are normally used to create new HTML documents. Document conversion tools, on the other hand, can be used to convert existing documents to HTML format.

LaTeX2html

One of the first sophisticated document conversion tools to be developed was the LaTeX2html conversion program. This program was written by Nikos Drakos, Computer Based Learning Unit, University of Leeds. It and set the standard for document converters, providing a wide range of feature including:

- Automatic generation of hyperlinks at user-defined section breaks.
- Automatic conversion of graphics and formulae to embedded graphics.

Figure 4-10 illustrates a document which has been converted by the LaTeX2html conversion program.

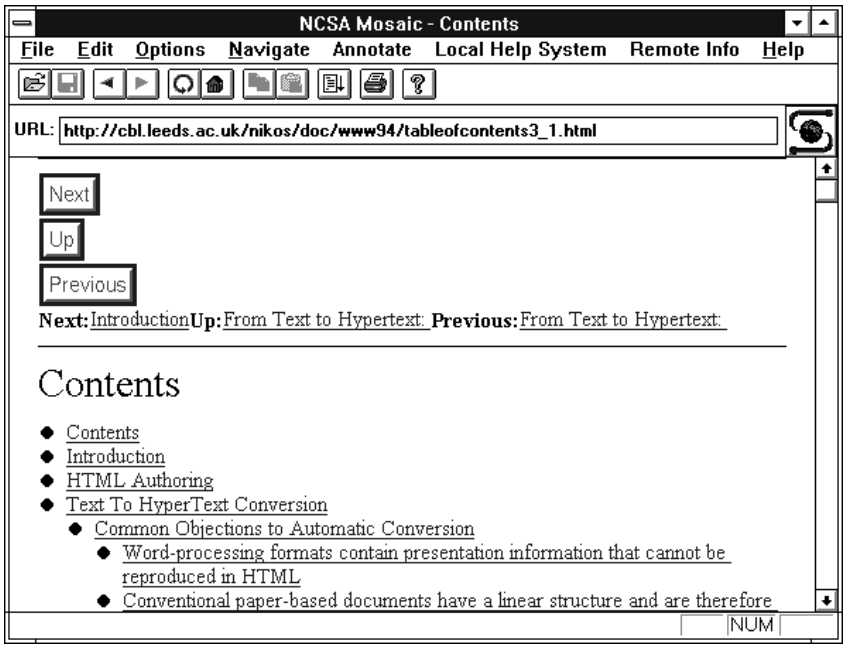


Figure 4-10 A Document Converted Using LaTeX2html.

LaTeX2html is available at the URL
`ftp://src.doc.ic.ac.uk/packages/WWW/tools/translators/latex2html`
Further information is available at the URL
`http://cbl.leeds.ac.uk/nikos/doc/www94/www94.html`

RTFtohtml

The RTFtohtml conversion program enables RTF files (which can be produced by word processing packages such as Word For Windows) to be converted to HTML. The program was written by Chris Hector(Cray) based on RTF parsing software developed by Paul DuBois.

RTFtohtml is available as a command line tool for a number of Unix platforms. In addition an Apple Macintosh implementation is available. A beta version of an MSDOS implementation was announced in November 1994.

An extension of the RTFtohtml program is known as RTFtoweb. This provides a number of additional features, including creation of hypertext links at user defined section breaks. Figure 4-11 illustrates a document on *Exploring The World-Wide Web Using Mosaic For Windows* which is available at the URL
`http://www.leeds.ac.uk/ucs/docs/tut50/tut50.html`

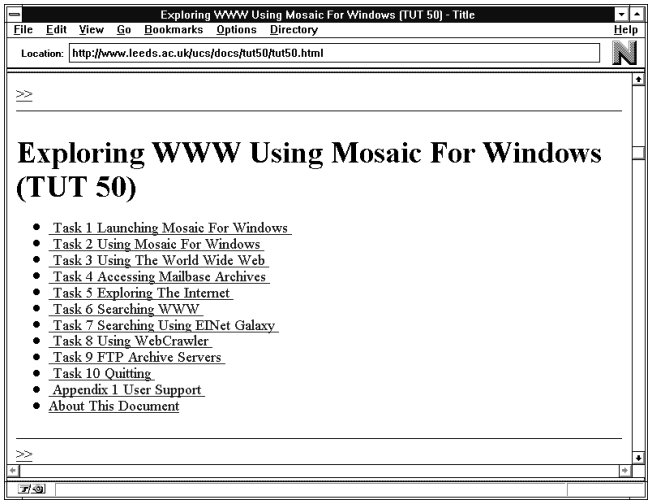


Figure 4-11 Document Converted Using RTFtoweb.

In Figure 4-11 it should be noted that the document is automatically split into a number of files. A hypertext table of contents is automatically generated. Chevrons (>> and <<) are also generated automatically which can be used to move to the next or previous section.

Further information about RTFtohtml is available at the URL `ftp://ftp.cray.com/src/WWWstuff/RTF/rtftohtml_overview.html` The software is available at the URL `ftp://ftp.cray.com/src/WWWstuff/RTF/latest/` In the UK it is available at the URL `ftp://src.doc.ic.ac.uk/packages/WWW/tools/translators/rtftohtml` RTFtoweb is available at the URL `ftp://ftp.rrzn.uni-hannover.de/pub/unix-local/misc/rtftoweb/html/rtftoweb.html`

HTML Quality Tools

The HTML specification states that "*HTML parsers should be liberal except when verifying code. HTML generators should generate strictly conforming HTML.*" Put simply this means that browsers should be capable of displaying documents which contain invalid HTML, but HTML authoring tools and document converters should generate HTML which conforms strictly to the standard.

A number of HTML validation tools are available which can validate HTML documents. A number of popular tools are described below.

HoTMetal

HoTMetaL is an HTML authoring tool and validator. It will provide feedback if it encounters invalid HTML, as illustrated in Figure 4-12.

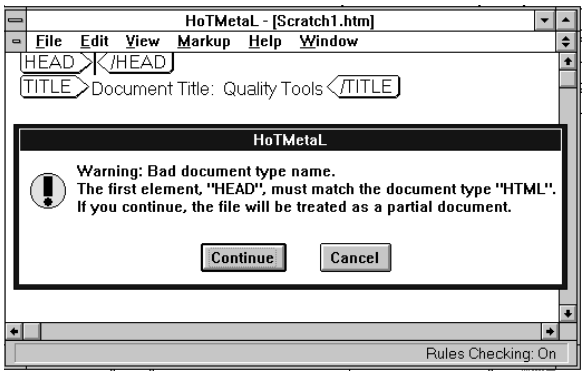


Figure 4-12 HoTMetaL.

HoTMetaL is available for the X and Microsoft Windows. Two versions of the software are available: a public domain version and a licensed version. HoTMetaL Pro, the licensed version, can be used to import and validate an existing document. The public domain version will give an error and refuse to load a document which contains invalid HTML.

HoTMetaL is available at the URL
`ftp://src.doc.ic.ac.uk/packages/WWW/Mosaic/html/hotmetal`

Weblint

A tool called **weblint** can be used to check for invalid HTML documents. This software is available from the URL `ftp://ftp.khoros.unm.edu/pub/perl/www/weblint-1.000.tar.gz` In the UK it is available at the URL
`ftp://src.doc.ic.ac.uk/packages/WWW/tools/weblint`

SGMLS

sgmls is a tool which can be used to validate SGML documents. It is available at the URL `ftp://sgml1.ex.ac.uk/pub/SGML/sgmls/` sgmls is used in a number of HTML validation services, such as those mentioned above. Information on installing sgmls and also **pgmls** (an SGML mode for emacs) is available at the URL
`http://web.nexor.co.uk/users/mak/doc/html/sgml-lib/html-sgml.html`

HTML Validation Service

An HTML validation service is available at the URL
`http://www.hal.com/%7Econnolly/html-test/service/validation-form.html` This service makes use of HTML forms and a CGI script which runs a HTML validation program. The service can be used to check HTML syntax by entering the HTML markup to be checked. It can also be used to check an existing HTML document by entering the URL of the document.

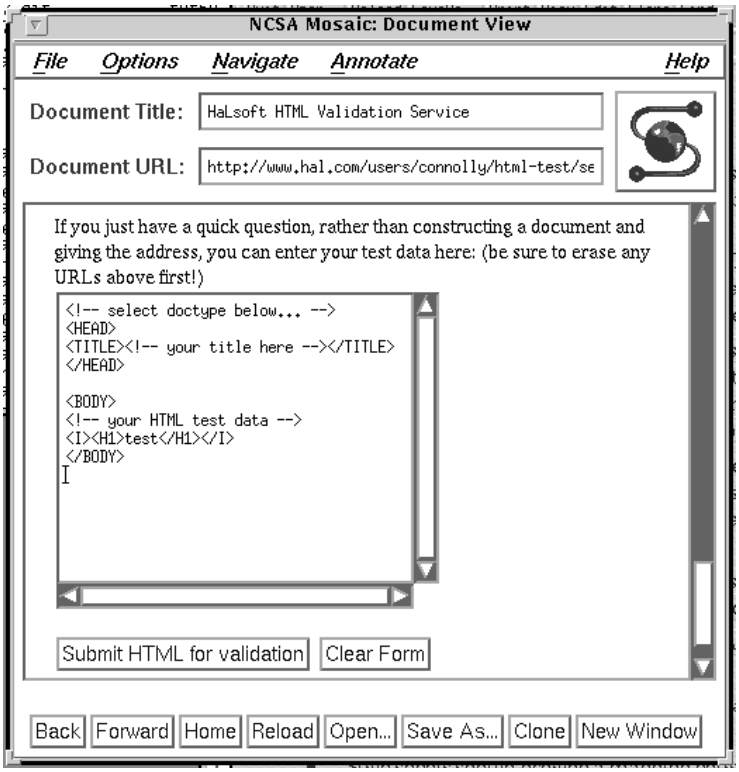


Figure 4-13 HTML Validation Service.

A variation on this service is available at the URL `http://www.cc.gatech.edu/grads/j/Kipp.Jones/HaLidation/validation-form.html`

These services make of the **sgmls** validation program.

The software can be installed on your local Unix system. It is available at the URL `ftp://ftp.hal.com/pub/CGI/check-html.tar.Z`

HTML Check Toolkit

The HTML Check Toolkit is another HTML validation program. The software can be installed using a WWW browser. The installation service, illustrated below, is based on the EIT Webmaster Starter's Kit. HTML Check Toolkit is available at the URL `http://www.hal.com/~markg/HaLSoft/html-check/`

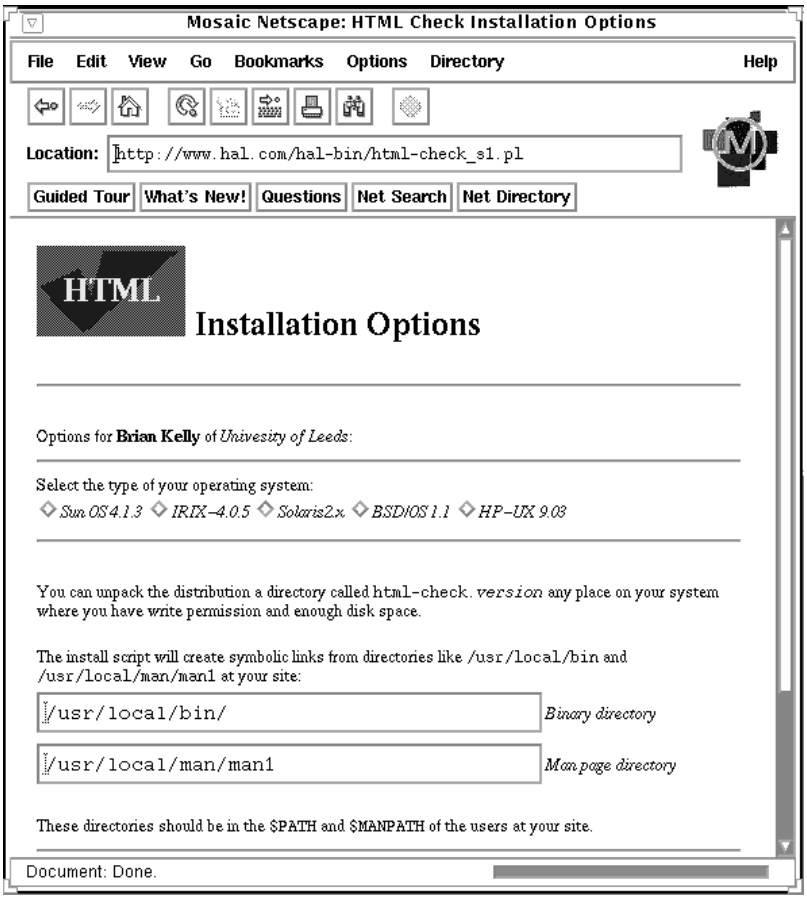


Figure 4-14 Installing The Check_HTML Script.

Review of HTML Tools

Before choosing HTML authoring tools, document converters or quality tools for institutional use the following issues should be considered:

Support Who wrote the software - an experienced software developer or a student as part of a computer project? Will the software continue to be developed and supported?

Quality Does the software produce valid HTML?

Functionality What facilities does the software provide?

Other Issues If the software is based on a word processing package, what happens if the word processed document needs to be used by another word processor?

Writing Style

Writing styles for WWW documents are still developing. However there are a number of guidelines which can be provided:

- Avoid saying "*Click on the option*". This is browser specific: with a command-line browser, such as Lynx, it is not possible to click to select hypertext links.
- Avoid saying "*Select here*".
- Avoid saying "*Back to University Home Page*". If the reader has gone directly to the page, the word back is not appropriate.
- Avoid saying "*University Home Page*". Which University? Try to give the reader some context to where they are in *cyberspace*.

Finding Out More About HTML

This document does not provide an in-depth tutorial on HTML. Many WWW resources are available which give details on writing HTML. Some of these are listed below:

- *A Beginner's Guide To HTML* is available at the URL `http://www.ncsa.uiuc.edu/demoweb/html-primer.html`
- *The Elements Of HTML Style* is available at the URL `http://bookweb.cwis.uci.edu:8042/Staff/StyleGuide.html`
- *HTML Documentation* is available at the URL `http://www.utirc.utoronto.ca/HTMLdocs/NewHTML/intro.html`
- Nathan Torkington's *An Information Provider's Guide To HTML* is available at the URL `http://www.vuw.ac.nz/who/Nathan.Torkington/ideas/www-html.html`
- Marc Andreessen's *HTML Primer* is available at the URL `http://www.ncsa.uiuc.edu/demoweb/html-primer.html`
- Peter Flynn's *Online Guide to HTML* is available at the URL `http://www.ucc.ie/info/net/htmlindex.html`
- *How to Write HTML files* is available at the URL `http://curia.ucc.ie/info/net/htmldoc.html`
- *Spinning the Web: An Introduction to HTML* is available at the URL `http://scholar.lib.vt.edu/reports/soasis-slides/HTML-Intro.html`
- *Composing Good HTML* is available at the URL `http://www.williamette.edu/html-composition/strict-html.html`
- A *Crash Course in HTML* is available at the URL `http://www.ziff.com/~eamonn/crash_course.html`
- An *HTML Tutorial* is available at the URL `http://www.charm.net/~web/Tutorial/HTML/`

In addition to these documents the following resources are also available.

A **review of Microsoft Windows HTML authoring tools** is available at the URL `http://werple.apana.org.au/~gabriel/html-editors/index.html`

A **list of HTML tools** is available at the URL `http://info.cern.ch/hypertext/WWW/Tools/Filters.html`

Dan Connolly's **HTML Design Notebook** is available at the URL `http://www.hal.com/%7Econnolly/html-design.html`

The **HTML specification** is available at the URL `http://www.hal.com/%7Econnolly/html-spec.html`

5 Graphics

The World-Wide Web is, of course, a graphical system. This section describes how graphical objects can be incorporated in an HTML document, how external graphical files can be used and how to create and use interactive maps. The section also considers the performance aspects of using graphics.

HTML Graphical Tags

Inline images are defined in an HTML document using the <IMG> tag. For example:

```
<IMG SRC="portrait.gif">
```

The full syntax for the <IMG> tag is:

```
<IMG SRC="source file" ALT="textual description"
ALIGN="option">
```

The SRC attribute is used to specify the URL of the graphical file. At the time of writing graphical files should normally be in GIF format, although support for other graphical file formats may be available in certain browsers. The SRC attribute is mandatory.

The ALT attribute is used to specify text which should be displayed by a browser which cannot display graphics, or a browser which has the display of inline images option switched off. Use of the ALT attribute is highly recommended.

The ALIGN attribute can take the values TOP, MIDDLE or BOTTOM. It is used to define whether the top, middle or bottom of the graphic should be aligned with the text. Use of the ALIGN attribute is optional.

Using External Viewers

You can use the <A> anchor tag to refer to a graphical file. When the link is selected the graphical file is normally passed to a graphical viewer (such as xv or LVIEW) for displaying.

One common use of the <A> tag is to provide a link to a large colour graphic from a small thumbnail image. For example:

```
<A HREF="full-image.jpeg"><IMG SRC="thumbnail.gif"
ALT="Portrait of John Smith"></A>
```

It is also possible to use this technique to provide links from thumbnail images to video clips. For example:

```
<A HREF="fluidflow.mpeg"><IMG SRC="fluidflow.thumb.gif"
ALT="Video clip of fluid flow"></A>
```

Active Maps

An *active map* (also sometimes refered to as a *clickable image*) is an inline image in an HTML document. An area of the image can be selected, usually by clicking with a mouse. The coordinates of the image that has been selected are sent to a program which can then process the information. An active map can be used to provide a graphical menu, in which selecting a menu option will retrieve a specified HTML document. Active maps can also be used in developing teaching and learning software - for example a medical student could be asked to click on an area of an xray which shows a cancerous growth. If an incorrect area is selected a HTML document giving further information can be displayed.

An active map can be specified as shown in the HTML document below.

```
Please select an area of the xray showing cancerous
growths.
<A HREF="cgi-bin/htimage/xray.config">
<A IMG SRC="xray.gif" ISMAP></A>
```

Figure 5-1 HTML Document Containing Markup For An Active Map.

The file `xray.config` will contain the coordinates of regions in the image, as illustrated below.

```
default error.html
rectangle (100,100) (500,500) cancer.html
circle (50,50) 25 homepage.html
```

Figure 5-2 Configuration File For Active Map.

When the user clicks on an area of the image the coordinates are sent to the `cgi-bin/htimage` CGI program. The name of the configuration file for the image (in this case (`xray.config`) is also sent to this program. The `htimage` program will then retrieve the HTML document specified in the confuguration file. If, for example, the user has clicked in a circle defined by the centre at position 50,50 with a radius of 25, the file `homepage.html` will be sent to the browser. If the user has clicked in a rectangle with vertices at the position `100,100` and `500,500` the file `cancer.html` will be sent to the browser.

Mapedit

Mapedit is an editor for creating image map files. Image map files are a feature of NCSA and CERN servers; they enable you to turn a GIF image into a clickable map by designating areas using polygons and circles within the GIF and specifying a destination URL for each area. The software is not public domain. Commercial users must pay a licence fee; non-profit and educational users are asked to send the author a postcard. The software is available from the URL `ftp://sunsite.unc.edu/pub/packages/infosystems/WWW/tools/mapedit` In the UK it is available at the URL `ftp://src.doc.ic.ac.uk/packages/WWW/mapedit` Mapedit was written by Thomas Boutell (`mailto:boutell@netcom.com`).

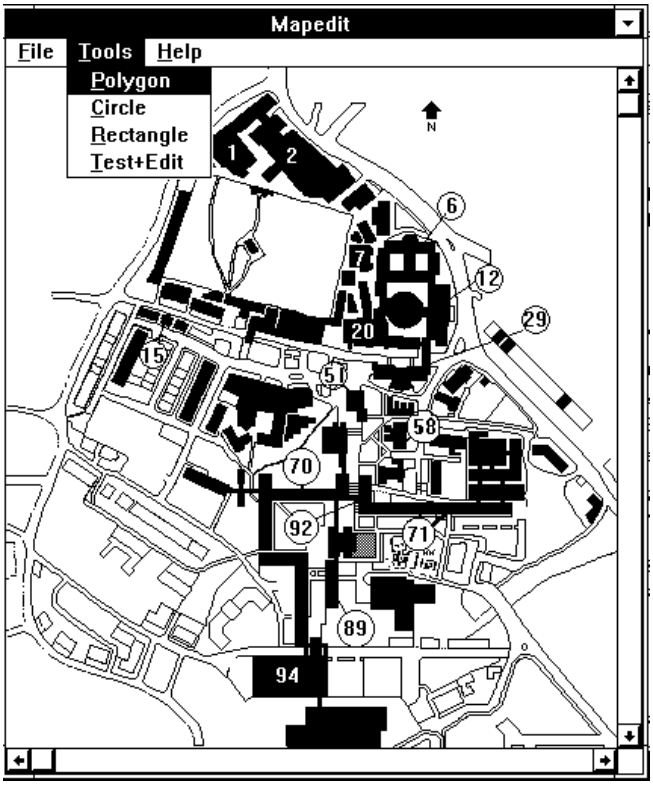


Figure 5-3 MapEdit.

Graphical Tools

Paintshop Pro

Paintshop Pro is an example of a Microsoft Windows tool which can be used to manipulate graphics files for use on WWW. Paintshop Pro can be used to convert file formats, to reduce colour depth and to convert colours.

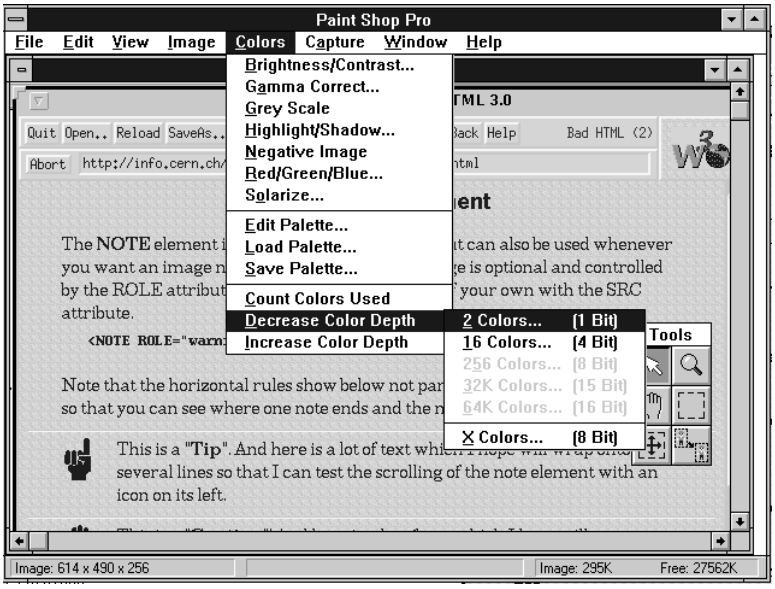


Figure 5-4 Paintshop Pro.

The image being manipulated by Paintshop Pro contains information for 256 colour (as shown in the bottom left of the screen). The colour depth of the image should be reduced to decrease the size of the file to an appropiate level (e.g. a line drawing should not contain 256 colours), and thus reduce the network traffic when the image is retrieved on WWW.

Other Graphical Tools

Imagizer can generate high-quality thumbnail images, among other things, on-the-fly. It is available for SunOS, Solaris, and HPUX, and soon for Windows and NT. Further details are available at the URL `http://pc.inrird.com/imagizer.html`

San Diego Supercomputer Center's **imtools** package converts many file formats, including GIFs.

ImageMagick is a multi-purpose raster converter and manipulation package. The convert program handles many file formats including GIF. The software is available at the URL `ftp://ftp.x.org/contrib`

Graphics Workshop from Alchemy Mindworks is a DOS package for converting graphical files.

Appropriate Use Of Graphics

Novice information providers may be tempted to fill their HTML documents with inline graphical images. More experienced computer users will remember the large numbers of poorly designed paper documents which were produced once desktop publishing packages became widely used.

Before making use of graphics you should consider the following points:

- Will the image look attractive and be readable on computers with monochrome or limited colour capabilities?

- Since many computers currently used have limited colour maps, images which contain many colours may not be displayed correctly.
- Since the http protocol needs to reestablish a connection for every image, using many small images (such as coloured balls for bullets) can be inefficient.
- The use of images to compensate for deficiencies in the current version of HTML can limit the effectiveness of browsers for the disabled (e.g. a browser for the visual impaired which provides audio output will not be able to handle graphical instructions and navigational aids).
- The use of the same image (such as a logo) but held on different departmental servers results in unnecessary traffic, since the image will not be cached in the client.
- Excessive use of graphics can act as a barrier to people in developing countries who may have to pay for their network traffic.

Look at the following URL. See how long it takes for the information to be delivered. Note that if you retry the URL it is likely to be quicker if it is cached by your client or by your server (if your server supports caching).

`http://www.leeds.ac.uk/ucs/people/BKelly/uniras94/uk_logos.html`

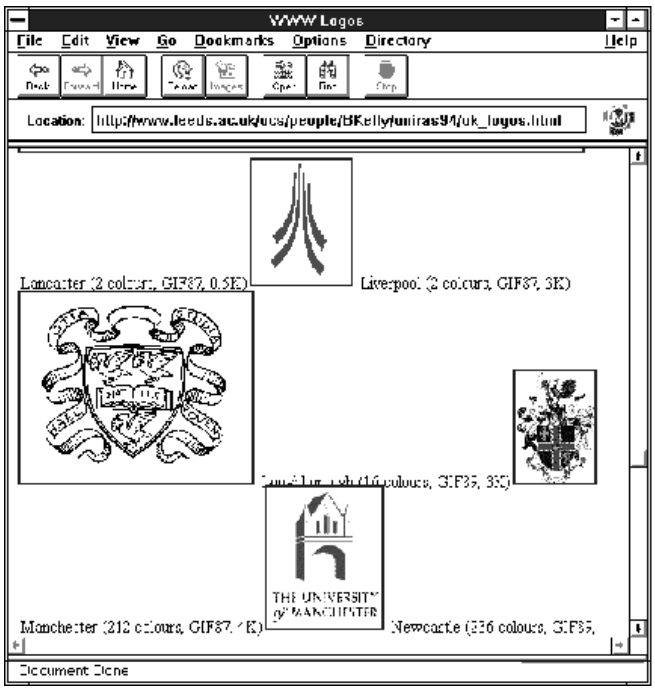


Figure 5-5 UK University Logos.

This page contains pointers to logos on institutional UK university WWW servers. Details of the numbers of colours and the file size are also provideed.

Further Information

A tutorial on imagemaps is available at the URL `http://wintermute.ncsa.uiuc.edu:8080/map-tutorial/image-maps.html`

A good example of use of graphics is the Xerox Parc Map viewer which is available at the URL `http://pubweb.parc.xerox.com/map`

6 Searching And Indexing

The tremendous growth in the numbers and extent of information services on WWW has made *net-surfing* an ineffective way of finding useful information. Fortunately sophisticated indexing tools are being developed. Figure 6-1 shows a page which contains pointers to a number of searching tools.

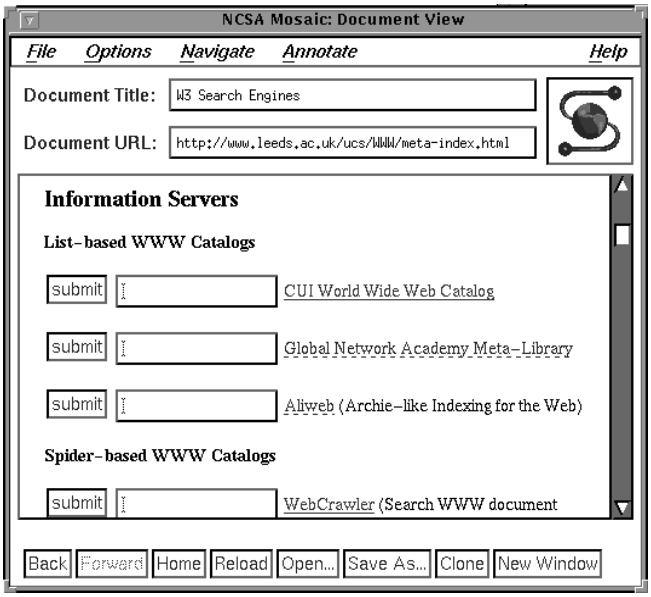


Figure 6-1 A Collection Of WWW Search Engines.

A collection of WWW search engines is available at the URL `http://cuiwww.unige.ch/meta-index.html` Some of the main searching tools are listed below:

CUI WWW Catalog
`http://cuiwww.unige.ch/cgi-bin/w3catalog`

Yahoo
`http://akebono.stanford.edu/yahoo/`

Globewide Network Acadamy
`http://uu-gna.mit.edu:8001/cgi-bin/meta/`

EINet's Galaxy
`http://galaxy.einet.net/`

Aliweb
`http://web.nexor.co.uk/public/aliweb/aliweb.html`

Lycos
`http://fuzine.mt.cs.colorado.edu/mlm/lycos-all.html`

World-Wide Web Worm
`http://www.cs.colorado.edu/home/mcbryan/WWWW.html`

WebCrawler
`http://webcrawler.cs.washington.edu/WebCrawler/WebQuery.html`

RBSE URL
`http://rbse.jcs.nasa.gov/eichmann/urlsearch.html`

Nikos
`http://www.rns.com/cgi-bin/nomad`

Jumpstation Robot
`http://www.stir.ac.uk/jsbin/js`

World-Wide Web Wanderer
`http://www.mit.edu:8001/cgi/wandex`

Robots, Spiders and Worms

During 1993 many WWW users discovered resources by *net-surfing*: going to one WWW server, exploring what was available, and then following links to other WWW servers. A number of software developers produced software which automated this process, so that a program went from server to server, indexing information, such as contents of the `<TITLE>` tag or the contents of server home pages. Such programs became known as robots or spiders; one robot was called WWWW, the World-Wide Web Worm.

There are a number of problems with this approach to global indexing:

- **Server performance** When a robot arrives at your server it can place a heavy load on the server.
- **Network performance** Robots can place heavy loads on the network infrastructure.
- **Quality** Robots index all files. They are unable to differentiate between valuable information resources (such as a collection of research reports) and low quality, transient information (such as an undergraduate's personal interests).
- **Maintenance** The robot will not know if an information source that has been indexed is withdrawn from service.

A number of these issues have been addressed. Martijn Koster's *Guidelines For Robots,* which is available at the URL `http://web.nexor.co.uk/mak/doc/robots/robots.html` provides guidelines for developers of robots.

A list of robots is kept at the URL `http://web.nexor.co.uk/mak/doc/robots/active.html`

Aliweb

Aliweb (Archie Like Indexing In The Web) provides another approach to the indexing of WWW resources. With Aliweb each site is responsible for indexing files. The server administrator is responsible for choosing and describing the services to be indexed.

Further information about Aliweb is available at the URL `http://web.nexor.co.uk/aliweb/doc/aliweb.html` The paper *ALIWEB - Archie-Like Indexing In the Web*, which was presented at the WWW 94 conference in CERN, is available at the URL `http://web.nexor.co.uk/mak/doc/aliweb-paper/paper.html`

SWISH

SWISH, which stands for Simple Web Indexing System for Humans, was announced on 16 November 1994. It is a program that allows you to index your Web site and search for files using keywords in a fast and easy manner. Documentation is available at the URL `http://www.eit.com/software/swish/swish.html`. The software is available at the URL `ftp://ftp.eit.com/pub/web.software/swish/`

WAIS

WAIS (Wide Area Information Server) is another mechanism for indexing resources. WAIS is used by the Computing Service, University of Leeds to index its documents and newsletters. An example of how the WAIS server and WAIS indexing software is used is given below.

The command:

```
waisserver -p 210 -d /apps/info/WWW/WAIS
```

is used to start the WAIS server software. The `-p 210` argument specifies the name of the port on which the server runs while the `-d` argument gives the name of the directory which will contain WAIS databases. Note that since the WAIS server will normally be running continuously it will normally be initiated by the system administrator.

Newsletters are indexed by giving the command

```
waisindex -export -d /apps/info/WWW/ucs/newsletter/wais-
sources/computing-service-newsletter -T HTML *.html
```

The name of the WAIS database is `computing-service-newsletter` This long name is used since a single directory is used for all WAIS databases - it will save confusion if other departments wish to index their own departmental newsletters.

The WAIS database can be accessed by a dedicated WAIS client or by a WWW browser which contains support for the WAIS protocol. The WAIS database can be accessed by giving the URL `wais://www.leeds.ac.uk/computing-service-newsletter`

WAIS Utilities

A number of utilities are available which can post-process the output from WAIS.

wais.pl is a CGI script which is distributed with the NCSA httpd server.

Son of wais.pl is a CGI script which is based on the wais.pl script.

SFGate is a CGI script which interfaces to WAIS servers. SFGate provides a forms interface which can be used to access a number of WAIS databases. It is available at the URL `http://ls6-www.informatik.uni-dortmund.de/SFgate/SFgate.html` A demonstration is available at the URL `http://ls6-www.informatik.uni-dortmund.de/SFgate/multiple.html`

wwwwais is a small ANSI C program that acts as gateway between waisq or waissearch (programs that search WAIS indexes) and a forms-capable World-Wide Web browser. With the freely distributable freeWAIS package, this program, and your local Web site, you can:

- Create searchable databases of the information on your Web site
- Allow users to search multiple databases via their Web browser with customizable options
- Create a custom pop-up menu of servers to search through
- Produce hypertext search results, with file information and links directly to the relevant HTML documents
- Retrieve WAIS source descriptions and files
- Specify URLs and filters to map results to
- Only allow users from certain sites to search documents

Documentation is at the URL
`http://www.eit.com/software/wwwwais/wwwwais.html`

You can FTP the source and related files from the URL
`ftp://ftp.eit.com/pub/web.software/wwwwais/`

You can see how it looks at the URL `http //www.eit.com/cgi-bin/wwwwais`

A WAIS Application

One interesting application of the use of WAIS is the multimedia archive prototype developed by Andy Walker, formerly of the CBL/Multimedia Unit, University of Leeds. The prototype was developed to investigate the feasability of providing a archive of multimedia objects for use in CBL applications by members of the University of Leeds.

A directory is created for each multimedia object. The directory contains the multimedia object itself (e.g. a graphical file, video clip or sound file) together with a keyword file which

describes the object. The keyword files are indexed using WAIS. A WWW browser which supports forms is used to run a CGI script. The CGI script invokes the **waisq** command to search the WAIS database. The output from waisq is then used to create a HTML file which contains a pointer to a thumbnail image of matching multimedia objects.

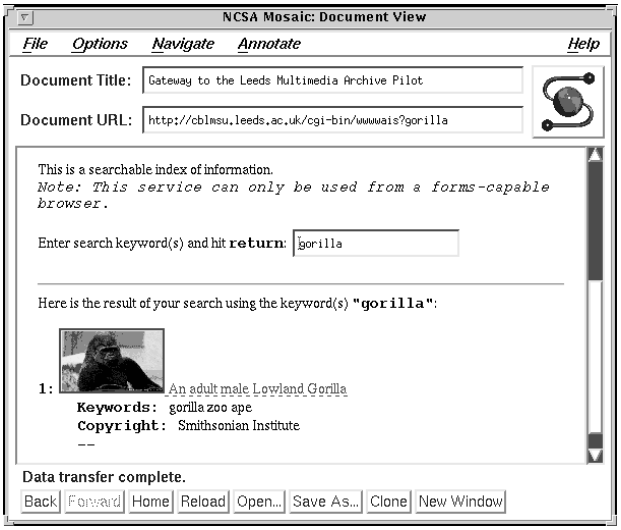


Figure 6-2 Multimedia Archive.

Which WAIS?

A number of WAIS servers are available. The **freeWAIS** software is currently used at the University of Leeds. This software is maintained by CNIDR, the Clearinghouse For Networked Information Discovery and Retrieval. The freeWAIS software, however, is based on the 1988 version of the Z39.50 protocol. An implementation of WAIS based on the 1992 version of Z39.50 is also believed to be available from CNIDR. freeWAIS is available at the URL `ftp://ftp.cnidr.org/pub/NIDR.tools/freewais`

freeWAIS-sf is an implementation of WAIS developed at Dortmund University. It is available at the URL `ftp://ls6-www.informatik.uni-dortmund.de/pub/wais/freeWAIS-0.2-sf-beta.tar.gz`

CNIDR Isite

CNIDR Isite is an integrated software package including a text indexer, search engine and Z39.50 communication tools to access databases. Isite includes the CNIDR ZDist, Isearch and Search API distributions.

A mailing list has been established to discuss Isite. To join, send an email message to `listserv@vinca.cnidr.org` with the body of the message as **`subscribe ISITE-L`** *`your name`* To post messages to the list, send to `isite-l@vinca.cnidr.org`.

Further information is available at the URL
`http://vinca.cnidr.org/software/Isite/Isite.html`

Further Information

A tutorial on *Mosaic and WAIS* is available at the URL
`http://wintermute.ncsa.uiuc.edu:8080/wais-tutorial/wais.html`

A WAIS overview is available at the URL
`http://info.cern.ch/hypertext/Products/wais/sources/Overview.html`

A list of resources about the Z39.50 information discovery protocol is available at the URL
`http://ds.internic.net/z3950/z3950.html`

7 WWW Servers

If you wish to make information available you will need to run a WWW server. The server software is known as httpd - the hypertext transport protocol daemon. Just as their many WWW browsers available there are also many servers, including ones for Unix, MS Windows, Windows NT and the Apple Macintosh.

This section gives an example of how to install and run a server for the Microsoft Windows environment. The section then goes on to illustrate a number of server management issues which are based on the CERN server for the Unix platform.

Example of Installing A Server On A PC

An example illustrating how easy it is to install a WWW server is given below. The example assumes that you have access to a networked PC.

Connect to the NCSA server software from the anonymous FTP server at `ftp.ncsa.uiuc.edu` Then change directory to `/Web/httpd/Uni/ncsa_httpd/contrib/winhttpd` Finally retrieve the file `whtp13p1.zip` An example of how to do this using the FTP software is illustrated below.

```
ftp src.doc.ic.ac.uk
image
cd /Web/ncsa/httpd/Windows/winhttpd
get whtpp13p1.zip
```

Create a directory called `C:\HTTPD` on the `C:` drive of your PC and then move to the directory using the `CD \HTTPD` command. Then uncompress the file by giving the command:

```
PKZUNIP -D WHTPP13P1.ZIP
```

The `-D` option will preserve the directory structure from the compressed file.

Run Microsoft Windows and create a program icon using the `New` option on the `File` menu. The icon should point to the file `C:\HTTPD\HTTPD.EXE`

Set the time zone in the `AUTOEXEC.BAT` file so that `TZ=GMT`.

Run the server program. The window shown below should be displayed.

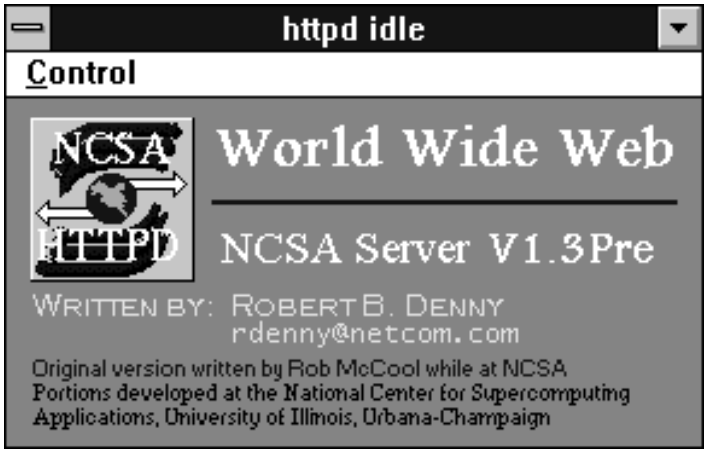


Figure 7-1 Running The Windows HTTPD Server.

Run a World-Wide Web browser and then enter a URL containing the IP address of your PC. For example if your PC has an IP address of 192.11.1.1 you should enter the address:

```
http://192.11.1.1/
```

The following diagram illustrates NCSA Mosaic for X accessing a server running on a PC.

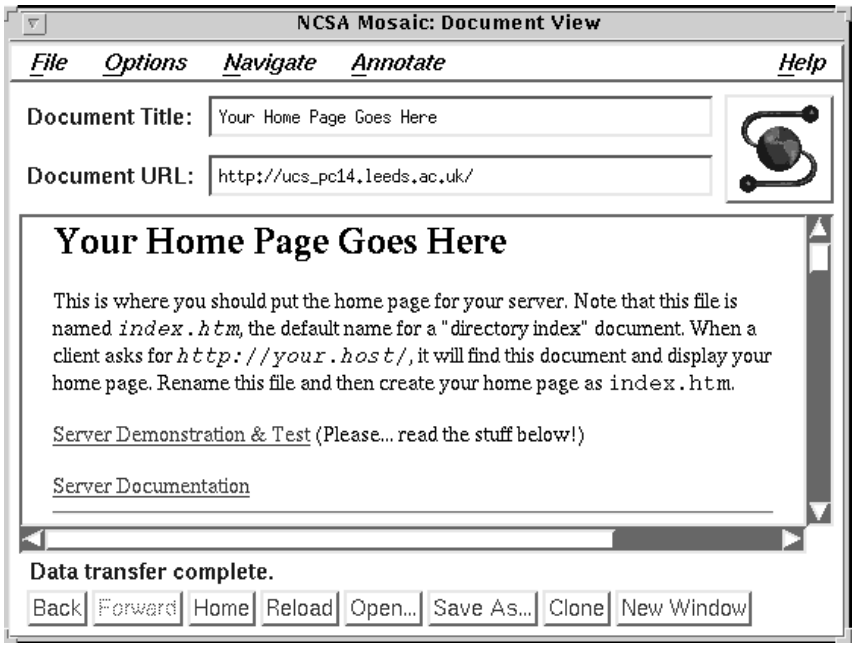


Figure 7-2 Accessing The MS Windows HTTPD Server.

This example is meant to illustrate the installation of a WWW server. In practice the server software is likely to run on a more robust system than a PC running MS DOS, such as a Unix or Windows NT system.

Server Configuration Files

World-Wide Web server software will normally have a configuration file which is used to:

- Specify an area of the file store on the server machine from which files can be retrieved.
- Restrict access to certain files.
- Specify caching, which enables remote files to be stored on the local server for a period of time.
- Act as a proxy gateway, which provides a means for unauthorised machines to access resources on WWW.

As WWW develops, additional features will be provided in the server software and the configuration files are likely to grow in complexity. An example of a simple configuration file is shown below.

```
map   /      file:/apps/WWW/homepage.html
map   /*     file:/apps/WWW/*
pass  file:/apps/WWW/*
fail  *
```

Figure 7-3 A Simple `httpd.conf` Configuration File

Figure 7-3 shows a simple configuration file for the CERN httpd server. Line 2 specifies that files located under the directory `/apps/WWW` should be available to the WWW server software. Line 1 specifies that file `/apps/WWW/homepage.html` is the default file to be displayed when the WWW server is accessed.

```
protection        prot-proxy {                        # Part 1
                  serverid www.leeds.ac.uk
                  mask @(129.11.*.*)
}
protect        http:*    prot-proxy                   # Part 2
protect        gopher:*  prot-proxy
protect        ftp:*     prot-proxy
protect        wais:*    prot-proxy

pass           http:*                                 # Part 3
pass           gopher:*
pass           ftp:*
pass           wais:*

Exec /cgi-bin/ucs/*      /apps/WWW/cgi-bin/ucs/*     # Part 4
Exec /cgi-bin/bionet/*   /apps/WWW/cgi-bin/bionet/*
Exec /cgi-bin/bmb/*      /apps/WWW/cgi-bin/bmb/*

map  /             file:/apps/WWW/homepage.html      # Part 5
map  /*            file:/apps/WWW/*
pass file:/apps/WWW/*
fail *

AccessLog          /var/adm/httpd.log                # Part 6
LogFormat   Common
LogTime     LocalTime

Caching On                                           # Part 7
CacheRoot /usr/info/WWW_cache
CacheSize 300
CacheAccessLog /var/adm/httpd_cache.log

# Delete files from cache after specified number of days
CacheClean http:*              10 Days
CacheClean gopher:*            10 Days
CacheClean wais:*              10 Days
CacheClean ftp:*               10 Days

# Don't cache local files                            # Part 8
NoCaching http://*.leeds.ac.uk/*

# If a file hasn't been accessed within the last specified
# number of days delete from cache
CacheUnused *      5 days
CacheUnused http://info.cern.ch/* 10 days
CacheUnused http://www.ncsa.uiuc.edu/* 10 days

# ensure dynamically changing documents are only kept for short
# periods e.g. one modified 10 days ago will only last 2 days
CacheLastModifiedFactor  0.2

# If a file was retrieved more than 5 days ago do a
# 'conditional get' request to the source server to check
# that it hasn't been updated in the meantime.
CacheRefreshInterval http://* 5 days
CacheRefreshInterval gopher://* 5 days
CacheRefreshInterval ftp://* 5 days

# CacheDefaultExpiry ensures that Gopher and FTP files are
# cached. The default is 0 which is what we want for http
# documents with neither an expiry nor a last-modified stamp.
CacheDefaultExpiry ftp://* 5 days
CacheDefaultExpiry gopher://* 5 days

# Remove unwanted cached files daily at 3 am (garbage collection).
Gc On
GcDailyGc 3:00
```

Figure 7-4 A `httpd.conf` Configuration File

Figure 7-4 shows another configuration file (this is for illustrative purposes - some options may have been superceeded). The various features are summarised below:

Parts 1 and 2 provides a mechanism for ensuring that the proxy gateway cannot be accessed from outside the local domain. Without these options it would be possible for a browser on an external system to use the proxy gateway to gain access to files which are restricted to local use.

Part 3 passes requests for the httpd, gopher, wais and ftp protocols.

Part 4 specifies the location for CGI files.

Part 5 specifies the area of the filestore which can be accessed by the server.

Part 6 describes the location and format of the server log file.

Part 7 specifies that server caching is to be available, and gives the location of the cache and the cache log files, together with the size (in Mbytes) of the cache.

Part 8 specifies the purging frequency for files in the cache.

An example of a typical `httpd.log` file is shown below.

```
abc.cs.xyz.edu - - [21/Nov/1994:21:58:58 +0000] "GET
   /music.html HTTP/1.0" 200 4375
gps0 - - [21/Nov/1994:21:59:48 +0000] "GET / HTTP/1.0" 200
   2782
abc_pc99.leeds.ac.uk - - [21/Nov/1994:21:59:47 +0000] "GET
   http://www.leeds.ac.uk/ HTTP/1.0" 200 2782
abc.nt.com - - [21/Nov/1994:22:00:03 +0000] "GET
   /music/NetInfo/MusicFTP/ftp_sites.html HTTP/1.0" 200 13175
```

Figure 7-5 Example of a `httpd.log` File.

Note that the names of the machines accessing files from the server have been altered in the diagram. This has been done because it could be argued that such information should be confidential.

Caching

Many clients provide client-side caching. This means that if you retrieve a file and then retrieve another file, when you return to the initial file it will be retrieved from the client's cache, thus saving a subsequent network transfer.

A number of servers also support caching by the server. This is illustrated in Figure 7-6.

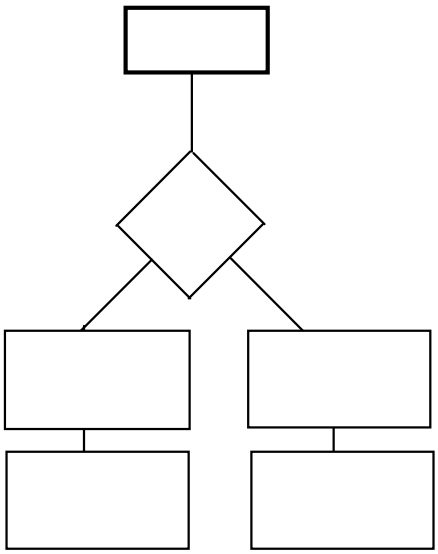


Figure 7-6 Caching By The Server.

Caching can improve the performance of a WWW service by ensuring that frequently requested files will tend to be stored in the local cache. There is, of course, a danger that if the

file on the remote server is updated that an out-of-date file will be retrieved from the cache. In practice, however, httpd server software which supports caching can deal with this issue by, for example, looking at the date of the file on the remote server and, if the remote file is newer than the file in the cache, replacing the file in the cache with the new version of the file.

```
[Proxy Information]
http_proxy:       www.leeds.ac.uk
gopher_proxy:     www.leeds.ac.uk
wais_proxy:       www.leeds.ac.uk
```

Figure 7-7 Client Configuration File To Support Caching.

It order for a client to make use of a cache on a server, the client configuration file (e.g. the `MOSAIC.INI` file) must be suitably configured. Figure 7-7 illustrates the relevant options for the `MOSAIC.INI` file.

Accesses of the cache are recorded in the cache log file. A typical log file is illustrated in Figure 7-8.

```
xyz_pc77.leeds.ac.uk - - [21/Nov/1994:00:43:35 +0000] "GET
   http://white.nosc.mil/gif_images/NM_Sunrise_s.gif HTTP/1.0" 200 18673
xyz_pc77.leeds.ac.uk - - [21/Nov/1994:00:43:38 +0000] "GET
   http://white.nosc.mil/gif_images/glacier_s.gif HTTP/1.0" 200 6474
xyz_pc77.leeds.ac.uk - - [21/Nov/1994:00:43:40 +0000] "GET
   http://white.nosc.mil/gif_images/rainier_s.gif HTTP/1.0" 200 18749
```

Figure 7-8 The httpd_cache Log File.

Note that the names of the machines accessing files from the cache have been altered in the diagram. This has been done because it could be argued that such information should be confidential.

Caching Strategies

As well as using a local server cache, it is also possible to use a national caching service. The Unix HENSA service at the University of Kent at Canterbury run a national caching service. To use this service the local client should define `www.hensa.ac.uk` as the proxy. Another national caching service is available at `sunsite.doc.ic.ac.uk` Further information is available at the URL `http://src.doc.ic.ac.uk/WWW-Cache.html`

An institution will need to decide whether to use a caching service and, if so, whether to have caching services running on a number of departmental system, to have an institutional caching service, or to use the national caching service at HENSA. In the future it may be possible to chain caches. The possibility in the long term of having institutional, metropolitan, national and continental caches should be considered.

Proxy Gateways

In many academic institutions off-campus access to the Internet is restricted to authorised computers. Depending on the institution's local policy, authorisation may be restricted to computers located in offices in which there is an individual who is responsible for use of the machine. Such a policy may be enforced in order to provide some means of security against hacking remote services. However this policy would appear to prevent students from accessing remote information services from computers in open access cluster areas.

In practice there is a technique known as *proxy gateways* which can be used to provide access to services off-campus, without compromising local security. With a proxy gateway a trusted system (typically a Unix system which is more secure to hacking than a desktop machine) will have Internet access. Machines in open access cluster can point to the proxy gateway, which will then retrieve information from off-campus services.

It should be noted that with increasing usage of Internet services such as the World-Wide Web, the author believes that the provision of security mechanisms, such as proxy gateways, will be increasingly important.

Further information

Further information on caching and proxies is available at the following URLs:

`http://www.hensa.ac.uk/www94/`
`http://info.cern.ch/httpd/Proxies/ManyProxies.html`
`http://www.hensa.ac.uk/new.cache.html`

Security

The httpd server also handles a number of security issues. It is common practice to restrict access to a certain area of the filestore. For example if the server configuration files contains the lines:

```
map   /*    file:/apps/WWW/*
pass  file:/apps/WWW/*
fail  *
```

Figure 7-9 Server Configuration File.

then clients will only be able to access files held under the directory `/apps/WWW/`.

Note This statement refers to clients running on remote machines. If the client is running on the same machine as the server, the client will normally be able to access files on the server to which it has read access.

Additional levels of security can also be specified:

- Restricting access to specified computer systems.
- Restricting access by username and password.

The method of implementing such security tends to be server dependent, and will not be described in this document.

Summary of Server Software

A brief summary of server software is given below. This summary is based on Thomas Boutell's WWW FAQ.

Unix Servers

CERN httpd

Information about CERN's server is available at the URL
`http://info.cern.ch/hypertext/WWW/Daemon/Status.html`

NCSA httpd

NCSA's server is available at the URL
`ftp://ftp.ncsa.uiuc.edu/Web/ncsa_httpd`

EIT httpd

EIT have created a *Webmaster's Starters Kit* which installs their server using a forms interface from a WWW browser. Further information is available at the URL
`http://wsk.eit.com/wsk/doc/`

GN Gopher/http

The GN server can serve both WWW anbd Gopher clients. It may be useful for sites wishing to migrate from Gopher to WWW, although it does not have the server-script capabilities of the CERN and NCSA servers. Further information is available at the URL
`http://hopf.math.nwu.edu/`

Plexus perl server

The Plexus server is written in Perl. Further information is available at the URL
`http://bsdi.com/server/doc/plexus.html`

WebWorks Enterprise server

The is a commercial server marketed by Quadralay Inc. Further information is available at the URL
`http://www.quadralay.com/products/WebWorks/Server/index.html`

Netsite Communication Server and Netsite Commercial Server

These servers have been developed by Netscape Communications Corporation. Further information is available at the URL
`http://home.mcom.com/MCOM/products_docs/server.html`

Macintosh Servers

MacHTTP

Information about the MacHTTP server for the Apple Macintosh is available at the URL
`http://www.biap.com/`

Novell Netware Servers

httpd nlm

The httpd NLM server for Novell Netware is available at the URL
`ftp://ftp.glaci.com/pub/netware/http/`

Microsoft Windows and Windows NT Servers

https

HTTPS, developed at Edinburgh University, is a Windows NT server which runs on Intel, MIPS and Alpha CPUs. It is available at the URL
`ftp://emwac.ed.ac.uk/pub/https/`

NCSA httpd For Windows

The NCSA httpd for Windows server provides most of the features of the Unix version, including scripts (which generate pages on the fly). It is available at the URL
`ftp://ftp.ncsa.uiuc.edu/Web/ncsa_httpd/contrib/`

SerWeb

SerWeb is a Microsoft Windows server. It is available at the URL
`ftp://emwac.ed.ac.uk/pub/serweb/`

Web4Ham

Web4Ham is a Microsoft Windows server. It is available at the URL
`ftp://ftp.informatik.uni-hamburg.de/pub/net/winsock/`

Server Strategies

An institution needs to decide on its server strategies. For example, should it support:

1. A central server
2. A number of departmental servers

If option 2 is chosen then how is indexing across servers to be achieved, and what caching strategy is to be adopted. What are the skills levels needed by the server administrator? An institution, such as a university, needs to recognise that adopting a server strategy is more than simply installing the server software.

Which Server?

The most widely used servers are probably those developed at CERN and NCSA for the Unix platform. Unix is probably the best platform for running an institutional WWW service, since it is a mature, pre-emptive multi-tasking operating system. In addition, Unix provides a wide

range of tools which can be used to assist in system administration. Servers are available for the PC and Macintosh platform, but, due to the inherent deficiencies in the operating system environments which are currently used on the platform, such servers are probably not recommended if you wish to run a large-scale, stable service.

Servers have been developed for the Windows NT environment. This may provide a robust operating system environment which can be used for providing a WWW server on an Intel platform.

Further Information

Further information about **HTTP** is available at the URL `http://info.cern.ch/hypertext/WWW/Protocols/`

Information about **HTTP/NG** is available at the URL `http://info.cern.ch/hypertext/WWW/Protocols/HTTP-NG/http-ng-status.html`

The **HTTP/1.0 specification** has been submitted as an Internet-Draft and is available for comment at the following URLs: `http://www.ics.uci.edu/pub/ietf/http/draft-fielding-http-spec-00.txt` and `ftp://www.ics.uci.edu/pub/ietf/http/draft-fielding-http-spec-00.txt`

The document *Setting up a World-Wide Web Server*, which is available at the URL `http://scholar.lib.vt.edu/reports/Servers-web.html`, gives advice on setting up a server.

A **collection of utilities** intended especially for WWW system administrators is available at the URL `ftp://src.brunel.ac.uk/WWW/managers/`

A **list of server software** is available at the URL `http://www.cern.ch/hypertext/WWW/Daemon/Overview.html`

A **list of server software** is available at the URL `http://www.charm.net/~web/Vlib/Providers/Servers.html`

A **list of server software** is available at the URL `http://akebono.stanford.edu/yahoo/Computers/World_Wide_Web/HTTP_Servers/`

A hypermail archive of the **HTTP-WG mailing list** is available at the URL `http://www.ics.uci.edu/pub/ietf/http/hypermail/`

8 Extending WWW

External Viewers

Access to WWW can be achieved by using a client such as NCSA Mosaic to display HTML documents and inline images in GIF format. However the World-Wide Web is an extensible system: clients can access information which is in other formats than HTML.

When a client receives a file from a server it checks on the file type. If the file type indicates that it is an HTML document, the file will be displayed by the browser. Otherwise the browser's configuration file can specify an external viewer which can be used to display the file. A list of widely used external viewers is given in Table 8-1.

File Format	Viewer
JPEG	LVIEW (MS Windows) xv (X Windows)
Postscript	Ghostview
DVI	xdvi (X Windows)
MPEG	mpeg_play (X Windows and MS Windows)

Table 8-1 Popular Viewers.

The association between the file type and the viewer is given in the browser's configuration file. A typical configuration file for Mosaic for Windows is given in Figure 8-1.

```
[Viewers]
TYPE0="audio/wav"
TYPE1="application/postscript"
TYPE2="image/gif"
TYPE3="image/jpeg"
TYPE4="video/mpeg"
TYPE5="video/quicktime"
TYPE6="video/msvideo"
TYPE7="application/x-rtf"
TYPE8="audio/x-midi"
TYPE9="audio/basic"
TYPE10="image/x-action"
TYPE11="application/x-w3launch"
application/postscript="L:\winapps\ghost\gsview %ls"
application/x-w3launch="n:\windept\bmb\w3launch\w3launch
%ls"
image/gif="L:\winapps\mosaic2\lview %ls"
image/x-action="n:\windept\bmb\action25\playact %ls"
image/jpeg="L:\winapps\mosaic2\lview %ls"
video/mpeg=""
video/quicktime=""
video/msvideo=""
audio/wav=""
audio/x-midi="mplayer %ls"
application/x-rtf="write %ls"
```

Figure 8-1 Part of a MOSAIC.INI File.

Running Client Applications

If a Postscript file is retrieved from a WWW server the browser program normally responds "*I don't know what to do with a Postscript file - but I know a program that does. I'll pass the Postscript file on to the Ghostview program*". If, for example, an Excel spreadsheet is retrieved from a WWW server the client could be configured to respond "*I don't know what to do with an Excel spreadsheet file - but I know a program that does. I'll pass the spreadsheet*

file on to the Excel program". This technique extends the functionality of the World-Wide Web from acting as a distributed file viewer to acting as a distributed program manager.

Security Implications

Unfortunately there are a number of security concerns with such an approach. For example an application developed using the Toolbook authoring system could be delivered using WWW. The application could then be launched using a local copy of Toolbook. The application could have a button marked **Start**. Clicking this button could then result in files held on the local machine being deleted! Even associating a word processed document with Word For Windows holds dangers, as many Microsoft applications, including Word For Windows, support the use of macros, including autostart macros, which could also cause files to be deleted.

As a general principle there are dangers in automatically invoking applications from WWW clients.

Implementing Security - W3Launch

There are security problems in using a WWW browser to download and run software from the Internet. It is generally **not** considered wise to configure a browser so that it recognises file types which contain programs. Jon Maber, Biochemistry and Molecular Biology, University of Leeds has developed a launching program for the Bionet TLTP project which provides a simple and secure method of launching only authorised software.

Further details on the W3Launch program is available at the URL
`http://www.leeds.ac.uk/bionet/student/pre-stud.htm`

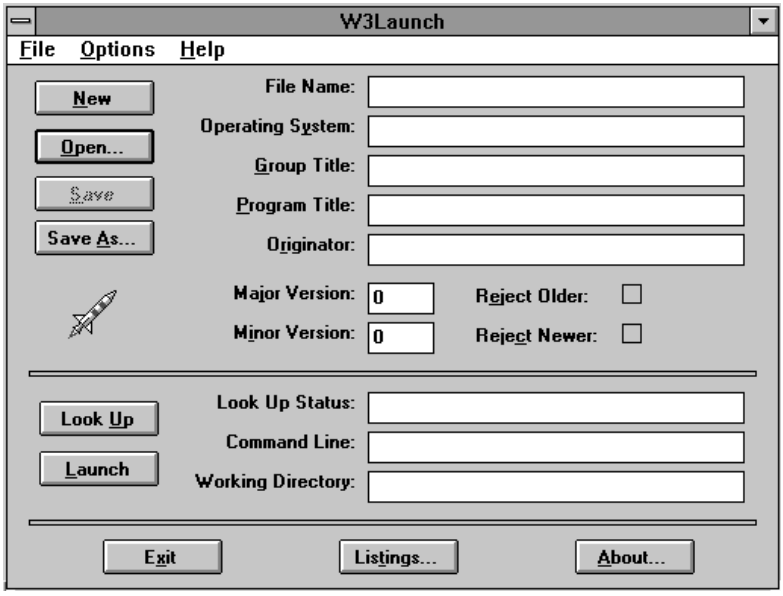


Figure 8-2 W3Launch.

It should be noted that W3Launch is an application developed at the Unievsrity of Leeds - it is not part of WWW itself.

Server-side Extensions

Example

The previous section described how it is possible to run applications on the client machine. It is also possible to run software on the server. A simple application running on the server is shown in Figure 8.3.

```
#!/bin/sh
echo Content-type: text/html
echo
if [ $# = 0 ]
then
   echo "<HEAD>"
   echo "<!-- Script written by Brian Kelly --!>"
   echo "<TITLE>Search University Phone Directory</TITLE>"
   echo "<ISINDEX>"
   echo "</HEAD>"
   echo "<BODY>"
   echo "<H1>Phone Directory</H1>"
   echo "Enter surname of the person you are searching for.<P>"
   echo "Script written by <A HREF=http://www.leeds.ac.uk/
ucs/people/BKelly/bk.html>Brian Kelly</A>."
   echo "</BODY>"
else
   echo "<HEAD>"
   echo "<TITLE>Results Of Search</TITLE>"
   echo "</HEAD>"
   echo "<BODY>"
   echo "<H1>Results of Search for $* </H1>"
   echo "<PRE><TT>"
   grep -i "$*" /apps/data/Telephone_Directory
   echo "</PRE></TT>"
   echo "</BODY>"
fi
```

Figure 8-3 Script To Generate An HTML Document.

The program, which is a C shell script which runs on the Unix server system, can be executed by selecting the URL `http://www.leeds.ac.uk/ucs/cgi-bin/phone`

When the URL is selected since no arguments are provided, the first part of the if statement is run. This will generate the following HTML document:

```
<HEAD>
<!-- Script written by Brian Kelly --!>
<TITLE>Search University Phone Directory</TITLE>
<ISINDEX>
</HEAD>
<BODY>
<H1>Phone Directory</H1>
Enter surname of the person you are searching for.<P>
Script written by <A HREF=http://www.leeds.ac.uk/
ucs/people/BKelly/bk.html>Brian Kelly</A>.
</BODY>
```

Figure 8-4 Virtual HTML Document.

The `<ISINDEX>` tag generates a search dialogue box. The HTML document is rendered as shown below.

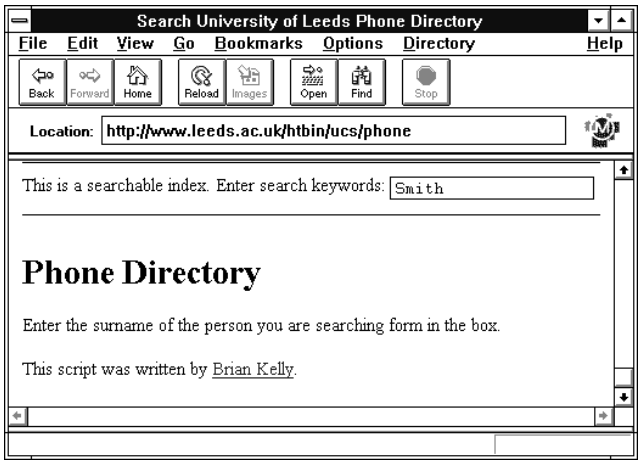


Figure 8-5 Running The Script.

When text is entered in the Search box and the <Enter> key pressed, the script in Figure 8.4 is executed again. This time, since the program will be given an argument, the second part of the `if` statement will be executed. This will generate the HTML tags and then invoke the Unix `grep` command to search a file for lines containing the search string.

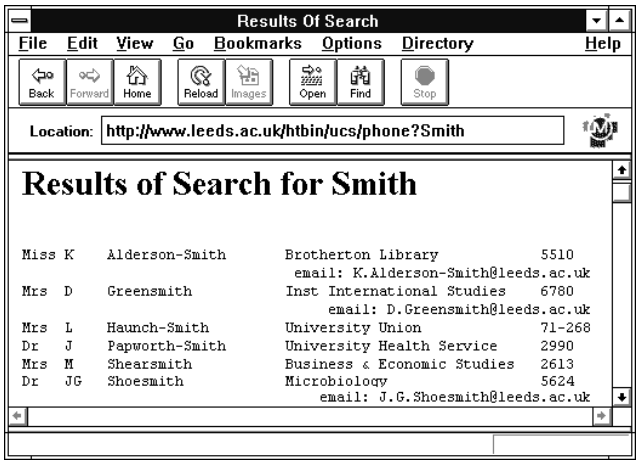


Figure 8-6 Output From The Script.

CGI Programs

The example described above is known as a CGI program. CGI stands for the *Common Gateway Interface*. It is a standard which has been adopted by a number of server developers (primarily developers of the CERN and NCSA server software) for running programs on the server machine. A definition of CGI is available at the URL
`http://hoohoo.ncsa.uiuc.edu/cgi/examples.html`

Examples of the use of CGI programs are available at the URl
`http://hoohoo.ncsa.uiuc.edu/cgi/examples.html`

A tutorial on CGI is available at the URL
`http://www.charm.net/~web/Tutorial/CGI/`

A *CGI Programmer's Reference* is available at the URL
`http://www.halcyon.com/hedlund/cgi-faq/`

An archive of useful CGI programs is available at the URL
`ftp://ftp.ncsa.uiuc.edu/Web/httpd/Unix/ncsa_httpd/cgi/`

Forms

Forms are often used to collect the information from a user which is used as input to a CGI program. A description of forms is given below.

Creating A Form

A form consists of areas of the screen in which the user can input data. The data is sent to the HTTP server, which can run a script or program to process the data in some way. One common use of forms is to provide feedback on a WWW service. Input to the form can be emailed to the service administrator. Forms can also be used to input search criteria to be input to a search engine, or to specify parameters for distributed teaching and learning services.

A form is defined by the <FORM ...> and </FORM> HTML tags. The <FORM> tag has the syntax:

```
<FORM METHOD="method" ACTION="url">
```

For example:

```
<FORM METHOD="post" ACTION="http://leeds.ac.uk/ucs/
cgi-bin/myscript">
```

will send the input data to be processed by `myscript`.

An example of a form is shown below:

```
<TITLE>Fill-Out Form Example #7</TITLE>
<H1>Fill-Out Form Example #7</H1>
This is another fill-out form example, with toggle buttons. <P>
<HR>
<FORM METHOD="POST" ACTION="http://hoohoo.ncsa.uiuc.edu/htbin-post/post-query">
<H2>Godzilla's Pizza -- Internet Delivery Service, Part II</H2>
Type in your street address: <INPUT NAME="address"> <P>
Type in your phone number: <INPUT NAME="phone"> <P>
Which toppings would you like? <P>
<OL>
<LI> <INPUT TYPE="checkbox" NAME="topping" VALUE="pepperoni">
     Pepperoni.
<LI> <INPUT TYPE="checkbox" NAME="topping" VALUE="sausage"> Sausage.
<LI> <INPUT TYPE="checkbox" NAME="topping" VALUE="anchovies">
     Anchovies.
</OL>
How would you like to pay?  Choose any one of the following: <P>
<OL>
<LI> <INPUT TYPE="radio" NAME="paymethod" VALUE="cash" CHECKED> Cash.
<LI> <INPUT TYPE="radio" NAME="paymethod" VALUE="check"> Check.
<LI> <I>Credit card:</I>
<UL>
<LI> <INPUT TYPE="radio" NAME="paymethod" VALUE="mastercard"> Mastercard.
<LI> <INPUT TYPE="radio" NAME="paymethod" VALUE="visa"> Visa.
<LI> <INPUT TYPE="radio" NAME="paymethod" VALUE="americanexpress">
     American Express.
</UL>
</OL>
Would you like the driver to call before leaving the store? <P>
<DL>
<DD> <INPUT TYPE="radio" NAME="callfirst" VALUE="yes" CHECKED> <I>Yes.</I>
<DD> <INPUT TYPE="radio" NAME="callfirst" VALUE="no"> <I>No.</I>
</DL>
To order your pizza, press this button: <INPUT TYPE="submit"
VALUE="Order Pizza">. <P>
</FORM>
```

Figure 8-7 HTML Document Defining A Form.

This example is available at the URL
`http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/fill-out-forms/example-7.html`

The way in which form is displayed is illustrated below.

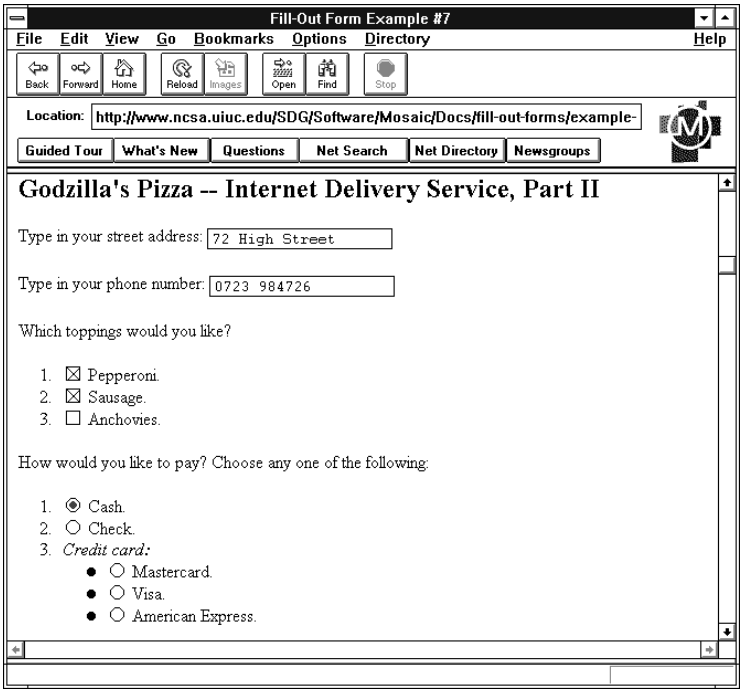


Figure 8-8 Using A Form.

Processing A Form

Once the form is submitted the data which has been entered is appended to the end of the URL given in the `ACTION` attribute of the `FORM` tag. This information is then processed by the script.

Further Information About Forms

Forms tutorials are available at the URL
`http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/fill-out-forms/overview.html`,
`http://hoohoo.ncsa.uiuc.edu/docs/cgi/forms.html`,
`http://www.webcom.com/html/tutor/forms/start.html` and
`http://kuhttp.cc.ukans.edu/info/forms/forms-intro.html`

A **forms testing suite** is available at the URL
`http://www.research.digital.com/nsl/formtest/home.html`

9 Utilities

A number of useful utility programs have been developed which will assist systems managers and information providers.

w3new is a program which will extract a list of URLs from the Mosaic client hotlist file or extract URLs from a HTML document. It will then retrieve the modification dates for each document listed and output a HTML file with the URLs sorted by their last modification date.

Information about the program is available at the URL
`http://www.stuff.com/~bcutter/home/programs/w3new/w3new.html`
The utility was written by Brooks Cutter (`mailto:bcutter@stuff.com`).

wusage is a WWW server usage meter which produces weekly activity reports in HTML. In addition it provides graphical displays of server usage.

Further information is available at the URL `http://siva.cshl.org/wusage.html`
The software is available from the URL `ftp://isis.cshl.org/pub/wusage` wusage was written by Thomas Boutell (`mailto:boutell@netcom.com`).

getstats (formerly called getsites) is a versatile WWW server log analyser. It is available at the URL `http://www.eit.com/software/getstats/getstats.html`

weblint is a Unix utility for checking the syntax of HTML documents. The checks include illegally nested, overlapped, unclosed and obsolete tags. Further details are available at the URL `http://www.khoros.unm.edu/staff/neilb/weblint.html` The software can be obtained from the URL `ftp://ftp.khoros.unm.edu/pub/perl/www/`. The utility was written by Neil Bowers, Khoral Research Inc. (`mailto:neilb@khoros.unm.edu`) The email list `weblint@khoros.unm.edu` provides announcements of new versions of Weblint. Email Neil Bowers if you wish to be added to the list.

Verify_links is a robot which performs link verification. Further information is available at the URL
`http://wsk.eit.com/wsk/dist/doc/admin/webtest/verify_links.html`

MOMspider (Multi-Owner Maintenance spider) is a tool which can be used to help information providers and system managers to maintain links to documents. MOMspider is available at the URL `http://www.ics.uci.edu/WebSoft/MOMspider/`

Hypermail is a program that converts a file of email messages to a hypertext WWW form. It is available at the URL
`http://gummo.stanford.edu/html/hypermail/hypermail.html`

The following utilities are available at the URL
`ftp://src.doc.ic.ac.uk/pub/packages/infosystems/WWW/tools`

checkweb	looks for dead links in your Web
html+tables.shar	creates preformatted text tables from HTML+ Table definitions
mosaic-wais-cli.pl	does a WAIS search using Mosaic from the command line
newslist/	compiles an HTML page of links to all newsgroup on your server
simon/	URL database to replace NCSA Mosaic's Hotlist
test-cgi/	sets up HTTP environment for a CGI script
url-get.pl	a perl script which brings in any document given its URL
w3get.pl	retrieves a HTML page named by a URL and all HREFs and IMGs in it

10 Legal and Ethical Issues

Is your WWW service legal? Who is legally responsible for the contents of a WWW service? Is pornography acceptable on a WWW service? If not, who defines what is pornographic and what is art? How do you reconcile control over the contents of a WWW server with intellectual freedom?

The author does not know the answer to these questions. Fortunately WWW is attracting the interests of lawyers, philosophers and artists who are starting to address these issues. Many of the papers which have been published address issues which affects WWW providers in the USA. The American Constitution, and in particular the amendment on free speech, means that much of the work published in the USA in this area is not relevant to the UK.

A future edition of this handbook may contain information relevant to UK legislation and culture. This edition covers a number of issues are raised which WWW service and information providers should consider.

Liability

It could be argued that the contents of a WWW service are the responsibility of the organisation which runs the service. So if an undergraduate has been granted to publish information and publishes libellous information the University may be legally responsible. An editorial in the Times Higher Education Supplement suggested that if the organisation has published guidelines covering acceptable and unacceptable use the organisation will have a strong defence if a case is brought to law.

Computer Misuse Act

It is likely that any material which incites, encourages or enables others to gain unauthorised access to a computer system would be found illegal under this act.

Pornography

Are pictures of naked women acceptable on a WWW service? It could be argued that similar guidelines which govern the contents of a University library should be developed for the WWW. Are pictures of naked women acceptable in books in the university library? The answer is probably "yes", especially if the university has a fine art department. Similar arguments could be made for textual pornography.

However even the most liberal individual is likely to be offended by some of the pornography which is believed to be available on the Internet. In addition UK legislation on computer pornography is likely to be introduced shortly. This could mean that universities have a legal obligation to concern itself with computer pornography.

Copyright, Designs and Patents Acts

In general the Copyright, Designs and Patents acts require that the permission of the owner of the intellectual property **must** be sought before any use of it is made whatsoever.

A WWW manager may have the responsibility to ensure that copyright materiel is not made available unless the copyright holder has granted permission. This may affect research papers which have been submitted for publication. It may also affect the use of photographs, drawings and maps, for which the copyright may be owned, for example, by the photographer or the organisation which commissioned the photograph.

Data Protection Act

Information about individuals which is available on WWW may have to be registered with the Data Protection Officer. The information provider may have to abide by regulations to ensure the accuracy of the information.

Equality Of Access To Information

WWW can provide global access to a wide range of information services. However including large logos and graphical icons on pages can act as a barrier to access to the information, especially for readers in developing countries will limited network access. In some developing countries access may be provided over local telephones lines. A health worker in a hospital in Africa who wishes to retrieve information about public health services may have to pay the additional in retrieving unnecessary graphics. If the local telephone company is owned by a multinational telephone corporation then accessing the information will result in a transfer of money from the developing country to the multinational corporation.

Advertising

As shown in Figure 10-1 some WWW service providers have sponsors for their pages. Is this currently acceptable within the UK academic community? Should it be acceptable?

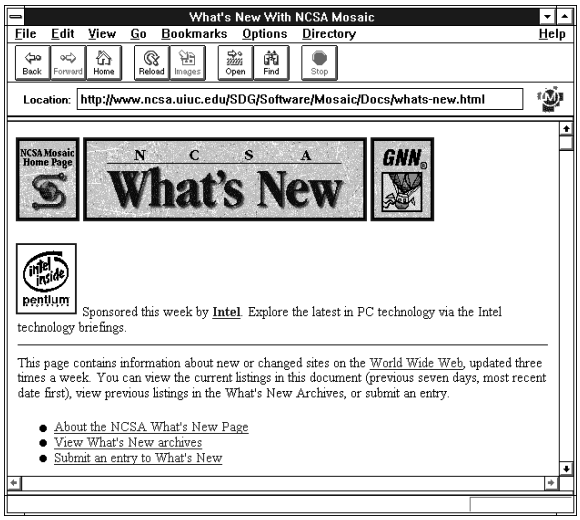


Figure 10-1 The "What's New On Mosaic" Page.

JANET Acceptable Use Policy

UK universities which make use of JANET (the Joint Academic Network) must abide by the rules and regulations governing he use of JANET. The following point should be noted.

JANET may be used for any legal activity in furtherance of the aims and policies of a connected organisation, subject to a number of rules. For example, the following uses are not permitted on JANET:

- Use for offensive, indecent, etc. purposes.
- Use for unsolicited mail.
- Deliberate unauthorised access to connected services.

What Is Your WWW Service For?

Formulating an institutional acceptable use policy for WWW information providers may not be a simple task. There are likely to be lively discussions over censorship and control. The formulation of the policy will be helped if the institution has a clear idea of what it expects from its WWW service. Is it:

- To provide easy access to high quality information?
- To enable students to gain experience in multimedia publishing?
- To provide researchers with a test bed for their research?
- To provide members of the University with access to cultural services and to enable them to publish cultural services?

Further Information

An interactive document called *Sex, Censorship and the Internet* is available at the URL `http://www.eff.org:80/CAF/cafuiuc.html` This document asks questions such as should universities carry alt.sex Usenet groups and should students be punished for using vulgarities on the Net. The document provides pointers to case studies.

Cranfield University have published guidelines for information providers which is available at the URL `http://www.cranfield.ac.uk/docs/publish_code.html`

Information about the Data Protection Act is available at the URL `http://www.open.gov.uk/dpr/dprhome.htm`

11 CWISes And WWW

WWW is an ideal system for developing a campus (or community) wide information system (CWIS). The world's first multimedia CWIS was developed at the Honolulu Community College and officially announced at the end of May 1993. It is available at the URL `http://www.hcc.hawaii.edu/`

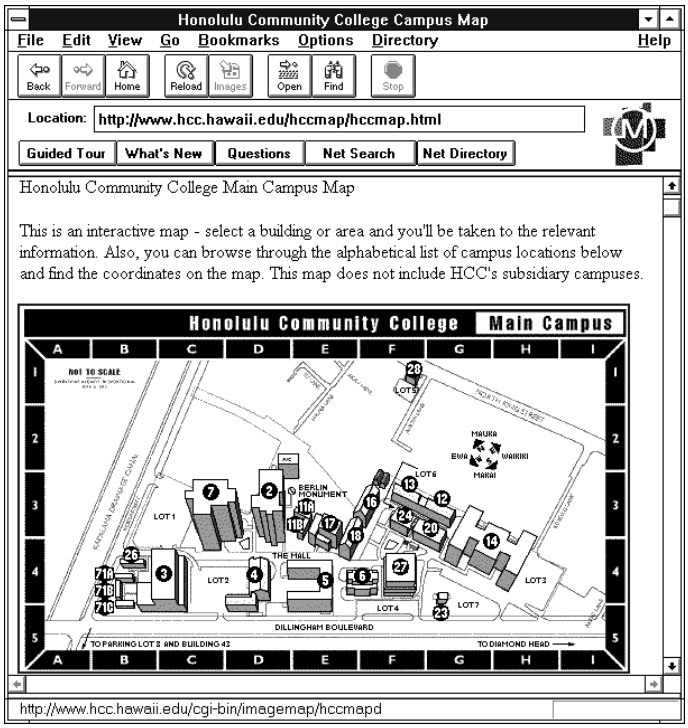


Figure 11-1 CWIS At HCC.

The HCC CWIS was developed to support its goal of becoming the "*Technological Training Centre of the Pacific*". The most important aspects of developing and managing an effective CWIS are managerial and not technical. Formulating the objectives of a CWIS, resourcing it and developing a training programme are key issues which an institutional needs to address.

The Universities and Colleges Teaching, Learning and Information Group (UCTLIG) are producing a *CWIS Manager's Handbook* which will address many of these issues.

Finding Out More

Papers by Judy Hallman about CWISes are available at the URL `ftp://sunsite.unc.edu/pub/docs/about-the-net/cwis/cwis-l` and `ftp://sunsite.unc.edu/pub/docs/about-the-net/cwis/hallman.txt`

Polly-Alida Farrington's listing of CWISes is available at the URL `http://www.rpi.edu/Internet/cwis.html`

Lists of (global) CWISes are available at the URLs `http://www.rpi.edu/Internet/cwis.html` and `http://kawika.hcc.hawaii.edu/ws94/cwis.html`

The CWIS-L Listserv mailing list provides a forum for the discussion of topics related to campus-wide information systems. To subscribe send the message **`SUB CWIS-L`** *your name* to the address **`LISTSERV@MSU.EDU`**

A **Framework for Administering NASA's Web Information Hypermedia** is available at the URL `http://naic.nasa./gov/www-framework.html`

12 Teaching And Learning On WWW

Although WWW was initially used as a distributed multimedia system techniques such as CGI scripts meant that interaction could be built into WWW applications. Much of the interest in the WWW within the academic community is based on its potential for developing distributed teaching and learning software rather than simply delivering information.

Examples of Teaching And Learning On WWW

An early example of a distributed multimedia teaching prototype was developed by Ben Whitaker, School of Chemistry, University of Leeds in 1993. As can be seen in Figure 12-1 this prototype is a simple hypertext application. It is of interest because it illustrates how *distributed* teaching applications can be developed.

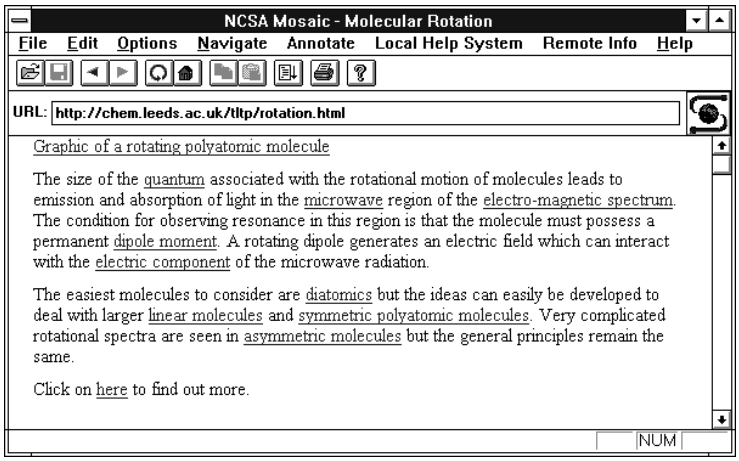


Figure 12-1 Early Example Of A Distributed Multimedia Teaching Application.

A more sophisticated teaching application was developed by the School of Chemistry in conjunction with Imperial College. The example illustrated in Figure 12-2 makes use of a chemistry MIME type.

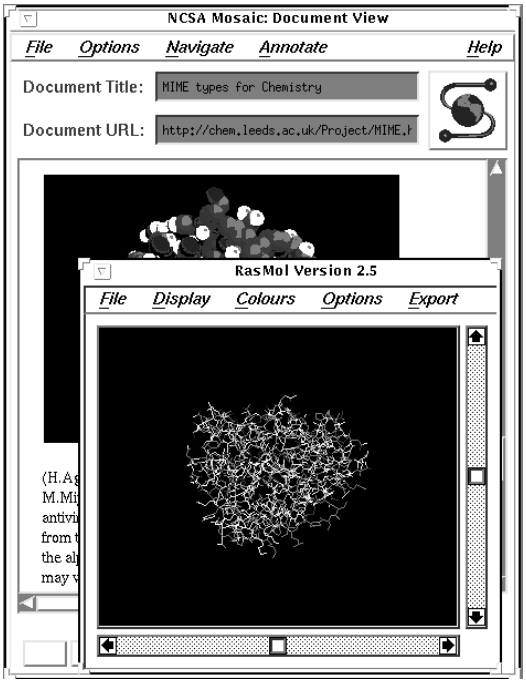


Figure 12-2 Using a MIME Chemistry Type.

In this example the WWW client is configured to associate the MIME type with the RasMol program. For example in NCSA Mosaic For X the line:

```
chemical/x-pdb; rasmol %s
```

is included in the `.mailcap` file. When a URL with the extension `.pdb` is selected the file will be downloaded and the Rasmol program launched, as illustrated in Figure 12-2.

Further information on this project is available at the URL `http://chem.leeds.ac.uk/Project/MIME.html`

The **Globewide Network Academy** (GNA) is a consortium of educational and research organisations. Its mission is to provide a central organisation in which students, teachers, scholars and researchers can meet and interact. Further information about GNA is available at the URL `http://uu-gna.mit.edu:8001/uu-gna/`

Mark Cox, Department of Industrial Technology, University of Bradford presented a paper at the *Mosaic and the Web* conference on *Robotic Telescopes: An Interactive Exhibit on the Web*. This paper is available at the URL `http://www.eia.brad.ac.uk/mark/wwwf94/wwwf94.html`

Mark also has a collection of pointers to hardware control services over the Web which is available at the URL `http://www.eia.brad.ac.uk/mark/fave-inter.html`

A **Virtual Frog Dissection Kit** has been developed at the LBL. It is available at the URL `http://george.lbl.gov/ITG.hm.pg.docs/Whole.Frog/Whole.Frog.html`

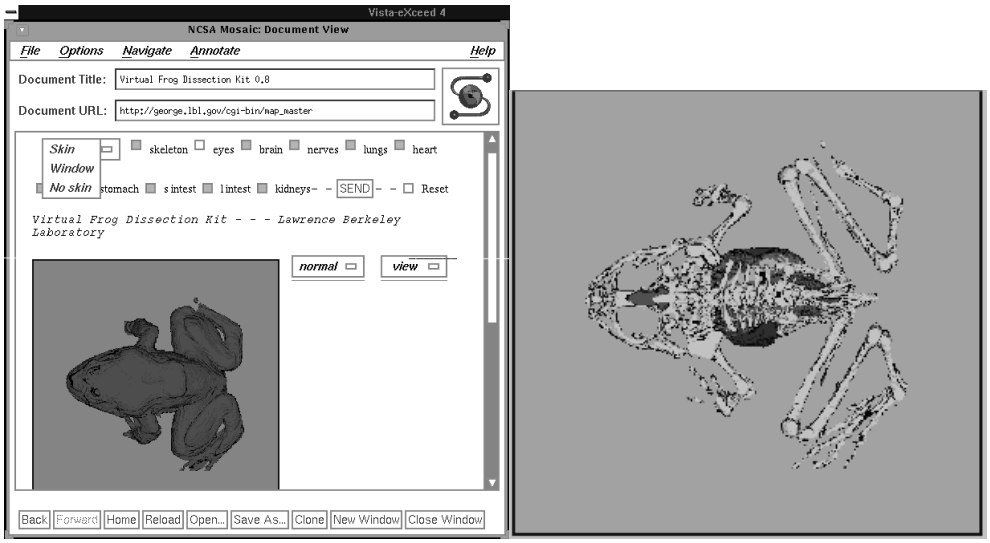


Figure 12-3 Frog Dissection.

CD ROM Facilities

Providing teaching and learning services on WWW does not necessarily deny access to those who do not have a network connection. Teaching and learning services developed on WWW can be transferred to a CD ROM and used on a standalone system. Such systems are typically developed so that there is a closed set of links. The files (which could include HTML documents, image, sound and video files) and the WWW browser software can then be transferred onto a CD ROM. This approach provides an updateable service for users with network connectivity together with a fixed service for users with access to a PC or Macintosh with a CD ROM player.

National Resources

A number of TLTP (Teaching and Learning Technology Programme), CTI (Computers in Teaching Initiative) and ITTI (Information Technology Training Initiative) projects are using WWW to disseminate information about their projects or, in some cases, to deliver their courseware.

CTISS is available at the URL `http://www.ox.ac.uk/cti/`

CTI Centre For Biology is available at the URL
`http://www.liv.ac.uk/ctibiol.html`

CTI Centre For Chemistry is available at the URL
`http://www.liv.ac.uk/ctichem.html`

CTI Centre For Law is available at the URL
`http://crocus.csv.warwick.ac.uk/WWW/law/default.html`

CTI Centre For Psychology is available at the URL `http://ctipsych.york.ac.uk/`

CTI Centre For Sociology is available at the URL
`http://lorne.stir.ac.uk/departments/cti_centre/`

CTI Centre For Textual Studies is available at the URL
`http://www.ox.ac.uk/depts/humanities/`

BioNet Project is available at the URL `http://www.leeds.ac.uk/bionet.html`

CLIVE Project is available at the URL `http://www.vet.ed.ac.uk/`

Insurrect Project is available at the URL `http://av.avc.ucl.ac.uk/`

Institute Of Computer Based Learning, Heriot-Watt is available at the URL
`http://www.icbl.hw.ac.uk/`

INTERACT Project is available at the URL
`http://medusa.eng.cam.ac.uk/~interact/`

Interactive Learning Centre, University of Southampton is available at the URL
`http://ilc.ecs.soton.ac.uk/welcome.html`

ITTI is available at the URL `http://www.hull.ac.uk/Hull/ITTI/itti.html`

PsyCLE Project is available at the URL
`http://ctipsych.york.ac.uk/Psycle/PsyCLEinfo.html`

STILE Project is available at the URL `http://indigo.stile.le.ac.uk/`

TLTP is available at the URL `http://www.hcbl.hw.ac.uk/tltp/`

TLTP Archaeology Consorteum is available at the URL
`http://www.brad.ac.uk/acad/archsci/homepage.html`

TLTP Mathematical Project is available at the URL
`http://othello.ma.ic.ac.uk/`

Further Information

Further information about a mailing list for teaching and learning is available at the URL
`http://tecfa.unige.ch/edu-ws94/ws.html`

Pointers to global uses of WWW for teaching is available at the URL
`http://wwwhost.cc.utexas.edu/world/instruction/index.html`

Harry Kriz's paper "*Teaching and Publishing in the World Wide Web*" is available at the URL
`http://learning.lib.vt.edu/webserv/webserv.html`

13 Collaboration On WWW

WWW was originally envisaged by Tim Berners-Lee as a groupware tool. In practice it grew in popularity as a publishing tool. However software developers are now working on tools which will facilitate collaboration on WWW. A brief summary of some of the collaborative tools is given below.

Asynchronous Systems

WIT

WIT, the WWW Interactive Talk system, was announced shortly after the WWW 94 conference in CERN. WIT can be accessed at the URL `http://info.cern.ch/wit`

Figure 13-1 WIT.

Access To Usenet

The Netscape browser can be used to post to Usenet newsgroups.

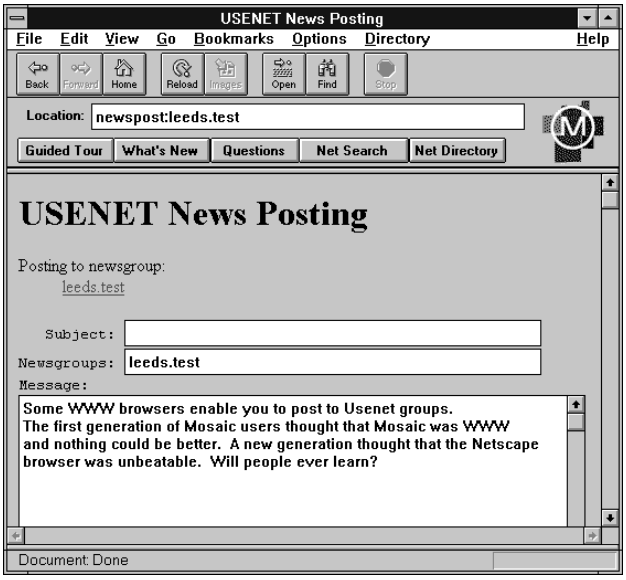


Figure 13-2 Posting To Usenet News.

Hypermail

Hypermail is a utility which can be used to convert mail archives to hypertext format on WWW. A example of a hypermail archive is illustrated below.

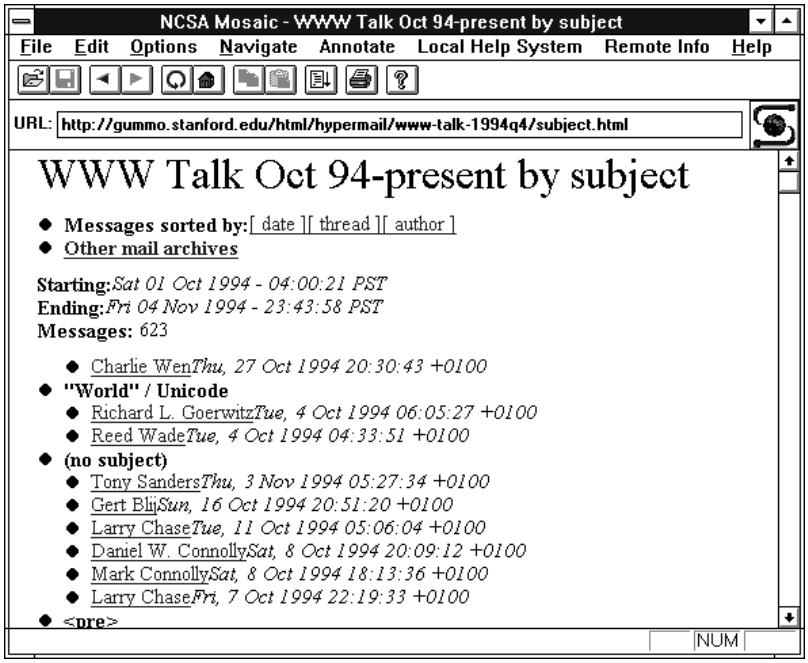


Figure 13-3 A Hypermail Archive.

Mailserv

Mailserv provides a forms interface to a number of mailing list servers. The software is available at the URL `http://iquest.com/~fitz/www/mailserv/` The software was written by Patrick M Fitzgerald (`mailto:pmfitzge@iquest.com`)

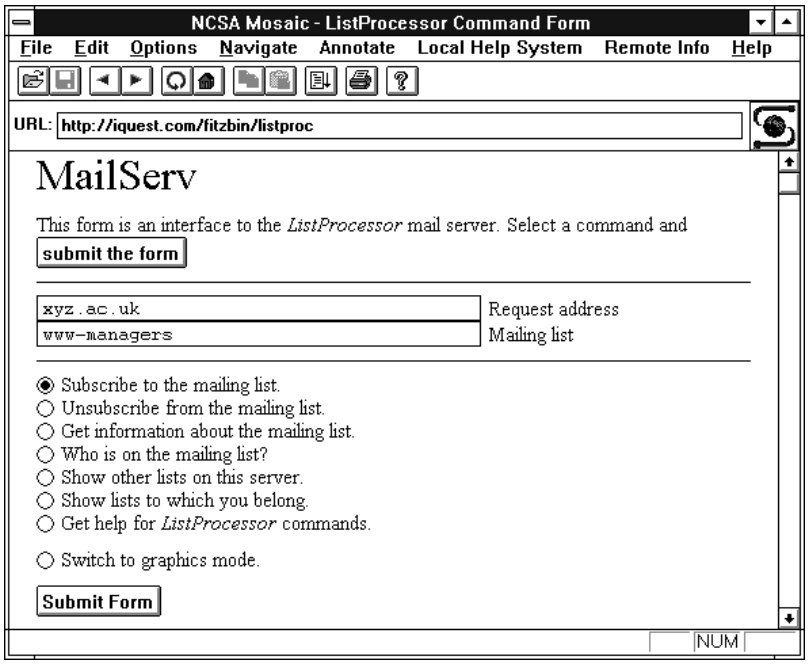


Figure 13-4 The Mailserv Interface To Mailing List Servers.

Synchronous Systems

Video conferencing facilities are being developed which can be integrated with WWW.

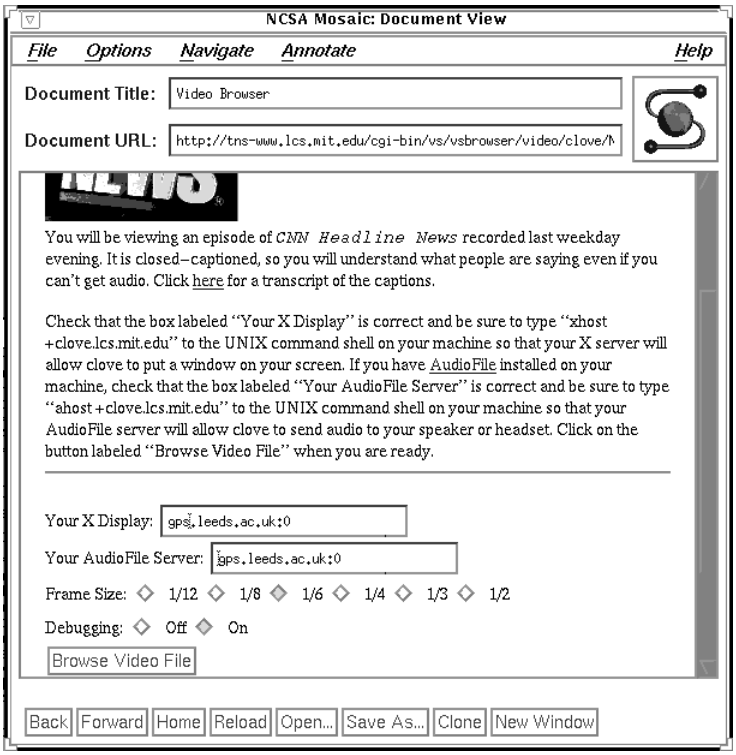


Figure 13-5 Accessing A Video On WWW.

One interesting application of a multimedia desktop conferencing systems is MONET (Meeting on the Network) which is described in *Applications of Mosaic in Health Care Delivery* by Srivasa *et al*. This paper, which was presented at the *Mosaic and The Web* conference, is available at the URL `http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings/MedTrack/srivasa/artemis.html`

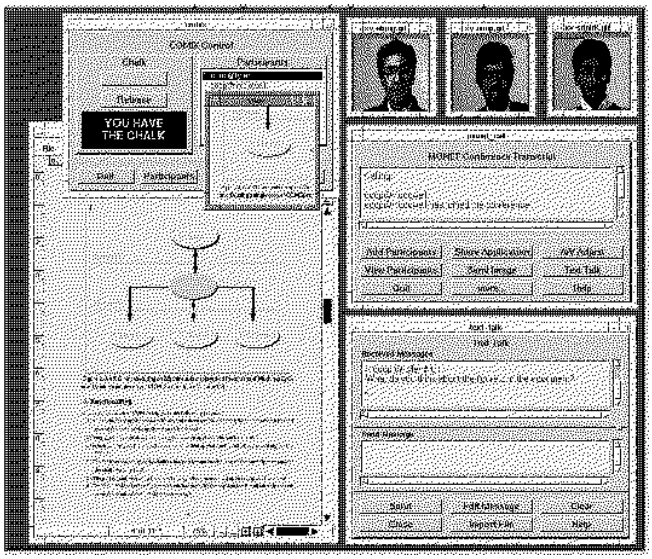


Figure 13-6 MONET.

At the time of writing many of these services are experimental. However, given the rapid growth of WWW and the extent of development work which is going on, such services may be mainstream in the near future.

Virtual Conferences

One form of collaboration within the academic community is through conferences, workshops and seminars. Whenever the author gives a paper at a conference or is involved in running a workshop or a course he makes his papers, OHP foils, etc. available on WWW under his personal page (sometimes referred to as a *vanity page*).

About 200 of the papers which were given at the second WWW conference, *Mosaic and The Web*, were available on WWW before the conference began. Perhaps one important question which the academic community should be addressing is whether it should be the standard practice for conference proceedings to be made available on WWW.

Further Information

A collection of WWW collaborative projects is available at the URL
`http://union.ncsa.uiuc.edu/HyperNews/get/www/collaboration.html`

Examples of conference proceedings available on WWW is given in Appendix 5.

14 Libraries And WWW

University Libraries should have a strong interest in WWW developments. This handbook provides a overview of the World-Wide Web which should be of interest to libraries which are considering using WWW.

Example Of A Gateway To A Library Catalogue

In the UK many university library catalogues are held in proprietary systems with old-fashioned user interfaces. It may be possible, however, to use WWW to provide an interface to the library catalogue which is consistent with other information services on WWW. At the University of Leeds a backup copy of the library catalogue is kept on a central Unix system using the BRS free text retrieval system. A gateway program, has been developed by Terry Screeton, Computing Service which provides access to the Library catalogue. This gateway is available at the URL `http://www.leeds.ac.uk/library/cats/backup.html`

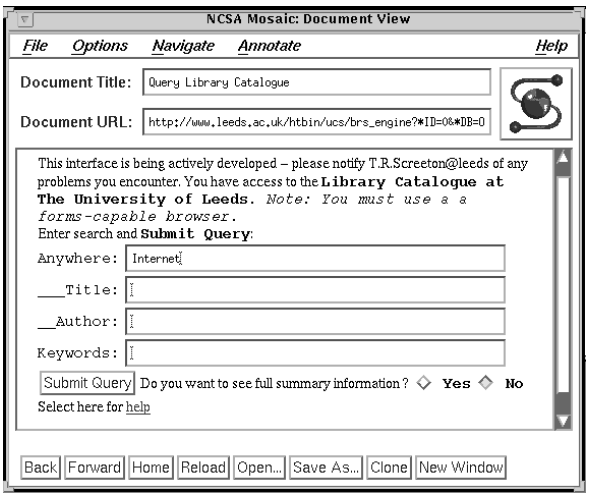


Figure 14-1 Gateway To A Library Catalogue.

In Figure 14-1 a form is completed. The term Internet is used as a search term. Once the form is submitted the data is sent to a CGI program. In this case the CGI program is a C program which invokes the BRS free text retrieval system. The output from the BRS program is then processed to generate the appropriate HTML markup. The output from the search is illustrated in Figure 14-2.

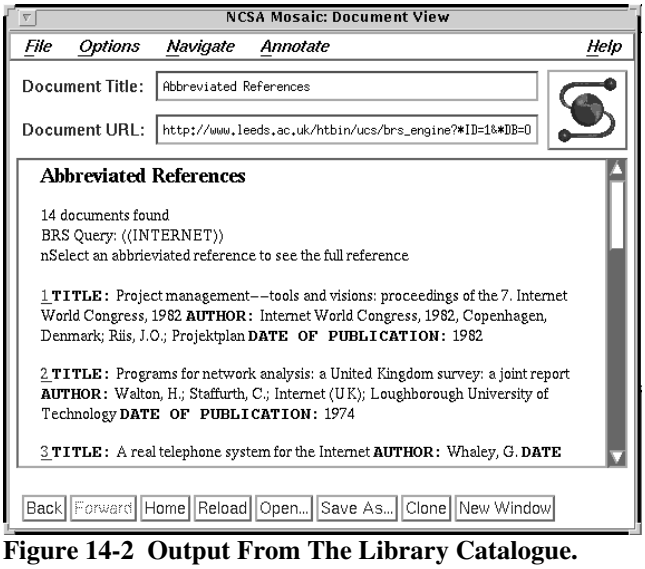


Figure 14-2 Output From The Library Catalogue.

Resources

Datalib provides an interface to a number of online information service hosted at Edinburgh University. It can be accessed at the URL `http://datalib.ed.ac.uk/`

SALSER is an online information service about serials held in Scottish academic and research libraries. It can be accessed at the URL `http://salser.ed.ac.uk/`

The **Clearinghouse for subject-oriented Internet resource guides** is available at the URL `http://http2.sils.umich.edu/~lou/chhome.html`

The **EINet Galaxy** collection of online resources is available at `http://galaxy.einet.net/galaxy.html`

The **CERN Virtual Library** is available at the URL `http://info.cern.ch/hypertext/DataSources/bySubject/Overview.html`

The **Boulder Community Network** service is available at the URL `http://bcn.boulder.co.us/` Its policy statement is available at the URL `http://bcn.boulder.co.us/bcn/policy.html` The policy statement includes a bill of rights, a freedom to read statement and a freedom to view statment.

The following Library resources may also prove useful:

- `http://www.lib.washington.edu/~tdowling/libweb.html`
- `http://www.library.nwu.edu/DOCS/LibResources.html`
- `http://www.rutgers.edu/lawschool.html`
- `http://badger.state.wi.us:/00/agencies/dpi/www/dpi_home.html`
- `http://akebono.stanford.edu/yahoo/Reference/Libraries`
- `http://cyclops.idbsu.edu`
- `http://www.loc.gov/`
- `http://alpha.acast.nova.edu/hytelnet.html`
- `http://www.arcade.uiowa.edu/arcade-www/gateway/search/library.html`

Finding Out More

Web4Lib is a mailing list aimed at library-based WWW managers and developers. To subscribe to the list email **`listserv@library.berkeley.edu`** with the message **`SUBSCRIBE Web4Lib`** *`yourname`*.

Eric Morgan's article on Libraries and the Web in *Public Access Computer Systems Review*, 5(6) 1994:5-26 is available at the URL `http://www.lib.ncsu.edu/staff/morgan/www-and-libraries.html`

15 Future Developments

This handbook describes how to run a WWW service using the technology which is available today. However the technology is developing so rapidly that it is important that WWW managers and information providers are aware of developments which may happen sooner rather than later.

Uniform Resource Identifiers

Uniform Resource Locators (URLs) describe the location of a resource on the Internet and the protocol which is used to access the resource. An object on WWW may be available in many locations: for example popular browsers, such as NCSA Mosaic, are available from anonymous FTP servers in many locations around the world. The mirroring of files helps to minimise network traffic over busy links, such as the trans-Atlantic link. Mirroring also reduces the load on the central server. *Uniform Resource Names* (URNs) will provide a mechanism for uniquely identifying a resource. In the future it is likely that a browser will request a URN rather than a URL. A URN to URL resolver will locate the nearest object (nearest in network terms).

Uniform Resource Characteristics (URCs) will provide meta-information about a document. This information could include information about the author, keywords, expiry dates (for caching servers), copyright and cost information. URCs could also provide information about the quality of the document. For example a seal of approval (SOAP) could be given by university publications group which confirms, by the use of a digital signature, that the document is a PhD thesis.

Uniform Resource Identifiers (URIs) includes URLs, URIs and URCs. The URI specification is available as RFC 1630. The mailing list `uri@bunyip.com` is used to discuss URIs. Send email to `uri-request@bunyip.com` to subscribe to this list. Archives of the list are available at the URL `http://www.acl.lanl.gov/URI/archive/uri-archive.index.html`

New Facilities

CCI

NCSA Mosaic For X (version 2.5) provides support for CCI (Common Client Interface). This will provide a standard mechanism by which WWW browsers can communicate with external programs. A number of demonstrations of this facility are available, including a slideshow program, which instructs Mosaic to display URLs which are specified in a file. A program called xwebteach provides a mechanism by which a teacher can control the display of Mosaic on student's machines. Further information about the CCI specification is available at the URL `http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/cci-spec.html`

W3A

W3A (World-Wide Web Applets) is a proposal for a standard API for dynamically linking applets (which can be defined as a piece of software that can be attached to a host program such as a WWW browser). Further information is available at the URL `http://www.let.rug.nl/~bert/W3A/W3A`

Appendix 1 Mailing Lists

This section contains information on mailing lists and Usenet groups on topics related to the World-Wide Web.

Please note that before sending a message to any of these lists please listen to the discussions first and, where possible, read the information about the list. You should not send simple questions about, say, installing Mosaic on your home PC to a list for developers of the WWW protocols.

Usenet News

comp.infosystems.www.users

comp.infosystems.www.users provides a forum for discussion of WWW client software (such as Mosaic, Cello and Lynx), new user questions, client setup questions, client bug reports, resource discovery questions on how to locate information on WWW that can't be found by the FAQ and comparisons between various client packages are among the acceptable topics for this group.

comp.infosystems.www.providers

comp.infosystems.www.providers provides a forum for the discussion of WWW server software and the use of server software to provide information to users. General server design, setup questions, server bug reports, security issues, HTML page design and other concerns of information providers are among the likely topics for this group.

comp.infosystems.www.misc

comp.infosystems.www.misc provides a general forum for discussing WWW issues which are not covered by the other comp.infosystems.www groups.

comp.infosystems.announce

comp.infosystems.announce is for announcement of new information services (e.g. new WWW sites) and new software products (new server software, new clients, new document converters, etc.) An archive of the `comp.infosystems.announce` Usenet group is available at the URL
`http://www.cs.rochester.edu/users/grads/ferguson/announce/`

comp.infosystems.wais

comp.infosystems.wais covers WAIS topics, including integration of WAIS with WWW.

comp.text.sgml

comp.text.sgml covers SGML, including HTML.

Archives of These Groups

Archives of the **www-announce**, **net-happenings** mailing lists and **comp.infosystems.www.*** Usenet groups are available at the URL `http://cair-archive.kaist.ac.kr/Archive/Announce/`

CERN Mailing Lists

To join a list at CERN send electronic mail to **listserv@info.cern.ch** with the line **subscribe** *www-list your name*.

For example if John Smith wanted to subscribe to the **www-announce** list he would send the message **subscribe www-announce John Smith**

An overview of CERN mailing lists is available at the URL
http://info.cern.ch/hypertext/WWW/Administration/Mailing/Overview.html
Alternatively send an email message to **listserv@info.cern.ch** containing line **lists** to receive a list of lists or **review** *list* to receive a list of subscribers to *list*.

www-announce

www-announce is for anyone interested in WWW, its progress, new data sources, new software releases. Please refrain from posting administrivia to this list! The list owners want to keep it low volume, large membership.

www-html

www-html is for technical discussions of the HyperText Markup Language HTML. Design discussions only, please, not newcomer questions.

This list is archived at the URL http://198.92.133.3/menus/6581.htm and at the URL http://gummo.stanford.edu/html/hypermail/hypermail.html

www-rdb

www-rdb is for discussion of gatewaying relational databases into WWW. This list is archived at the URL http://info.cern.ch/hypertext/WWW/Archive/www-rdb

www-proxy

www-proxy is for technical discussion about WWW proxies, caching, and future directions. This list is archived at the URL
http://info.cern.ch/hypertext/WWW/Archive/www-proxy

www-talk

www-talk is for technical discussion for those developing WWW software or with that deep an interest. (Please keep this to WWW technical design only. **Not** general questions from non-developers, which should go to the newsgroup, **nor** for HTML topics which should go to www-html.)

This list is archived. A threaded version of the archive is available at the URL
http://gummo.stanford.edu/html/hypermail/archives.html.

Other Mailing Lists

atmwww-l

atmwww-l is an open and unmoderated discussion of the impact of Asynchronous Transfer Mode (ATM) technology and networking on the World-Wide-Web. To subscribe to the list send the message **subscribe atmwww-l** *your name* to the address **listserv@cmuvm.csv.cmich.edu** in the To send messages to the **atmwww-l** discussion list, email: **atmwww-l@cmuvm.csv.cmich.edu**

cello-l

cello-l is a discussion list for users of the Cello WWW browser. To subscribe to the list send the message **sub cello-l** *your name* to the address **listserv@cornell.edu** Further information is available at the URL ftp://ftp.law.cornell.edu/pub/LII/Cello/default.htm The Cello FAQ is available at the URL Archives of the list are available at the URL gopher://gopher.law.cornell.edu:70/11/listservs/cellol

html-wg

html-wg is a mailing list for an IETF working group which is discussing developments of HTML. To subscribe email **html-wg-request@oclc.org** with the message **SUBSCRIBE html-wg** *yourname* An archive of the list is available at the URL http://www.ics.uci.edu/pub/ietf/html/

http-wg

The HTTP working group (**http-wg**)will work on the specification of the Hypertext Transfer Protocol (HTTP). HTTP is a data access protocol currently run over TCP and is the basis of the World-Wide Web. The initial work will be to document existing practice and short-term extensions. Subsequent work will be to extend and revise the protocol. Directions which have already been mentioned include:

- improved efficiency,
- extended operations,
- extended negotiation,
- richer metainformation,
- ties with security protocols.

To subscribe email **http-wg-request@cuckoo.hpl.hp.com** with the message **SUBSCRIBE http-wg** *yourname* An archive of the list is available at the URL http://www.ics.uci.edu/pub/ietf/http/hypermail/

libwww-perl

libwww-perl is a library of Perl4 packages which provides a simple and consistent programming interface to the World-Wide Web. This library is being developed as a collaborative effort to assist the further development of useful WWW clients and tools.

A mailing list has been established for technical discussion about libwww-perl, including problem reports, interim fixes, suggestions for features, and contributions. The mailing list address is libwww-perl@ics.uci.edu and administrivia (including subscribe requests) should be sent to **libwww-perl-request@ics.uci.edu**

A Hypermail archive of the mailing list is also available at the URL http://www.ics.uci.edu/WebSoft/libwww-perl/archive/

mosaic-l

mosaic-l is a Listserv list for the NCSA Mosaic WWW browser. To subscribe send the message **`subscribe mosaic-l`** *`firstname lastname`* to the address **`listserv@uicvm.uic.edu`**

NOTE This list is now believed to be defunct since it was being used for basic Mosaic questions, rather than providing a forum for Mosaic developers.

MacHTTP-talk

MacHTTP-talk is a mailing list for MacHTTP users has been set up. It provides an open forum for any questions, answers, suggestions, announcements, etc. about the MacHTTP server software. To subscribe to the list send a mail message to the address **`MajorDomo@academ.com`** containing the message **`subscribe MacHTTP-talk`** *`firstname surname`*

Further information is available at the URL `http://www.uth.tmc.edu/mac_info/machttp/mailing_list.html`

moo-www

moo-www is a mailing list to discuss links between MUDS, in particular systems based on Pavel Curtis's MOO server, and the World-Wide Web. Subjects for discussion include:

- In-MOO clients to access the web.
- Web clients that understand how to talk to MUDs.
- Using the Web to provide Multimedia capabilities to MUDs.

The list is at `moo-www@maths.tcd.ie` Subscription requests should go to **`moo-www-request@maths.tcd.ie`**

Netscape

Netscape is a Listserv list for the Netscape WWW browser. This list is for the purpose of discussing features and bugs contained in this new browser, as well as the new tags Netscape implements. To subscribe send the message **`subscribe netscape`** *`firstname lastname`* to the address **`listserv@irlearn.ucd.ie`**

Quality

Quality is a mailing list for the discussion of quality issues. To subscribe to the list send the message **`subscribe quality`** to the address **`listmanager@naic.nasa.gov`**.

An archive is available at the URL `http://naic.nasa.gov/naic/archives`

unite

unite is a Mailbase list which can be used for discussions about a *User Network Interface To Everything*. Based in the UK with an international membership.
`mailbase@mailbase.ac.uk` with the message **`join unite`** *`yourname`*

The archives are available at the URL `gopher://mailbase.ac.uk/11/lists-u-z/unite`

Web4Lib

Web4Lib is a list for Library-based WWW managers and developers. To subscribe email **`listserv@library.berkeley.edu`** with the message **`SUBSCRIBE Web4Lib`** *`yourname`*

web-support

web-support is a Mailbase list which can be used for discussions about WWW issues. Based in the UK. To subscribe email **mailbase@mailbase.ac.uk** with the message **join web-support** *yourname*

The archives are available at the URL `gopher://mailbase.ac.uk/11/lists-u-z/web-support` A Hypernews version of the archive is available at the URL `http://www.mailbase.ac.uk/hypermail/web-support/archive.html`

WebServer-NT

The **WebServer-NT** mailing list is intended as a forum where users of Windows NT can discuss World-Wide Web server issues. Likely topics might include (but are not limited to):

- ideas for novel ways of using WWW servers on Windows NT
- security issues on NT Web Servers (encryption, proxy servers, ACLs)
- integration of Web Servers with NT applications
- CGI scripts wanted/available for Windows NT
- features which people would like to see in future Web Server
- products for Windows NT
- the Freeware Web Server (HTTPS) for Windows NT
- configuration and document management issues for NT-based Web Servers

To subscribe, send message to **webserver-nt-request@mailserve.process.com** and in the message body type **SUBSCRIBE webserver-nt**

To get help on the mailserver commands put **HELP** in your message body To receive a list of the available mailing list put **LISTS** in the body To receive a list of subscribers in a list put **SEND/LIST webserver-nt** in the body.

www-buyinfo

Discussions of issues of commercial transactions of information via the Web take place on the **www-buyinfo** mailing list. To subscribe send the message **subscribe www-buyinfo** to the address **www-buyinfo-request@allegra.att.com**

The archives are held at the URL `http://www.research.att.com/www-buyinfo/about.html`

www-courseware

www-courseware is a list dedicated to courseware on WWW. To subscribe send mail to **www-courseware-request@eit.com** containing the message **subscribe**

An archive of the list is held at the URL `http://www.eit.com/mailinglists/www-courseware/archive/`

wwww-literature

This is a list dedicated to literature on the WWW. To subscribe send mail to **www-literature-request@eit.com** containing the message **subscribe**

An archive of the list is held at the URL `http://www.eit.com/mailinglists/www-literature/archive/`

www-managers

The aim of this list if to provide a high signal-to-noise, quick turn-around forum for managers of WWW servers and sites to get answers to specific questions about the setup and

maintenance of http servers and clients. The mailing list is managed by a utility called majordomo. To subscribe send the message **`subscribe www-managers`** to the address **`majordomo@lists.stanford.edu`**

www-security

www-security is a list to discuss different methods of providing a secure WWW service. The list will focus on how to secure HTTP and/or HTTP-like protocols to provide privacy, user authentication, service certifications and document checking (digital signatures).

To subscribe send mail to **`www-security-request@nsmx.rutgers.edu`** containing the message **`subscribe www-security`**

An archive of the list is held at the URL `http://www.verity.com/www-security.html`

Information about the www-security list is also available at the URL `http://www-ns.rutgers.edu/www-security/index.html`

www-speed

The **www-speed** list is dedicated to the proposition that the web is just too darned slow, and that some of its key components have inherent performance problems that cannot be dealt with without changes to protocols. Topics appropriate to the list are:

- Performance issues in existing implementations
- Managing busy servers
- Benchmarking servers
- High performance Server design
- High speed (130Mbps+) issues
- Client implementation
- HTTP - beyond 1.0
- Other performance and scalability issues

The list address is **`www-speed@tipper.oit.unc.edu`** The request address is **`www-speed-request@tipper.oit.unc.edu`**

www-vrml

VRML (the Virtual Reality Markup Language) is an evolving specification for a platform-independent definition of 3-dimensional spaces within the World-Wide Web. It is designed to combine the best features of virtual reality, networked visualization, and the global hypermedia environment of the World-Wide Web.

To subscribe to the Virtual Reality Markup Language (VRML) list send mail to **`majordomo@wired.com`** containing the message **`subscribe www-vrml`**

Further information is available at the URL `http://www.wired.com/vrml/`

www@unicode.org

www@unicode.org is intended for indepth technical discussions of the possibility of modifying the WWW protocols to support Unicode. It is going along the same lines as some of the Unicode discussions on **www-talk**, just a more focused group with no other WWW issues. If interested in joining this list, send email to `to` **`www-request@unicode.org`** with a subject line of **`subscribe`**, and a message body of subscribe **`www@unicode.org`** *`your name`*

Appendix 2 WWW Resources

A wide range of resource materials about the World-Wide Web are available on the World-Wide Web. A number are listed below.

WWW Online Resources

The **World-Wide Web Developer's Library** is available at the URL
`http://www.stars.com/`

Spider's Web is available at the URL
`http://gagme.wwa.com/~boba/spider.html`

Yahoo is available at the URL
`http://akebono.stanford.edu/yahoo/Computers/`

Computers: World-Wide Web is available at the URL
`http://akebono.stanford.edu/yahoo/Computers/World_Wide_Web/`

One World is available at the URL
`http://oneworld.wa.com/htmldev/devpage/dev-page1.html`

Web Weaver's Page is available at the URL
`http://www.nas.nasa.gov/RNR/Education/weavers.html`

WebStars: Astrophysics in Cyberspace is available at the URL
`http://guinan.gsf.nasa.gov/`

Pointer's to WWW resources (Toronto University) is available at the URL
`http://www.utirc.utoronto.ca/`

PC Week's pointers to WWW resources is available at the URL
`http://www.upcweek.ziff.com/~pcweek/pointers.html`

Oslonett is available at the URL
`http://www.oslonet.no/html/demo/WWWinfo/html`

CGI Programmer's Reference is available at the URL
`http://www.halcyon.com/hedlund/cgi-faq/`

The WWW Locator Guide is available at the URL
`http://groucho.gsfc.nasa.gov/Code_520/locator/locator.html`

A list of **World Wide Web FAQs and Guides** is available at the URL
`http://cuiwww.unige.ch/OSG/FAQ/www.html`

WWW Icons and Clip Art

A list of online resources of icons and clip art which can be used to produce HTML documents containing graphics is given below. Note, however, that before using graphics in HTML documents you should be aware of the additional loads which will be placed on network and servers.

`ftp://ftp.cica.indiana.edu/pub/win3/icons`

`http://white.nosc.mil/images.html`

`http://guinan.gsfc.nasa.gov/Alan/Richmond.html`

`http://www.cli.di.unipi.it/iconbrowser/icons.html`

`http://www.jsc.nasa.gov/~mccoy/Icons.index.html`

`http://www.cs.yale.edu/HTML/YALE/CS/HyPlans/loosemore-sandra/clipart.html`

WWW Conferences

Conference proceedings from the first WWW conference, *WWW '94*, held at CERN on 25-27 May 1994 are available at the URL `http://www.elsevier.nl/`

Further information about the second WWW conference *Mosaic and The Web*, held at Chicago on 17-20 October 1994 is available at the URL `http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings/` A searchable index of the papers is available at the URL `http://www.verity.com/spidersearch.html`

The third WWW conference will be held at Darmstadt, Germany on 10-14 April 1995. Further details are available at the URL `http://www.igd.fhg.de/www95.html`

Other Resources

WWW Information At CERN

Information about the World-Wide Web Initiative is available at the URL `http://info.cern.ch/hypertext/WWW/TheProject.html`

Best of the Web

The Best of the Web awards promotes WWW to new and potential users and helps information providers by demonstrating what can be done on WWW. The award winners and entrants are available at the URL `http://wings.buffalo.edu/contest/`

WWW FAQ

The WWW Frequently Asked Questions (FAQ) is available at the URL `http://sunsite.unc.edu/boutell/faq/www_faq.html`

Entering The World-Wide Web: A Guide to Cyberspace

Kevin Hughes' *Entering The World-Wide Web: A Guide to Cyberspace* is available at the URL `http://www.eit.com/web/www.guide`

Information Superhighway in the UK

Information about the Information Superhighway in the UK is available at the URL `http://tin.ssc.plym.ac.uk/up.html`

Appendix 3 National UK Services

Services

The **Bulletin Board For Libraries** (**BUBL**) holds a wide range of information of interest to anyone involved with libraries in education. Further information is available at the URL `http://www.bubl.bath.ac.uk/BUBL/home.html`

The **Mailbase** mailing list service run a WWW server which is available at the URL `http://www.mailbase.ac.uk/`

The **Micros Hensa** service run a WWW server which is available at the URL `http://micros.hensa.ac.uk/`

The **Unix Hensa** service run a WWW server which is available at the URL `http://unix.hensa.ac.uk/`

CTISS run a WWW server which is available at the URL `http://www.ox.ac.uk/cti/`

The **Office for Library and Information Networking** (**UKOLN**) runs a WWW server which is available at the URL `http://ukoln.bath.ac.uk/UKOLN/home.html`

NISS runs a WWW server which is available at the URL `http://www.niss.ac.uk/`

A **TLTP** specific Web Server is available at the URL `http://www.icbl.hw.ac.uk/tltp`

The **Social Sciences Information Gateway** is available at the URL `http://sosig.esrc.bris.ac.uk/`

CCTA, the UK Government computer agency, runs a WWW server which is available at the URL `http://www.open.gov.uk/`

Directories

A list of **United Kingdom Based WWW servers** is available at the URL `http://src.doc.ic.ac.uk/all-uk.html`

A **UK tourist guide** is available at the URL `http://www.cs.ucl.ac.uk/misc/uk/intro.html`

A **UK sensitive map** is available at the URL `http://scitsc.wlv.ac.uk/ukinfo/uk.map.html` This service is maintained by the School of Computing and Information Technology, University of Wolverhampton (email `jphb@scitsc.wlv.ac.uk`)

WAIS Resources

The following WAIS services are provided by NISS.

NISS Bulletin Board

A wide range of information of interest to varying sectors of the academic community. This service is available at the URL `wais://gopher.niss.ac.uk/NISSBB`

World Factbook

Basic details (population, climate, main industries and so on) for the countries in the World. Use a search term such as the name of a country to locate particular records. This service is available at the URL `wais://wais.niss.ac.uk/World_Factbook`

Roget's Thesaurus

The 1911 edition (enhanced with an additional 1,000+ words not included in the original version) of the ever-useful thesaurus of the English language. Use any word as your search term. This service is available at the URL `wais://wais.niss.ac.uk/Roget`

JANET News

JANET News contains material about the JANET computer network, such as registered domain names and addresses, and information about gateways to other networks. This service is available at the URL `wais://news.janet.ac.uk/JANET.news`

CHEST Directory

The CHEST Directory of software is available at the URL `wais://wais.niss.ac.uk/CHEST_Directory`

Appendix 4 Conferences On WWW

Bruce Altner (mailto:ari@clark.net), the Director of Technical Services of ARInternet Corporation has a vision for gatherings at the electronic meeting hall combines the best features of the WWW (browsing, multimedia and hypertext capabilities, searching and information retrieval, file downloading and e-mail communication, to name just a few) within the format of the traditional poster paper session.

Electronic Conferences and Workshops

Here are some real life examples of Electronic Conferences and Workshops:

ChemConf'93 is available at the URL
gopher://info.umd.edu:901/11/inforM/Educational_Resources/Faculty_Resources_and_Support/ChemConference

NASA High Alpha Conference IV (high angle of attack) is available at the URL
http://www.dfrf.nasa.gov/Workshop/HighAlphaIV/highalpha.html

The HIDEC Electronic Conference (the F-15 Highly Integrated Digital Electronic Control program) is available at the URL
http://mosaic.dfrf.nasa.gov/Workshop/HIDEC/Conf.DIRS/.htmllinks/ConfWeb.html

DL94:Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries is available at the URL http://atg1.wustl.edu/DL94

On-Line Proceedings of ACL-94 (Association of Computational Linguistics) is available at the URL http://xxx.lanl.gov/cmp-lg/ACL-94-proceedings.html

...and its post-conference workshops is available at the URL
http://xxx.lanl.gov/cmp-lg/ACL-94-post.html

1st Electronic Conference in Computational Chemistry (ECCC) is available at the URL
http://hackberry.chem.niu.edu:70/0/ECCCinformation.html

Reviews of Electronic Conferences

A discussion of the pros and cons of this type of online gathering, written by the ChemConf'93 organizer Dr. Tom O'Haver, is available at the URL
gopher://info.umd.edu:901/00/inforM/Educational_Resources/Faculty_Resources_and_Support/ChemConference/BackgroundReading/OnlineConferencing.txt

And as a wonderful example of self-referencing, a la Douglas Hofstadter's Godel, Escher, and Bach, see the URL http://www.automatrix.com/conferences

An example of an "after-the-fact" online conference is available at the URL
http://stardust.jpl.nasa.gov/igarss/

TaTOO '95

TaTTOO '95 On-Line used state-of-the-art technology combining interactive multi-user virtual environments with the World-Wide Web to bring an International Conference and Trade Exhibition to the desktop. In the virtual conference objects such as delegates, rooms, personal business cards and leaflets could all be browsed on the Web. A virtual exhibition took place in TaTTOO/MOO, a virtual environment. The MOO was extended to make the objects in it Web-aware, so it was possible to browse the system using a Web client. TaTOO '95 is available at the URL
http://www.cms.dmu.ac.uk/Research/OTG/Online/live-announce.html

Appendix 5 References

Books

"*Spinning the Web: How to Provide Information on the Internet*" by Andrew Ford, to be published by Van Nostrand Reinhold, New York (ISBN 1-850-32141-8) and International Thomson Publishing, London (ISBN 0-442-01962-9). The book describes how to run a web site, which covers creating material for dissemination via the Web and setting up and running a web server. It describes HTML in detail and includes a tear-out HTML reference card and a resource guide. This book is recommended by the author of this handbook.

"*Mosaic Quick Tour For Windows*" by Gareth Branwyn, published by Ventana Press costs £7.95 (ISBN 1-56604-194-5). Further information available at the URL `http://www.vmedia.com/vvc`

"*The Internet via Mosaic and World-Wide Web*" by Steve Browne, published by ZD Press costs £22.99 (ISBN 1-56276-259-1).

"*The World-Wide Web, Mosaic and More*" by Jason J Manger, published by McGraw Hill costs £24.95 (ISBN 0-07-709132-9).

"*Teach Yourself HTML Web Publishing in a Week*" by Laura Lemay, to be published by Sams' Publishing (ISBN 0-672-30667-0). This book discusses not only the various aspects of HTML, Web servers, gateways, forms, and imagemaps, but also focuses strongly on style and structure and navigation. In other words, its not just a reference, its also a style guide.

"*HTML For Fun and Profit*" by Mary Morris, to be published by Prentice-Hall. It includes forms, clickable images, server includes, indexing, linking and basic formatting. It will have a CD-ROM with examples and tools on it. See the URL `http://www.sun.com/smi/ssoftpress/`

"*The Mosaic Handbook for the X Window System*" by Richmond Koman and Paula Feguson, published by O'Reilly (ISBN 1-56592-095-3), "*The Mosaic Handbook for Microsoft Windows System*" by Richmond Koman, published by O'Reilly (ISBN 1-56592-094-5) and "*The Mosaic Handbook for the Macintosh*" by Richmond Koman, published by O'Reilly (ISBN 1-56592-096-1). These books, which cost £22 each, contain a CD-ROM (the X Window book) or a floppy disk which contains a copy of the Mosaic software.

Magazines

Many magazines are being published which cover various aspects of the Internet. The following list gives some of the main ones, including ones published in the UK.

.net published by Fortune Publishing Ltd. Further details are available at the URL `http://www.futurenet.co.uk/home.html` or by sending email to `netmag@futurenet.co.uk`

infoHighway ISSN 1355-2465. For further details send email to `p.deacon@eurodollar.co.uk` or `david@pipex.net`

Wired. Further details are available at the URL `http://www.hotwired.com/` For subscriptions details send email to `subscriptions@wired.com`

3W cost £24 for 6 issues. Further information is available at the URL `http://www.3w.com/3W/`