



**The World Wide Web**  
**A Strategic Tool for UK Higher Education**



Part of the JISC New Technologies Initiative

---

## CONTENTS

---

Executive Summary and Recommendations	1
Programme of Conference and Workshop	7
Participants at the Conference and Workshop	9
Introductory Comments — Anne Mumford	11
An Introduction To The World-Wide Web — Brian Kelly	13
A Campus Wide Code of Practice for Mounting Material on a WWW — Bob Hynds	23
Other Guidelines	30
Legal issues of WWW & Electronic Publishing — Andrew Charlesworth	31
JANET Acceptable Use — Alistair Chalmers	53
ISSC - Spider or Fly? — Derek Law	56
Exploiting and Protecting the Network — Peter Kemp	59
Approaches to Wide Area Indexing — Martijn Koster	61
The provision of effective local, national and international Web services Neil Smith	63
Teaching and Learning Issues — Roger Rist	65
WWW as a Strategic Tool in Research and Teaching of Molecular Sciences Henry S Rzepa	69
World Wide Web as a Tool for Teaching and Research in Visualization Ken Brodlie	72
Global Network Academy — Alan Mills	74
WWW - opportunities and Obstacles for administrators — Peter Tinson	75
Web Futures — Brian Kelly	81
Discussion Sessions	83
Teaching, Learning and Research Issues	83
Supporting Information Providers	85
Management Issues	88
Technical Issues	89

---

## EXECUTIVE SUMMARY AND RECOMMENDATIONS

---

### Introductory comments

This is the report of the workshop organised by AGOCC as part of the Support Initiative for Multimedia Applications (SIMA) funded by the JISC New Technologies Initiative. The event was held at Loughborough University on 13/14 February 1995. The first day consisted of papers from experts in the field. The second day built on that through group discussions of the issues raised to result in recommendations to the community for activity and funding. The first day was attended by 80 participants from 54 institutions of whom 41 stayed for the second day.

The World-Wide Web (WWW) provides a distributed multimedia hypertext system which can be used in teaching, research and administration. Using WWW we can:

- retrieve multimedia documents from around the world
- publish documents globally
- run programs on remote servers
- download and run programs on local machines

Many institutions have developed Campus Wide Information Systems (CWISs) using WWW. Students can access resources from around the world to assist them in their learning. Also, and perhaps more significantly, a number of distance learning courses are now being provided on the WWW. Researchers are beginning to take advantage of control systems and collaborative tools which are emerging. Conferences, including reports, video, slide images etc. can be published on the WWW, allowing greater and more rapid dissemination.

The WWW is clearly an important tool for many applications and is expected to be so for some years to come. This workshop was concerned with addressing the issues in the context of UK Higher Education and with making recommendations for activity and funding which would help UK Higher Education make efficient and effective use of this important tool.

Despite the apparent popularity of the WWW, many sites have not installed a server and most do not have well developed strategies. One thing which did emerge from the workshop is the need for support of those sites in developing servers and strategies based on good practices and experiences elsewhere.

## The World-Wide Web — A Tool for Many Applications

### Providing Information on the WWW

The World-Wide Web was created to support remote collaborative research, but it has developed primarily as a means of providing information which can be linked to other information sources. It is an essential medium for accessing, delivering and exchanging information. There needs to be an emphasis on the provision of quality information using the tools in an appropriate way. The provision of information and effective use of the tools available requires specialist expertise which needs to be recognised at sites. Training and access to appropriate tools (browsers, authoring, editing, conversion, validation tools) are needed to ensure good use of the WWW for information provision.

## **Teaching and Learning**

The WWW is providing a number of opportunities for teachers and students. Resources can be accessed which might otherwise have been unavailable. These include virtual libraries and museums. Other resources can be built up and used by students, for example questions and answers which can be searched or routed through to an expert if it is a new query and then the answer logged for future use. Teaching programs can be accessed and used by students as part of their modules. There are also a number of complete distance learning modules being mounted on the WWW.

There are a number of issues relating to the use of the WWW in teaching and learning. These include: how widely do we wish to avail people in other institutions worldwide of our best resources; how can we judge the quality of the information; where is the competitive edge in this global information age; how can we make most effective use of the opportunities and how can we best train staff and students to do so?

We need to help teachers with the pedagogic issues associated with this medium. There is a need to equip students with the necessary research and information access skills. This points to the need for training of teachers, library and computing staff.

## **Opportunities for Administrators**

The WWW offers administrators opportunities for a number of their areas of work. Any such change would need a change of culture towards more open access to information while making use of emerging tools for security where appropriate. The exchange of information is clearly one area of expansion, although there may be security considerations for much of this. We are already seeing the Times Higher providing job advertisements on the WWW and sites advertising job details in a similar way. It is a potential way of promoting the institution through having the prospectus available — perhaps less to schools in the UK, but certainly to potential overseas postgraduates. The availability of papers from central online sources may make them more quickly available than traditional methods of departmental circulation on paper. There are plenty of possibilities and administrators need to be advised on how best to take advantage of the tools.

## **Developing National and Institutional Strategies**

### **WWW - an Important Network Tool**

The WWW is clearly becoming an important network tool and its use is growing. It has the potential to swamp network use with users accessing images and connecting to information round the world. In order to make best use of the medium, we need to have appropriate policies at national and institutional levels. These policies need to include: caching strategies which locate caches at bottlenecks; replication of frequently accessed resources; indexing of information to facilitate access. They also need to include strategies at national and institutional levels as to the kind of use which may be made of the WWW in terms of information provided, linked to and accessed. These need to be in line with JANET Acceptable Use policies and with Guidelines from UCISA.

### **National Caching and Indexing Strategies**

One of the problems many people find when using the WWW is the delay as (sometimes gratuitous) images - or sometimes just text - are downloaded, particularly from the US. Other bottlenecks may occur locally. It is important to minimise these bottlenecks for the sake of all network users. The WWW may be the "killer application" for the network but perhaps we do not wish it to soak up all the available bandwidth. In order to maximise efficiency we need to develop national and local caches. HENSA have a national cache, but national facilities need further consideration. Local sites also need to be advised on using caches. There is also the need to build on current research in this

area and consider more sophisticated strategies. Replication and cascading caches also need to be considered.

Indexing is a further way of making Web use more effective. We should encourage people to make use of tools when they develop pages and use the WWW. There is a need to make use of more advanced tools which can help automate the process. Good indexing does require a high level of manual resource. The community needs to address how best to allocate resources to making information accessible. The Social Sciences Information Gateway (SOSIG) is an example of effort being put in to making a wide range of quality information available by reviewing potential sources, indexing and describing them.

Some sites are providing indexing tools, such as ALIWEB, as part of their high level information.

## **Legal Issues**

There is considerable concern regarding the legal issues of using the network generally and WWW in particular. There are many issues regarding both individual, site and community responsibilities. The individual has a right to have their work protected under copyright law. A site has to consider issues such as freedom of expressions and the need to educate their staff and students in the use of new tools. Tight, restrictive use may not be the answer — education may be the best way forward. There is no case law in this area which makes advice difficult.

## **Institutional Strategies**

Institutions need to develop advice to staff and students. The advice needs to cover use of resources outside the institution and development of those within it, including personal home pages for both staff and students. A number of sites have developed such guidelines, including Imperial College and Edinburgh University. There are also the emerging guidelines on network use from UCISA and the JANET Acceptable Use Policy.

## **WWW Futures**

There is a great deal of momentum behind the development and support of the WWW. There are now development organisations based in Europe (EuroW3O based at INRIA in France) and the USA (W3O based at MIT). Fortunately the intention is to have membership which is linked. Developments of new versions of the mark up language and protocols are being progressed as well as standards for resource identification. Applications for collaborative work (for example using whiteboards) and remote teaching (through MOOs and MUDs) are pushing forward the possibilities for use of WWW. Significantly, we are also seeing commercial developments, for example the Netscape browser.

## **Recommendations**

### **Community Support on WWW**

There is a need to support the community in its introduction and use of the WWW. This could be achieved in a number of ways which might include: training; development of training materials; employment of people for specific tasks; ensuring dissemination of examples of good practices.

We should look to develop self-help on various aspect relating to the Web including:

- technical issues such as caching, indexing
- gathering and exchanging information on good practices

- exchange of advice and support on WWW tools

## **Training Programme**

We need to:

- develop (or build on existing resources) training materials on good practices which can be incorporated into courses at sites which should include social, legal and institutional issues as well as technological ones.
- run a seminar for Staff Development Officers about the WWW with the aim that this could be built into induction courses for staff at institutions
- hold subject-based symposia to discuss the ways in which the WWW could be used within specific disciplines including advice on pedagogic issues

## **Institutional Guidelines**

A mechanism should be established to provide exchange of information on institutional guidelines and good practices. These need to link to JANET Acceptable Use Policy and to UCISA Guidelines.

## **Legal Advice**

Legal advice on the use of network information should be sought and maintained. This might include taking advice now on particular issues in the absence of any case law. We should also look to set up some mechanism for access to ongoing advice and to keep a database of questions and answers. The CCTA mechanisms should be investigated.

## **Caching Strategies**

Caching could improve performance at bottlenecks thus improving on efficiency and providing a cost-effective way of delivering information. A national caching strategy and associated local caching strategies should be developed. In order to define the nature of this further discussion an evaluation should be carried out to define the location and requirements for caching at national level. In the future we need to address the use of cascading caches and replication.

We need to educate sites and individuals mounting information to use appropriate caching. This could be provided as an extension of the report by Brian Kelly on "Running a WWW Service" which is part of the AGOCG/SIMA output. We should document good practices.

## **Indexing**

We need to develop an indexing strategy to ensure efficient use of the WWW and thus of the network. We need to address how we can advise sites to link their information. There is a need for national indexing into UK HE sites with relevant information and links to site home pages.

## **Software Evaluation**

We should evaluate relevant software and advise CHEST on suitable software for which good pricing structures should be negotiated. The software includes: document creation/conversion and management tools; graphics tools; gateways; collaborative application tools; quality; validation; clients, helpers and servers.

## **Quality**

Quality of information is the key to the success of a WWW service. We should develop a code of practice relating to quality for providers of materials on the WWW.

## **WWW and Courseware Delivery**

We should encourage the production of good quality WWW-delivered courses. We should also promote examples of good practices which are currently available. One possible way of achieving this recommendation is to offer a competition and prizes rather than tendering for good material to be produced.

## **Site Contacts**

We should ensure that there is a WWW site contact at all sites who can answer questions both within the site and from outside (along the "postmaster" model). There needs to be local support at all sites.

## **W3O and EuroW3O**

The workshop welcomed the decision of JISC to join W3O and felt that the way the community can input ideas and receive information needs to be agreed and promoted widely.





---

## PROGRAMME

---

### Conference Day — 13 February 1995

- 09.30 Registration and Coffee
- 10.00 Welcome  
*Anne Mumford, Loughborough University*
- Providing Information on WWW  
*Brian Kelly, University of Leeds*
- 10.30 Coffee
- 11.00 Approaches to Wide Area Indexing  
*Martijn Koster, Software Engineer and Web Administrator, NEXOR*
- 11.30 A campus wide Code of Practice for mounting material on a WWW  
*Bob Hynds, Imperial College*
- 12.00 Legal Issues  
*Andrew Charlesworth, University of Hull*
- 12.30 Issues associated with the provision of effective local, national and international Web services  
*Neil G. Smith, HENSA Unix Administrator*
- 13.00 Lunch
- 14.00 WWW - Teaching and Learning Issues  
*Roger Rist, ICBL, Heriott-Watt University*
- 14.30 WWW — Opportunities and Obstacles for Administrators  
*Peter Tinson, City University*
- 15.00 The Role of the ISSC: Spider or Fly  
JANET Acceptable Use  
*Derek Law, Kings College, London*
- 15.35 Closing Remarks
- 15.45 Tea

## Workshop Day — 14 February

0915 Plenary Session

Introduction to the Issues

*Anne Mumford, Loughborough University*

Web Futures

*Brian Kelly, University of Leeds*

Introduction to Group Sessions

1015 Coffee

1045 Group Sessions

Teaching, Learning and Research  
Indexing Issues and Server Strategies  
Information Providers Tools  
Management Issues

1245 Lunch

1400 Plenary Session

Groups report back and make recommendations

National Strategies

*Peter Kemp, University of Glasgow*

1545 Tea and Depart

---

## PARTICIPANTS

---

Allison Alden	Goldsmiths College
Frances Allen	University of Edinburgh
John Arfield	Loughborough University
Emily Armstrong	Writtle College
Barry Barker	De Montfort University
Revel Barker	Cranfield University
Jackie Bettess	University of Durham
Amanda Boe	University of Sussex
David Boyd	R.A.L.
Angela Braithwaite	Oxford Brookes University
Ken Brodlie	University of Leeds
Chris Brown	University of Bath
Iain Burke	Glasgow Caledonian University
Malcolm Clark	University of Warwick
Sue Cunningham	University of Manchester
Andrew Dallgleish	University of Glamorgan
Steve Daniels	The Open University
Graham Daniels	Coventry University
Alan Dawson	Glasgow University
Peter Dunn	University of Warwick
Mary Faircloth	University of Hertfordshire
Brian Fitzpatrick	University College Salford
Tom Franklin	University of Manchester
William Gillies	University of Strathclyde
Mandy Goss	University of Plymouth
Ray Gregg	Queens University of Belfast
Martin Hamilton	Loughborough University
Mr M Hannavy	Essex University
Valerie Harmer	University of Surrey
David Harrison	Cardiff Institute of H.E.
Ronald Haynes	University of Bristol
Jim Hensman	University of Coventry
Clare Hetherington	Cheltenham & Gloucester College of H.E.
Bob Hopgood	R.A.L.
Roger Horton	University of Huddersfield
Max Hunt	Loughborough University
Demetra Katsifli	Kingston University
Jane Kirk	Nottingham Trent University
Rob Kirkwood	Loughborough University
Jon Knight	Loughborough University
Chris Lilley	Manchester University
Gary Martin	Sheffield Hallam University
Professor A R McElroy	Napier University
Alan Mills	Birbeck College
Peter Murray-Rust	Glaxo Research & Development
Jerry Niman	University of Birmingham
Alan Pibworth	Cranfield University
Andy Powell	University of Bath
Chris Prowting	Royal Military College of Science
Bronwen Reid	University of Hull
Ron Rogerson	JISC
Helen Sargan	University of Cambridge
Ted Smith	University of Central Lancashire

Alan Stanier  
Wilma Strang  
Dave Temple  
Ray Thompson  
John Townsend  
Owen Turney  
David Whitehurst  
Lili Yeadon

University of Essex  
University of Kent  
Loughborough University  
Sheffield Hallam University  
Edge Hill College of H.E.  
University of Westminster  
UMIST  
Loughborough College

---

## INTRODUCTORY COMMENTS — ANNE MUMFORD

---

An early paper about WWW by Tim Berners-Lee, Robert Cailliau and Jean-François Groff was presented at the 3rd Joint European Networking Conference in Innsbruck in May 1992. This conference was a networking conference that was addressing issues of users and user services on the network. There was a great deal of excitement around at the conference about the new tools called Gopher and Archie.

In that paper the authors say:

"The W3 project is not a research project, but a practical plan to implement a global information system"

and, again:

" The aims of the W3 initiative are twofold: firstly to make a single, easy user-interface to all types of information so that all may access it, and secondly to make it so easy to add new information that the quantity and quality of inline information will both increase".

The last few years have seen the massive popular growth in use of the network and the adoption of the ideas presented at the JENC conference worldwide. It is the World-Wide Web which has proved to be the popular tool. This has been particularly true as graphical interfaces to the structured hypermedia have been developed making the tool relatively accessible to a wide range of people.

Perhaps one could agree that the aim of the paper to increase quantity of information has been realised. The quality issue is a very real one. The issue of accessing relevant information from the quantity of information available is also a major problem we need to address.

This workshop is not about using the WWW — there seem to be many courses around on that topic. Rather, it is concerned with the issues presented to us as individuals, institutions and as a community as a result of the pervasive use of the WWW. It is about considering the role of the WWW and making recommendations to relevant bodies as to resource allocation and strategic directions which are needed if we are to make best use of the technology.

Reference

Berners-Lee, T.J., Cailliau, R., Groff, J, "The world-wide web", Computer Networks and ISDN Systems, Vol 25, No 4-5, p 454-459, 1992



---

# AN INTRODUCTION TO THE WORLD-WIDE WEB

## BRIAN KELLY

---

### Abstract

This paper gives an introduction to the World-Wide Web. It describes the client/server architecture of WWW and gives some examples of its use within the academic community.

## The World-Wide Web

The World-Wide Web (often referred to as WWW or W<sup>3</sup>) is a *distributed multimedia hypertext* system:

**Distributed:** Information on WWW may be located on computer systems from around the world.

**Multimedia:** The information can include sound, graphics and video, as well as text.

**Hypertext:** The information can be accessed by using simple hypertext techniques, such as clicking on a highlighted phrase.

## Client/Server Architecture

WWW is based on the *client/server architecture*, illustrated in Figure 1.

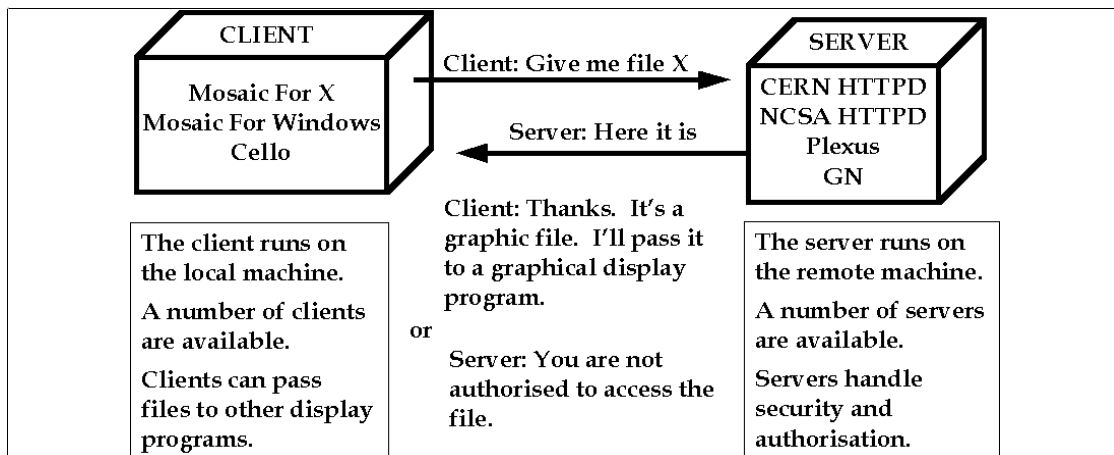


Figure 1 WWW Client/Server Architecture.

The client program (often called a browser) runs on the local workstation. When a hypertext link is activated, a request is sent to the server software, which normally runs on a remote server system. The server software (e.g. CERN httpd) will then transfer the file to the client.

### Servers

A variety of server software packages are available, including the widely CERN and NCSA httpd servers for Unix, commercial servers (which may be more reliable or provide additional features such as security) and servers for Apple Macintosh and the PC.

When setting up an institutional WWW service, the following server issues need to be addressed:

- Server hardware strategy. Will your institution run a central WWW system, or will the server software run on a number of departmental systems?
- Which server software?
- What server functionality will be required?
- Can the service be future-proofed, so that a change in the server software or server hardware strategy can be made with the minimum of inconvenience?

## Clients

A variety of browser clients are available, including

- NCSA Mosaic
- Netscape
- EInet (WinWeb, MacWeb)
- Enhanced Mosaic
- Lynx and DosLynx

When choosing a browser the following issues should be considered:

- The ease-of-use and robustness of the browser.
- Any security implications inherent in making the browser widely available.
- Performance implications.
- The background of the software developer and their target audience.
- The likelihood that desirable new facilities will be included in future releases.

## Extending WWW

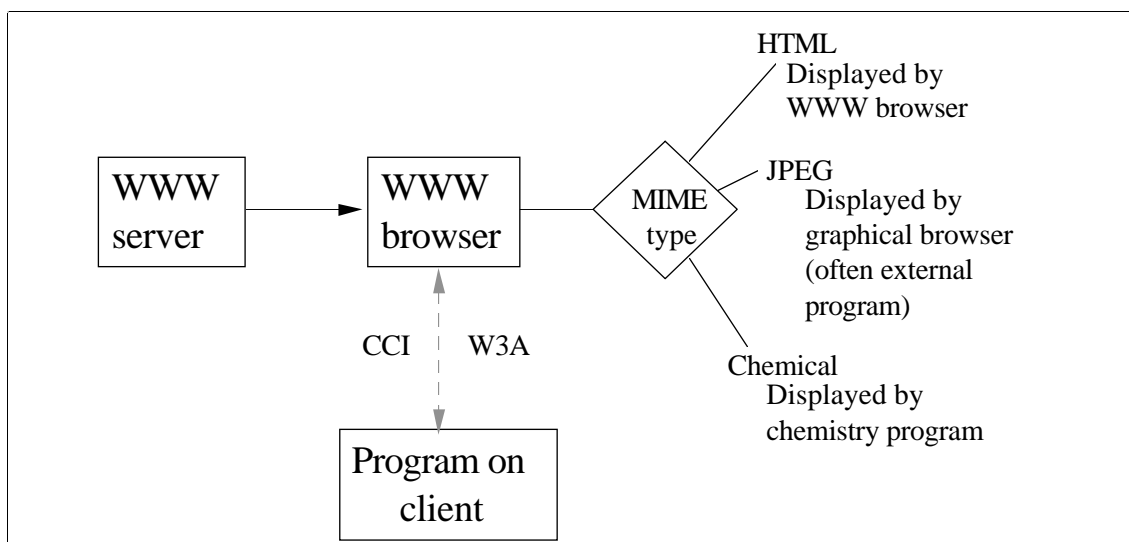
### Running Software On The Server

WWW can not only be used to deliver multimedia documents. The Common Gateway Interface (CGI) is a mechanism by which software running on the remote server can generate a *virtual document*. The use of CGI programs was used initially for developing searching tools. However it is now being used for developing distributed teaching and learning applications.

### Running Software On The Client

A WWW server delivers a MIME document to the browser. The browser will be configured to process MIME documents. For example HTML documents will be displayed by the browser, whereas the JPEG MIME type will normally be passed to an external graphical viewer. A Chemical MIME type has been developed by Henry Rzepa (Imperial College) and others. When such a MIME type is delivered, the browser can be configured to pass the document to a molecular modelling program.





**Figure 2 WWW MIME Types.**

This mechanism enables documents to be transferred from a remote server and passed on to an application on the local machine. New mechanisms are under development which will enable applications on the client to communicate with WWW browsers. The new mechanisms include CCI (Common Client Interface) and W3A (WWW Applets).

## Examples

WWW is many things to many people. It can be regarded as a system for the academic world, for the commercial world and for the cultural and social world. It can be regarded as a distributed file viewer or as a distributed program manager. It is also, ambiguously, referred to as the "Internet killer application". A number of examples of the use of WWW in the academic world are given below.

## Research Use

Many research papers, results, proposals, etc. are being made available on WWW [1]. One potential benefit of this is in making state of the art research material available to undergraduate students.

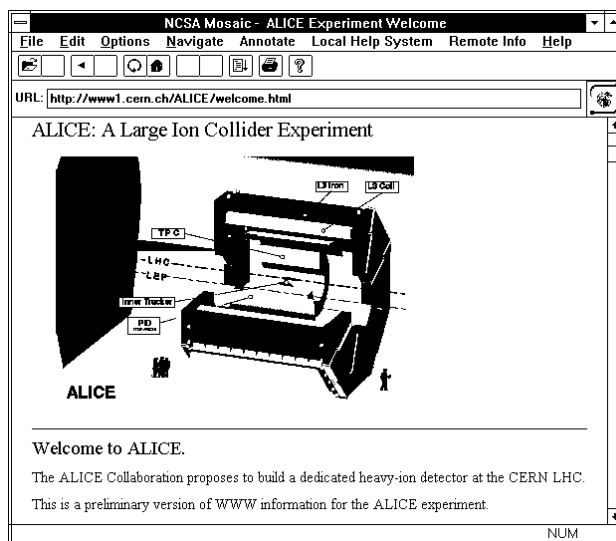


Figure 3 Experimental Research at CERN.

## Libraries

Libraries are beginning to make material available on WWW. This may provide a consistent interface between a library catalogue and other services on WWW.

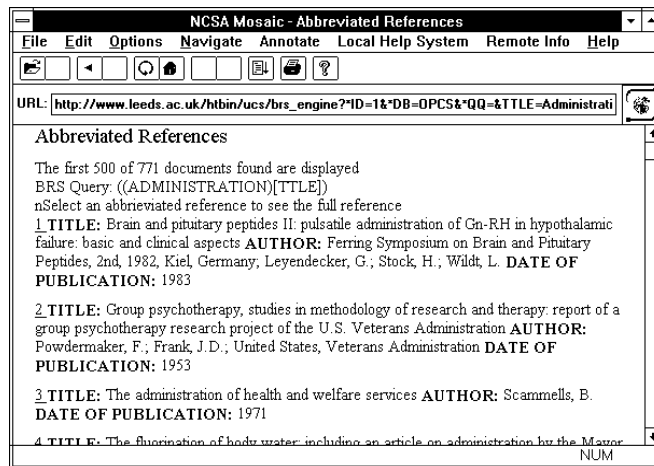


Figure 4 Libraries.

Figure 4 illustrates access to a backup copy of the Library catalogue at the University of Leeds [2]. The backup copy, held in a BRS free text retrieval system, is accessed on WWW by using a CGI gateway.

## CWIS

Campus Wide Information Systems (CWISes) are being developed on WWW. Figure 5 illustrates the CWIS at the University of Bradford [3].

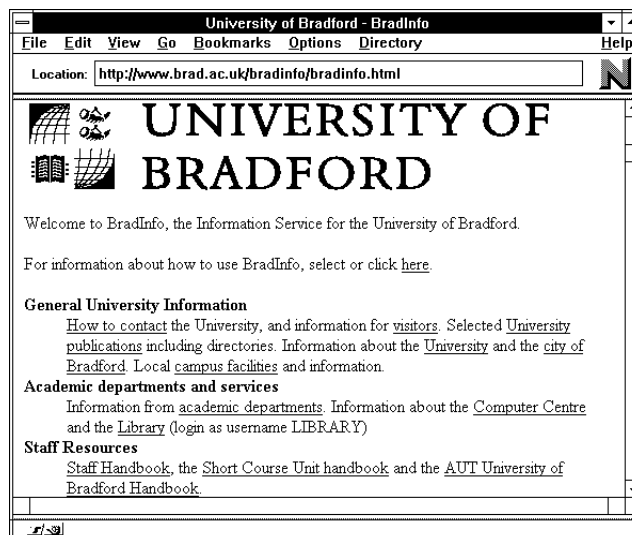


Figure 5 Bradford CWIS.

## Distributed Teaching and Learning

WWW is being used to develop distributed teaching and learning systems. The Frog Dissection Kit, illustrated in Figure 6, is one of the best known examples of a distributed teaching application.

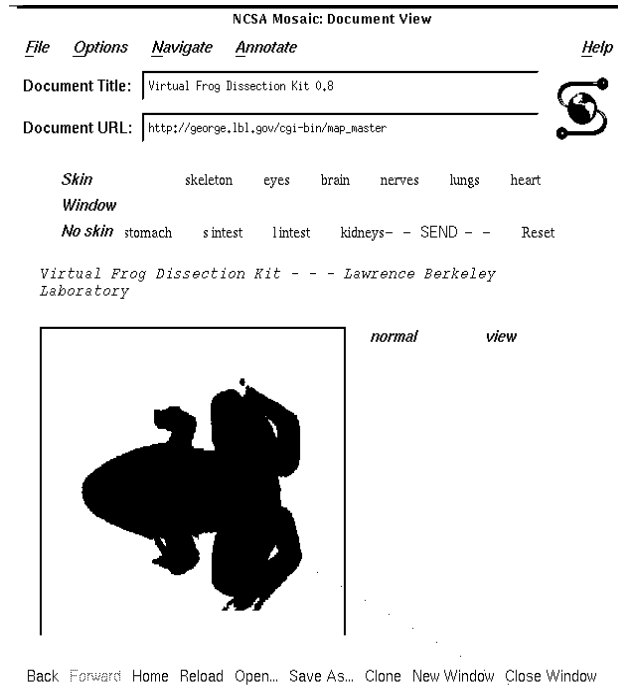


Figure 6 Frog Dissection Kit.

Students are being given the opportunity to develop their own authoring skills. As illustrated in Figure 7, Fine Arts students at the University of Leeds write multimedia essays. The essays will be used as an information resource by subsequent student groups.

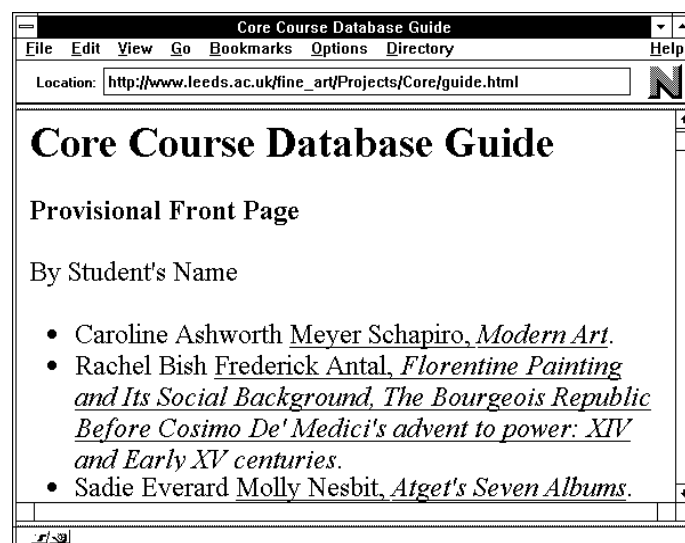


Figure 7 Student Authoring Skills.

## Visualisation

WWW is being used as a scientific visualisation tool. It may be difficult to visualise a molecule from simple reading text and viewing a 2D image.

As illustrated in Figure 8, the data used to produce the 2D image may be retrieved and loaded into a local molecular modelling package, enabling the molecule to be manipulated [6]. The technique can also be used for teaching and learning purposes.

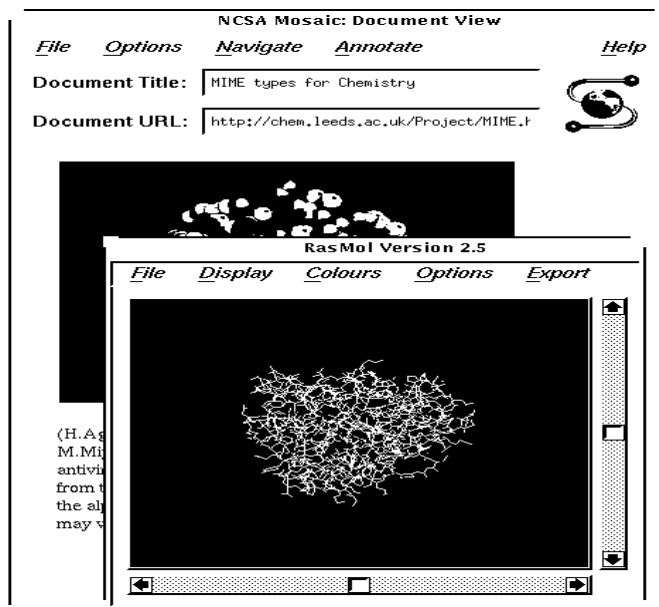


Figure 8 Visualisation.

## Control Systems

WWW is being used to provide integration with control systems. Mark Cox, University of Bradford has developed a robotic telescope which can be controlled using a WWW browser [7].

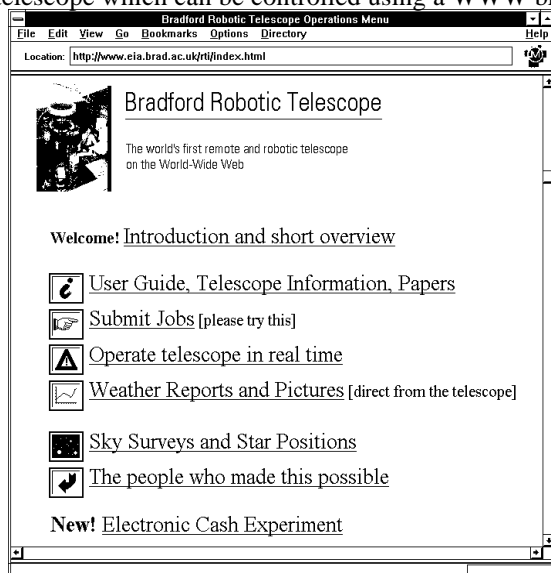


Figure 9 Control Systems.

## Conferences

Many conference organisers are making their conference proceedings available on WWW. The conference papers for the *Mosaic and the Web* conference held in Chicago in October 1994 were available on WWW at the start of the conference [8].

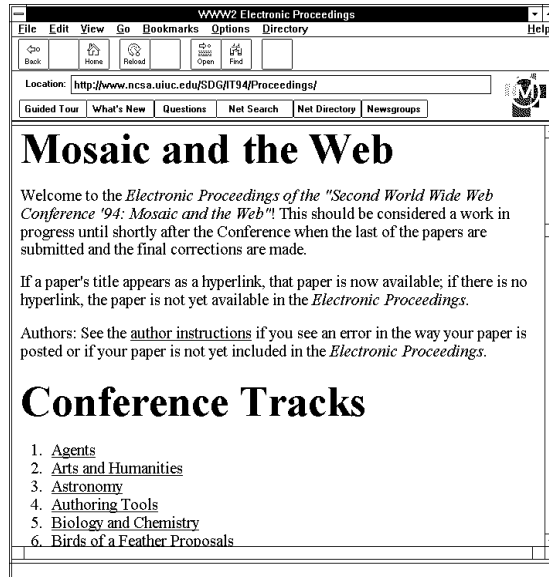


Figure 10 Conferences.

## Collaborative Systems

Collaborative systems are being developed on WWW. WIT (WWW Interactive Talk) is one example of a collaborative system which resembles Usenet, but provides a mechanism for voting on a topic.

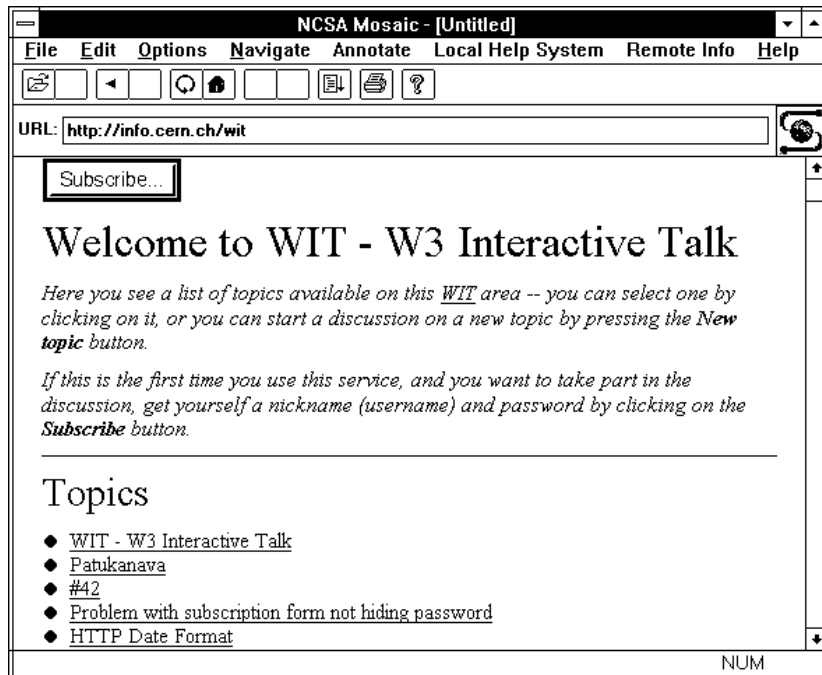


Figure 10 Collaborative Systems.

## WWW And Open Systems

It is likely that not only will we see increasing uptake of existing types of services on WWW, but also the development of new types of services. The development of WWW is helped by the open nature of WWW:

- Clients available on multiple platforms.
- Servers available on multiple platforms.
- Public domain clients and servers available.
- Specifications (HTML, HTTP) freely available.
- Specifications not controlled by single company/organisation.
- Involvement in specification developments open to anyone.

## References

- [1] Information about CERN research activities is available at the URL <http://www.cern.ch/>
- [2] The University of Leeds library catalogue is available at the URL <http://www.leeds.ac.uk/library/cats/backup.html>
- [3] The University of Bradford CWIS is available at the URL <http://www.brad.ac.uk/bradinfo/bradinfo.html>
- [4] The Frog Dissection Kit is available at the URL <http://george.lbl.gov/ITG.hm.pg.docs/dissect>
- [5] The Fine Art service is available at the URL [http://www.leeds.ac.uk/fine\\_art/Projects/Core/guide.html](http://www.leeds.ac.uk/fine_art/Projects/Core/guide.html)
- [6] Information about the Chemical MIME type is available at the URL <http://www.chem.leeds.ac.uk/novel.html>
- [7] The Robotic Telescope is available at the URL <http://www.eia.brad.ac.uk/mark/wwwf94/wwwf94.html>
- [8] The Mosaic and the Web conference proceeding are available at the URL <http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings>
- [9] WIT is available at the URL <http://info.cern.ch/wit>

---

## **A CAMPUS WIDE CODE OF PRACTICE FOR MOUNTING MATERIAL ON A WWW — BOB HYNDS**

---

The WWW provides unparalleled opportunities for universities to display their strengths world-wide, and all wish to seize the opportunity.

Because of its associations with UNIX and the Internet, WWW is perceived by senior academic managers as being in the province of the Centres. Thus they look to the Centre to see that the University has a good quality WWW which reflects credit on the institution.

The reality of the technical situation is that it is rather simple for anyone with a UNIX workstation to mount a WWW server, and any campus can expect to have a significant number of WWW servers on campus, with each one able to mount what ever the Webmaster chooses on the server. Thus unless there is an agreed set of campus guidelines for the mounting of material the Centres face a classical dilemma i.e. being held responsible by the university management without the power to control the situation. However, even if they were given such powers, they would still face the Catch 22 situation of being regarded by their users as exercising academic censorship, which indeed they would be.

Thus while WWW offers Centres the chance to earn brownie points with university senior managers, it also carries with it the danger of being crushed between a rock (of the University) and a hard place (the users).

How can Centres best deal with such a situation? In simple terms 'if you cannot beat 'em, join 'em'! Thus the need is to establish a set of guidelines to govern the material mounted on a WWW server that is acceptable to both the University and the users in general. How do we achieve such a situation? There are a number of ways, but this talk is concerned only with the way the issue has been dealt with at Imperial College.

An essential preliminary is to identify the issues that govern the operation of a successful WWW server and the possible role of the Centre in managing the overall operation of the WWW servers on campus. Once this is done a set of guidelines need be established that should be observed by contributors to any WWW server on campus. Then comes the hard part of persuading everyone from the Vice Chancellor downwards that the guidelines are fair, reasonable and acceptable to the majority of people on campus. Finally it is necessary to place the responsibility for observing the guidelines with the person who is designated as responsible for a particular WWW server.

So what are the relevant issues that need to be taken into account?

- WWW has tremendous potential to promote the University in a very positive way.
- WWW also has the potential to harm a University, the harm varying from a 'poor image' to a possible criminal or civil legal action.
- There are going to be a number of WWW servers on campus over which the Centre has no control.
- The Centre's role must be that of a technical supporter and co-ordinator of WWW.
- Any policy on what is permissible material to mount on a WWW server has got to be seen to be fair, reasonable and generally acceptable
- In practice the implementation of such a policy must be at the level of someone directly involved with a specific WWW server.
- It is not possible for the Centre to monitor all the material on all the WWW servers on campus.
- There are laws applicable to 'normal' publishing e.g. on Race relations, Equal opportunities, Libel and Copyright. Digitising the way the information is displayed does not negate these laws.
- It is necessary to establish College information which must be supplied via the central College WWW server e.g. details of fees, conditions for student eligibility for courses. Where these issues are mentioned on WWW servers on campus, there must be an active link back to the appropriate page on the central University WWW server.

There are some additional points that a Centre is wise to recognise:

- The success of a given server is directly related to the amount of effort expended in setting it up and particularly in the subsequent support.
- Only those people directly interested in the material mounted on a WWW server are going to be prepared to put effort into support, and in the long run this is likely to be people with modest computing skills.
- On a university campus there are going to be faculties e.g. Arts, Humanities who may not have the technical resources or skills to mount a WWW. Where this is true, they will need to rely on the Centre to set-up WWW servers for them, and to supply simple tools for the addition and editing of material.
- Students have a keen interest in WWW, and the Students Union should be provided with WWW facilities for itself and the clubs and societies.
- It is in the interests of the University to see that the home page of each WWW server on campus has some standard format that clearly labels it as a part of the University.



Taking into account the above issues, Imperial College assembled a Code of Practice which was discussed at a Heads of Departments meeting, agreed, and then authorised as College policy by the Rector. The Code itself, shown below, was formulated in the most general terms possible i.e. it tried to avoid proscriptive detail. Its form is also related to the management structure of the College i.e. as many functions as possible are devolved to departments, with Heads of Departments operating under a Memorandum of Understanding (MOU) between themselves and the Rector. Such a form may not be suitable in those institutions where e.g. faculties play a large part in the management of the University.

## **The Code of Practice**

There are three aspects of the use of WWW systems that need to be considered. These are:

1. The form of the first image presented by a WWW server i.e. the so called Home Page.
2. The authorisation of a WWW server for display on campus.
3. Ground rules and legal constraints on material used for a WWW server.

## **Form of the Home Page**

A Home Page usually contains the College Crest and the title Imperial College of Science, Technology and Medicine. It needs to be remembered that the College Crest and the Imperial College title are the property of the College and their style of use needs to comply with the conditions that are laid down by the College to protect its property and interests, including those imposed by the College of Arms. It is proposed that all Home Pages have a 'standard form' of the College Crest and the Imperial College of Science, Technology and Medicine banner (provided by the Centre for Computing Services), arranged in the form shown on the attached sheet, with any additional title banner e.g. department title, as shown.

Departmental WWW systems will obviously be concerned with information about their own department, but since the College is the 'sum of its parts' it is desirable that on the Home Page there should be three 'pointers' corresponding to :

- IC general information
- Other Academic departments
- Other UG and PG courses

All three of these pointers could point to the College WWW, which will contain all this information.

## **Authorisation of a WWW server**

The use of WWW offers the opportunity to make significant progress to the 'electronic campus'. However we need to recognise that the facility can be abused. If we are to exploit the opportunity then we need to avoid excessive bureaucracy, but the College must also be a position to show that it has taken all reasonable steps to prevent illegal activities associated with WWW.

There are three areas of the College that are likely to be implementing WWW systems. These are the central College i.e. Administration, Centre for Computing Services and the Library. The Departments (and research groups), Centres and Units. The ICU and its associated clubs.

The Centre for Computing Services is discussing with the Administration and the Library how best to implement a central College WWW. The Centre will provide the facility, and initially implement the mounting of material. When it is fully established it should represent a significant College asset, and the Rector will need to decide how best it should then be managed.

Several departments are already running WWW systems, and several research groups in one department. It would seem most appropriate if any WWW system associated with a department obtained the Head of Department's agreement before it was established, and the Head of Department informed the Centre of its existence. The Centre would then ensure that the Marketing Division was informed, and able to view it. Similar rules should apply to Centres and Units.

The ICU and its clubs will no doubt wish to use a WWW for informing students of campus activities. The Centre will provide a WWW facility for use of the Union, but would recommend that its viewing be restricted to the IC campus, and that the President of the Union be required to authorise its use by any club, and inform the Centre of each authorisation. The Union would need to be responsible for seeing that the use made of the WWW met the various legal and College requirements concerning material mounted on the system.

## **Ground Rules and Legal Restraints for Mounting Material on WWW systems**

Section IV of the Academic and Related Staff Handbook covers the general area of communications. The overall requirement is that staff members should seek the advice of the Press and Public Relations Office before producing material in the name of the College. In the context of WWW this is not practicable. The general rule must therefore be the requirement that no member of the College should act in any way that could endanger the good name or reputation of the College. Thus the Head of Department who is essentially responsible for the WWW systems in the department should ensure that the material mounted does not risk criminal prosecution or civil legal action, or that material, even if legal, is not inappropriate for publication by the College. Particular care should be taken over any material which breaches copyright or is potentially libellous, and in such cases the Marketing Division should be consulted first.

Failure to observe this code of practice by either students or staff will be considered a serious matter by the College. Where College regulations are breached the College will invoke the appropriate disciplinary procedures. For students this could involve fines, suspension of access to computing facilities or, in extreme cases, rustication or being sent down.

In either case, upon authorisation of the Rector or Deputy-Rector, the Centre will isolate the WWW system to its local subnet until the offending material is removed. Any breaches of the criminal or civil law are beyond the remit of the College. If the DPP decides upon a criminal prosecution this will be a matter for the department/individual. Similar considerations apply to any civil law case.

## Further Comments

The full document is available for viewing on the Imperial College WWW server, with an URL of <http://www.ic.ac.uk/iss/wwwcode.html>

In addition to the above statement it contains some general guidance on the legal requirements that should be borne in mind by those responsible for a WWW server.

Initial response to the Code has been very good. Departments have tried to observe the guidelines without being unduly bureaucratic, and apart from the one or two enthusiasts who were not aware of the Code, and mounted some rather unsuitable material (in terms of the image of their department) it has worked well.

However, in an environment that is developing as rapidly as WWW, new issues are going to arise for which the Code is not particularly well formulated. One such issue that has arisen in the last few months concerns Personal Home Pages (PHP). These are pages on a WWW server which are devoted to information about an individual. They seem to have originated in the USA, and can contain information varying from purely professional information e.g. research interests, CV, lists of publications, to very personal details of the individual's lifestyle.

The issue of whether these should be permitted on Imperial College WWW servers arose after the Code of Practice had been issued. While the Code covered the issue in a very general way, a situation arose where some departments banned them, while others allowed them with various conditions attached. The situation also arose where UGs started to request them. The Centre felt that a situation was arising where individual users in this particular instance were being treated very unequally.

The Centre consulted with the Departmental Computer Representatives Committee (the main channel of consultation between the Centre and the departments) on the issue and found widely varying views. After consultation with the Chairman of the College Information Systems Strategy Committee (ISSC) the Centre wrote to Heads of Departments and asked them not to agree to any further PHPs until the issue could be discussed by the ISSC. The Centre then wrote a brief paper summarising the views expressed by the Departmental Computer Representatives, listed the possible options from a complete ban to a situation with no limitations, and provided some examples of PHPs which varied from the sober to the frivolous!

If PHP were to be permitted then the issues were concerned with who should be allowed to mount a PHP e.g. academic staff, RAs, PGs, UGs, and what sort of material should be mounted. A particular issue on which there were strong views concerned RAs and final year PGs who were very keen to put up their CVs and publications lists with the information that they were in the market for a job!

The ISSC met in January 1995 and agreed a policy. This policy is shown below.

## Personal Home Pages (PHPs) on WWW Servers

At the ISS meeting on January 19th. the issue of the use of PHPs on Imperial College campus WWW servers was considered. Following from this meeting, the College has set the following guidelines for the use of PHPs on Imperial College WWW servers.

1. Permanent academic staff, RAs and Research PGs are permitted, at the discretion of the Head of Department, to mount PHPs on a departmental WWW server, subject to adequate disc space being available.
2. Such PHPs should be confined to professional information only i.e. research interests, publication list, CV, and in the case of RAs and Research PGs, whether they are in the market for a job. Hypertext links within all such PHPs should be internal only i.e. restricted to material within the

PHP. The Head of Department retains the right to veto any material considered not to be in the best interest of the department or College.

3. Except for PHPs related to 'official' College positions e.g. Departmental UG Tutor, the use of pictures should be strongly discouraged as their transmission loads the network.
4. In general terms UGs are not permitted to have PHPs. Taught PGs and final year UGs represent a special case. The Centre for Computing Services will investigate, with the Careers Service, the possibility of creating specific 'Available for Employment' WWW pages for taught course PGs and final year UGs. If this is practicable, entries will consist of a 'standard format' CV provided to the Careers Service on floppy disc in a form that could be simply processed for mounting on the WWW pages.

First reactions from the campus are just beginning to come in. They largely relate to two points. One is relatively trivial, and concerns the use of 'thumb nail' size images. The College is most likely to modify the guidelines to request a limit on the size of any images used. The second is more significant. It concerns the issue of links being confined to within the PHP document. There are two schools of thought. One suggests that the limit should be to links 'on campus' so that research papers mounted on research group WWW servers can be linked to the PHP. This is not unreasonable except that once links are external to the PHP they cannot be limited to just 'on campus'. The second suggestion is that links to anywhere in the world should be allowed, since links can then be made to research papers at other sites.

The original constraint on links was made because the College had an example of a PHP where links were made world wide to material that had no conceivable College purpose. However the basic issue seems to be that while the suggestion being made, which is to create a sort of electronic journal, is a perfectly sensible one, its place is not a PHP but possibly on a research group WWW server. This issue will be debated in the College, and may result in the guidelines being changed on the links issue.

---

## **OTHER GUIDELINES**

---

Other sites are developing guidelines and the online version of this document will point to these. The guidelines from the University of Edinburgh were available at the workshop. These can be found through:

<http://www.ed.ac.uk/>

If any site has guidelines which they feel may be useful to add as a pointer from the online version of this document, please send an email giving the details to Sue Cunningham, the Multimedia Support Officer:

[sue.cunningham@mcc.ac.uk](mailto:sue.cunningham@mcc.ac.uk)

---

# LEGAL ISSUES OF WWW & ELECTRONIC PUBLISHING

## ANDREW CHARLESWORTH

---

It should be noted that this is not a truly exhaustive list of the issues that may need to be taken into account. In electronic publishing, as in traditional forms of publishing, there are a myriad pieces of legislation which must be taken into account, including those concerning liability for content of publications, covering such topics as defamation, obscenity, blasphemy, and provisions with regard to sex discrimination and advertising standards. Defamation is discussed in this piece and obscenity touched upon, but space and time constraints prevent a full examination of the others. However, as the magazine *Gay News* discovered in the 1970's, there is still life left in the blasphemy laws, and Salman Rushdie and his publishers found out that, even though UK blasphemy law does not provide for Islamic sensibilities, publishing still holds many risks. The areas covered here are those in which there have been either ongoing or recent developments which have been of relevance to those working with the WWW. Readers in Scotland should be aware that certain of the topics discussed here are discussed only in relation to the law of England and Wales. This document should not be considered as a substitute for more comprehensive legal advice.

### Intellectual Property

In any discussion of electronic publishing, it is inevitable that the issue of the role of intellectual property in this sphere will be raised, for it is probably the single largest problem facing electronic publishers, and, of course, those who actually hold, and wish to protect, those intellectual property rights.

### Copyright and Associated Rights

#### What may be protected?

- *Literary Works*  
Literary works for this purpose include not only novels, poetry and non-fiction books but also all sorts of other written works which are original. Their literary merit is unimportant. This means that letters, memoranda, directories, e-mail messages and WWW pages may be protected. It should also be remembered that computer programs and code are protected as literary works.
- *Dramatic Works*  
Dramatic works for this purpose include plays and instructions for dance or mime. There must be some spoken words or described actions to perform to distinguish a dramatic work from a literary work. However the fact that a play does not contain any dialogue does not prevent its qualification as a dramatic work
- *Artistic Works*  
Artistic works for this purpose include graphic works, photographs, sculptures, collages, maps, charts and plans. These are protected regardless of artistic merit. However, works of architecture and of artistic craftsmanship require artistic quality in the work to qualify for protection
- *Musical Works*  
Musical works for this purpose include musical scores including any annotations and directions. Lyrics, however, are not, as they are protected as literary works.
- *Sound Recordings*  
This category covers every type of sound recording on any type of medium from which sounds can be reproduced.
- *Films*

'Films' in this context covers any medium from which a moving image may be reproduced. As the definition of film under the 1956 Act was similar video recordings appear to be covered from the time of their development.

- *Broadcasts*  
'Broadcasts' includes any transmission by wireless telegraphy which is capable of lawfully being received by members of the public. This clearly therefore includes satellite transmissions.
- *Cable Programmes*  
These are defined as transmissions carried as services via cable, including on-line services.
- *Published Edition*  
There is copyright in the typography and layout of a literary, dramatical and musical work.
- *Performers' Rights*  
While these rights are not technically copyrights, they provide protection to performers and persons who hold recording rights in a performance. These rights are included in the CDPA 1988.

### **Who owns the rights?**

In theory, the original owner of copyright in a given work is the person who created it. There are however, exceptions; in many cases, works created in the course of employment will be owned by the employer - universities are unusual in that much of the literary copyright remains with academic authors. Ownership of copyright in a work can change hands after its initial creation, and like any property, can be sold or assigned and may be passed on in a will.

Subject to the provisos above:

- *Literary Work*  
is owned by the author.
- *Sound Recordings and Films*  
are owned by the person by whom the arrangements necessary for the making for the recording or film are undertaken.
- *Broadcasts*  
are owned by the person making the broadcast.
- *Cable Programmes*  
are owned by the person providing the cable programme service in which the programme is included.
- *Published Editions*  
published editions of literary, dramatical and musical works are owned by the publisher.

As noted above this is subject to the proviso that works created prior to 1988 may be covered by different regimes. Thus, under the 1956 Act, copyright in a photograph belonged to the person who owned the negative film, unless the photograph was taken under commission.

### **How long do the rights last ?**

All works will eventually emerge from copyright protection. However, different types of works have different lengths (or terms) of copyright protection. Also despite the role played by international agreements such as the Berne Convention, different countries apply different lengths of copyright protection to works.

The CDPA 1988 made changes to the length of protection for various in the UK works but, as it does not apply retrospectively, it still remains necessary to be aware of the relevant provisions in the 1956 Act and the 1911 Act. Equally, there are various variations and exceptions, a particularly irksome one being Crown Copyright, which can be longer than normal copyright term.

### **A rough and ready guide to terms of copyright:**

- *Literary, Dramatic and Musical Works*  
The author's life and 50 years after his/her death.
- *Works of Joint Authorship*  
50 years from death of last author to die.
- *Anonymous Works*  
50 years from first publication.
- *Artistic Works*  
The author's life and 50 years after his/her death.



- *Computer Generated Works*  
50 years from first creation
- *Photographs*  
prior to 1988 - 50 years from when taken  
after 1989 - 50 years from the year of release.
- *Sound Recordings*  
prior to 1988 - 50 years from when made  
after 1989 - 50 years from the sound of release.
- *Films*  
prior to 1988 - 50 years from when made (reportage)  
prior to 1988 - 50 years from the death of the author of film script  
after 1989 - 50 years from the year of release.
- *Broadcasts and cable programme services*  
50 years from when broadcast first made or programme included in a cable service.
- *Published Editions*  
25 years from first publication of that edition.

### **Directive 93/98/EEC**

From July 1995, the term of copyright in the EC will be harmonised, extending the basic term from author's life + 50 years to author's life + 70 years. It appears at present that this will have retrospective effect, which will mean that some material which has fallen out of copyright in the UK will be, as it were, recopyrighted. It is unclear as to how this may effect those who have acted in reliance on the material being out of copyright.

### **What does this mean for multimedia/WWW publishers ?**

A multimedia or WWW publication may include some or all of the following copyrightable components:

- Literary elements - protected as literary works
- Dramatic elements - protected as dramatic works
- Musical elements - protected as musical works
- Artistic work (graphics, photographs, drawings and models) - protected as artistic works
- Moving images - protected in the same way as films
- Sound recordings - protected as sound recordings
- Typographical arrangements of published editions of literary, dramatic or musical work
- Computer program - protected as a literary work
- Choreographic routine - protected as a literary work

### **Worst case scenario**

A WWW home page dedicated to a biography of Margot Fontayn, including mpeg excerpts from a film of Swan Lake, photographs of her performances and ballet music by several composers, living and dead, including pictures of the musical scores.

### **What rights does the multimedia producer/WWW publisher have to obtain under copyright law?**

- the right to copy the work
- the right to issue copies of that work to the public and to let them copy it (limited)
- the right to adapt the work
- the right to perform the work in public
- the right to broadcast that work

## Where things start to get complicated

### *Performers' consents*

Where a multimedia product includes recordings of performances, various consents will be required from differing categories of performers. At UK law (but not US law) performers rights exist in:

- Musical works
- Dramatic performances
- Readings or recitations of literary works
- Performances of variety Acts

Obtaining this consent may not be easy especially if dealing with older material. Problems include:

- Identifying performers (includes not just featured performers, but also backup musicians & singers chorus girls and crowd extras - note rock concert exclusions).
- Performer may be dead - right of consent is transmittable to heirs

In certain circumstances, the Copyright Tribunal under CDPA 1988 can give consent on behalf of untraceable or unreasonable rights holders - but only for UK.

### *Rights of employees of multimedia producers and independent contractors*

While copyright in literary, dramatic, musical & artistic works belong to the employer, and independent contractors can be required to assign copyright, performance rights are more complicated - in the case of the employee they do not automatically transfer to the employer, consent should therefore be obtained contractually. Note that consent applies only to 'first fixing' although this problem can be overcome by copyright.

### *Copyrights in Films & Sound recordings*

These belong to the person or persons who undertook the arrangements which resulted in the film being made. Multimedia producers should therefore ensure that it is clear in contracts of employment that such arrangements are undertaken on the understanding that they are on behalf of the producers.

### *Outside Hours Work by Employees*

Where employees undertake work outside their strict working hours, be that making of arrangements under C above or creating works under B above, these may not be covered by the employers' right to the copyright. Contracts should thus ensure that this eventuality is catered for.

### *Personality Rights (US & other jurisdictions - not UK)*

The right of a person to prevent their name, likeness or biography to be used without their consent. therefore individuals mentioned in credits and packaging should give consent as should those featured in a multimedia product.

### *Music in Multimedia*

Synchronisation licences - the consent of the copyright owner to use music and lyrics in synchronisation with or timed relation to moving images.

Mechanical royalties - payment above and beyond that for the synchronisation licence.

Commissioned music by the multimedia producer can avoid both these problems, and may lead to potentially lucrative spin offs - Super Mario music.

### *Trademarks*

Many companies throughout the world have symbols and logos which they have protected by registering them as trademarks (1). For example, Windows, MS and MS-DOS are registered trademarks of the Microsoft Corporation. Companies register such symbols & logos in order to prevent other companies from using them to 'pass off' their product or service as a product or service of the original company. As a result, many firms are touchy about individuals using trademarked logos on WWW pages etc. unless their explicit permission is given, and the logo is stated to be a trademark of the holder.

### *Patent*

Patent did not appear to be a major influence in the area of the WWW, until the case of Unisys's patent on the mathematical algorithm which underlies the .gif picture compression format widely used by WWW users. It is clear that the US Patent Office's decisions with regards to the granting of computer software patents (often for algorithms for which there would appear to be 'prior art'), and the use of 'submarine' patents, may well have a significant effect in the way in which software in general, and WWW software, which has generally been available as freeware, in particular, develops.

## **Libel**

It is clear that the law of defamation will be applied to the various forms of Internet usage, including the creation and management of WWW sites, and will inevitably affect the way in which individuals and institutions use it. There are, however, a number of unresolved questions as to its application in this area. These include, who may be liable ?

While definitive answers are difficult to come by, the answer to this in terms of e-mail may well include:

- the poster of the defamatory statement to a Bulletin Board AND the System Operator of the Bulletin Board;
- the poster of the defamatory statement to a moderated mailing list AND the moderator, where the moderator reads all the messages for content;

with regard to the WWW:

- the creator of a defamatory WWW page AND the owner of the WWW server on which it is based;
- potentially, under certain circumstances, the creator of a link to a defamatory WWW page where that link will have the effect of spreading the defamation..

In the US it seems that the Courts have been inclined to hold the sysops of bulletin boards liable for material held on their boards, be it pornography, illegally copied software or other copyrighted material.<sup>1</sup> The debate in the US thus also touches on First Amendment issues and whether bulletin boards etc. should be granted the same type of privileged status as newspapers with regard to publication of allegations about public figures. In the UK, the issues raised by electronic defamation inevitably have a less constitutional bent. As yet there seem to be no WWW based cases or ongoing actions.

---

<sup>1</sup> A federal district court in Florida has held that a BBS operator is liable for infringing Playboy's copyright distribution and display rights by making available Playboy pictures in machine readable format. The interesting part is that the operator alleged that a subscriber had uploaded the files without the operator's knowledge, and the files had been removed as soon as the operator was aware of their presence. See *Playboy Enterprises, Inc. v. Frena*, No. 93- 489-Civ-J-20 (D.C. M. Fla. 12/9/93).

What then are the matters which need to be considered in regard to an action for libel?:

- “Libel consists of a defamatory statement or representation in permanent form ... Any thing temporary and audible only is slander. Statements in books, articles, newspapers and letters are libels.”<sup>2</sup> It might be possible to argue therefore that electronic communications/publishing are not permanent and thus cannot be libellous. However, as electronic communications/publishing are often downloaded as hardcopy for dissemination and reading, I suspect that this line of argument would be unlikely to succeed, despite the problems that the courts have often had in applying the law to modern technology.
- Is the allegation complained of defamatory as opposed to vituperative/abusive ? Those involved in flame wars on the Internet are frequently abusive about their opposite numbers - however, venting one’s feelings, even if they injure the other person’s self esteem, is not sufficient for libel - defamation is only made out where the plaintiff is held in lower esteem by others as a result.
- Does the defamatory statement refer to the plaintiff ? That is, would the ordinary man in the process of scanning his newspaper, e-mail, bulletin board or WWW page be led to believe that the plaintiff was being referred to.
- Has the defamatory statement been made known to others - has it been published ? Communication to the party named in the defamatory statement is not publication, as libel is concerned with how third parties view the plaintiff. It has been said that “the question of publication of a libel contained in a letter will depend on the state of the defendant’s knowledge, either proved or inferred, of the conditions likely to prevail in the place to which the libel is destined”.<sup>3</sup> In the case of an electronic mailing list the individual posting the communication will be well aware of the fact that it will be widely disseminated by the listserver, and still further by automated forwarding devices and the actions of others. In the case of a BBS, the individual posting the communication will similarly be aware that it will be available to anyone who dials up that BBS. It would seem that WWW pages might well have the widest potential dissemination rate of all. Thus the original poster cannot claim involuntary republication, for he is aware that this will occur.
- An important point to note is that “every repetition is a fresh publication giving rise to a fresh cause of action against each successive publisher. Thus not only the author of an article, but the editor, printer and publisher are also liable. Moreover, even mechanical distributors such as bookstalls, could be liable”<sup>4</sup> This is subject to the defence of unintentional defamation. Thus it could be argued that a BBS sysop could potentially be held liable for postings on his board, and that a moderator of a moderated e-mail list (where the messages are read by the moderator before posting to those on the list) almost certainly would be. Similarly the owner of a WWW server would appear to be caught by this, and even possibly an individual providing a link to the WWW page in question.

It seems likely that the courts in the UK would be inclined to accept that it is possible to libel individuals on BBS, mailing lists or WWW pages by posting untruthful and damaging statements about them in such fora. Thus a UK citizen posting untruthful and damaging statements about another UK citizen, on a list, BBS or WWW page where such a message would be read by others in the UK, and where this would be damaging to the plaintiff’s good name or reputation, would almost certainly be liable to an action for libel. The issue of a UK citizen libelling a US citizen is perhaps less clear as this might be considered by the courts to be out of their jurisdiction. However in the recent UK case *Godfrey v. Hallam-Baker* (ongoing) where the claim is for damages for libel or alternatively slander regarding seven Usenet messages posted in 1993, the defendant apparently works at CERN.<sup>5</sup>

As all litigation which has so far taken place appears to have concerned libellous statements via e-mail, it is to this that we have to turn to obtain a view of the way in which courts have reacted to defamation in the electronic forum.

---

<sup>2</sup> Dias & Markesinis - *Tort Law*

<sup>3</sup> *Theaker v Richardson* (1962).

<sup>4</sup> Dias & Markesinis *op.cit*

<sup>5</sup> *The Independent* 22 August 1994 at 22

The Australian case of *Rindos v Hardwick* (31 March 1994, Unreported)<sup>6</sup> seems to make it quite clear that the Courts in that country are willing to accept that an individual can be libelled via the medium of a bulletin board or mailing list. The case concerned, in part, an entry placed on the DIALx science anthropology computer bulletin board by the defendant. The plaintiff was an academic at the University of Western Australia who was sacked on the ground of insufficient productivity. This action drew protests from academics at a number of international archaeological institutions, including one to the bulletin board from US anthropologist Hugh Jarvis. This in turn was replied to by Gilbert Hardwick from a computer in Derby, Western Australia. This entry imputed that the plaintiff had engaged in sexual misconduct with a local boy, and that his academic reputation was not based on appropriate academic research but "on his ability to berate and bully all and sundry" from which the inference could be drawn that these had some bearing on the plaintiff's sacking.

The bulletin board in question was mainly used by academics and students, and according to the Court was accessible by upwards of 23 000 people worldwide. It noted that items placed on the board could also be printed out, and distributed in hard copy. The defendant made no effort to justify his comments, and did not defend his action in court. In his judgement Ipp J found that the remarks were clearly defamatory and had been widely published, and that the plaintiff had thus suffered serious harm to his reputation as a result of them, and awarded him \$40 000 dollars in damages. Dr Rindos' lawyer, Robert Castiglione said " Computer users who use these worldwide bulletin-boards should be aware that they could be exposing themselves to defamation actions ... It's an informal system where people say quite personal things, but making allegations of paedophilia and bullying is going too far."<sup>7</sup>

The case involved two Australian based scientists - as yet it is unclear as to what would happen if the issue had involved an Australian scientist, and say, a Canadian scientist. At this stage, the issue of jurisdiction comes into play, and this has several levels. We accept that it is clear that an Australian resident libelling an Australian resident on an internationally read bulletin board/ mailing list will be liable in the Australian courts, however would the Australian courts accept jurisdiction over:

- a libellous statement made by a national of another country against an Australian national on a bulletin board which would be read internationally, including by other Australians ?
- a libellous statement made by an Australian national against a national of another country on a bulletin board which would be read internationally, including by other Australians, and nationals of the libelled party's country ?
- a libellous statement made by a national of another country such as the US, against another US citizen, which by virtue of the Bulletin board was widely read throughout Australia ?

In short, because the standard for defamation would appear easier to meet in Australia than in the US and the defences available are fewer, would it be possible for a plaintiff pick his forum and his law because of the international nature of the Internet?

The nature of the electronic dissemination itself might be a factor - i.e. an individual sends a libellous communication to a listserver in Canada, intending for all the subscribers of the list (including some in Australia) to receive the column, as opposed to his posting his libellous communication on a BBS in Canada, and someone in Australia dialling up and reading it. In the first case the individual is deliberately placing the material into a foreign jurisdiction, in the second that intent would appear to be lacking. The very nature of the WWW however, would appear to place all such defamatory material placed on WWW servers into the first category.

There is of course a further twist to this with regard to the WWW. Take, for instance, the example of an individual who, while not placing defamatory materials on his WWW home page or server,

---

<sup>6</sup> See 'Computer libel wins academic \$40 000' M.Lang, *The West Australian*, 2 April 1994

<sup>7</sup> Ibid.

provides links to such material with the aim of directing others to them. Is the individual concerned, and perhaps his employer, thereby 'publishing' the material as regards any possible libel action?

The recent US lawsuit involving a journalist running his own on-line newsletter demonstrates possible problems for electronic publishing. The journalist, Brock Meeks, a resident of the Washington, D.C. area and an employee of a communications trade journal, created the on-line news service, Cyberwire Dispatch, to comment on developments in the IT field.

In one of his articles, he expressed his personal disapproval of the business activities of a particular company, Suarez Corporation Industries (SCI) concerning their activities in the direct mailing business.<sup>8</sup> SCI objected to the article and filed a defamation lawsuit claiming Meeks made defamatory remarks and sought to disparage its products. The owner claimed that the Dispatch article lost him business and he thus sought compensatory and punitive damages and demanded an injunction to block Brock from writing further about SCI or its owner.<sup>9</sup> In the event, the issue was settled out of court, with Meeks paying SCI's court filing fee, some \$64, and promising to contact SCI before publishing any further articles about it.

The case would, however, appear to demonstrate an interesting problem with the law of defamation. The cost of setting up a bulletin board, or Web server is not particularly expensive, however the cost of defending a legal action is likely to be very high. Concern has thus been expressed that as the number of libel suits internationally appears to be on the increase, the law of libel may well be used increasingly to stifle what many Internet users have long considered to be their right to absolute free speech on the networks, and that this will in turn have a "chilling" effect upon the willingness of individuals to carry out certain forms of electronic publishing, and may also make academic institutions less willing to allow staff and students a free hand in this area.

It is clear from the above that this area throws up a number of questions which the courts in the UK do not yet appear to have addressed.

- Can BBS sysops and WWW server owners be held liable when untruthful and damaging statements about individuals are made on systems under their control, even without their knowledge - where does the buck stop?
- If they can be held liable, are they then obliged to monitor every communication or web page - something that is probably next to impossible to do?
- Further, should they then go on to censor those communications or web pages which they suspect to contain untruthful and damaging statements about individuals?
- In allowing untruthful and damaging statements about individuals to be made on systems under their control, are they leaving themselves open to multi-jurisdictional liability?

These are topical issues in the US at present and I would suggest should be considered seriously in the UK as well. It appears that most UK Universities approach this problem via their rules for use of their computer systems, although others have specific rules about the setting up of individual as opposed to institutional WWW home pages. These type of rules and regulations may provide a defence should a user of the institutions machines, use them to disseminate defamatory material, subject to the actual response of the institution upon the defamation being brought to the attention of the relevant authorities.

---

<sup>8</sup> Notably that state and federal enforcement agencies had brought actions against SCI as result of their direct mailing practices.

<sup>9</sup> It should be noted that the suit against Meeks was filed in Ohio.

## Crime

The scale of computer crime, particularly as it affects the WWW, is extremely difficult to gauge. It would appear from surveys in the area of computer crime/misuse<sup>10</sup> that most crime/misuse is committed by individuals against their employers in a business environment. The most obvious (at least to the journalists) crime that might be carried out via the Web is the distribution/downloading of computer pornography. This may be covered by a number of provisions covered including the Telecommunications Act 1984<sup>11</sup> the Obscene Publications Act 1959, and with regard to child pornography, new legislation in the form of the The Criminal Justice and Public Order Act 1994, which, in amending other legislation including the Protection of Children Act 1978, is aimed specifically at computer generated and distributed pornography. That having been said, even when one takes the media hype into account<sup>12</sup> in the period 1991-1993 of the 976 obscenity cases handled by the Crown prosecution Service, only 11 involved computer pornography and only 7 of those went to court.<sup>13</sup>

In general, most WWW sites, particularly those based at academic institutions are keen to avoid any problems with hard or soft core pornography, and a great deal of control can be exercised by peer pressure from other institutions without the aid of the law, even where the law of the country involved does not forbid such material. For instance, WWW servers at the University of Delft (hard & soft core) and the Conservatoire National des Arts et Metiers (CNAM) (soft core), which carried such pictures downloaded by automatic newsfeed from Usenet groups, have both been forced to remove the offending material due to pressure from their governing bodies.<sup>14</sup> The US may be more problematic in this regard due to the First Amendment issues surrounding the area.<sup>15</sup>

---

<sup>10</sup> For example *Opportunity Makes a Thief: An Analysis of Computer Abuse*, the fifth triannual report of the Audit Commission on the extent of computer abuse and fraud in the UK (1994, HMSO Publications).

<sup>11</sup> s43 Improper use of public telecommunication system  
(1) A person who--  
(a) sends, by means of a public telecommunication system, a message or other matter that is grossly offensive or of an indecent, obscene or menacing character; or  
(b) sends by those means, for the purpose of causing annoyance, inconvenience or needless anxiety to another, a message that he knows to be false or persistently makes use for that purpose of a public telecommunication system, shall be guilty of an offence and liable on summary conviction to a fine not exceeding level 3 on the standard scale. [...]

<sup>12</sup> And there seems to be no end to the number of articles like the one run by *The Guardian*, 24 August 1994 somewhat unoriginally entitled 'Computer going down' which noted that a University of Wales computer was put out of action for two days due to an overload caused by a student downloading pornography from the US.

<sup>13</sup> 'Industry focuses on cleaning up its act' *The Guardian* 27 September 1994 at 8.

<sup>14</sup> And, it must be said, apparently due to massive overloads on the machines concerned as individuals attempted to access the pictures.

<sup>15</sup> Although there have been interesting developments with regard to 'jurisdiction hopping' by law enforcement agencies, see a recent US case reported las year (28 July 1994) on the cyberia-1 mailing list

"Jury Convicts Couple in Computer-Porn Trial" MEMPHIS, Tenn. (AP) -- A federal jury convicted a California couple today of transmitting obscene pictures over a computer bulletin board. The case has raised questions, in this age of international computer networks, about a 1973 Supreme Court ruling that defines obscenity by local community standards. Prosecutor Dan Newsom, an assistant U.S. attorney, said the trial was the first he knows of for computer bulletin board operators charged under federal law with transmitting pornography featuring sex by adults. Robert and Carleen Thomas, both 38, of Milpitas, Calif., were convicted of transmitting sexually obscene pictures through interstate phone lines via their members-only Amateur Action Bulletin Baord System. The Thomases were convicted on 11 criminal counts, each carrying maximum sentences of five years in prison and \$250,000 in fines. Thomas was acquitted on a charge of accepting child pornography

In terms of UK institutions, the solution is similar to that which has been outlined for libel above, in that institutions may make up their own rules to govern the content of WWW home pages, and can, via this route, explicitly forbid certain material being published on pain of expulsion from the system and perhaps even the institution.

It has also been suggested that publishing material which might be used in order to breach computer security or to facilitate unauthorised entry into computer systems may be caught by the Computer Misuse Act 1990, in terms of conspiracy to commit an offence. This does not yet appear to have been successfully tested in court.

## **Recent criminal law provisions**

### **The Criminal Justice and Public Order Act 1994**

#### **Child Pornography - s84-87**

**Amends:** Protection of Children Act 1978, Criminal Justice Act 1988, Civil Government (Scotland) Act 1982, Protection of Children (Northern Ireland) Order 1978, Criminal Justice (Northern Ireland) Order 1988, PACE 1984, Obscene Publications Act 1959.

#### **Procuring disclosure of, and selling, computer-held personal information - s161**

**Amends:** Data Protection Act 1984

#### **Access to computer material by constables and other enforcement officers - s162**

**Amends:** Computer Misuse Act 1990

## **Data Protection**

If a WWW home page or server contains, or has links to files within the same organisation which contain, personal details of named individuals, it would be wise for the owner to check the provisions of the Data Protection Act 1984 with regard to whether or not he should be registered with the Data Protection Registrar. It is also worth noting that a more rigorous regime has been proposed by the European Community in a draft Directive, which is currently the subject of negotiations regarding a common position by the Member States. If negotiations complete within the next couple of months, it has been suggested that this will result in new UK legislation by 1998.

The Data Protection Act 1984 applies to automatically processed information, but not to manually held information, which is contained in files or other paper records. Only information which relates to living individuals is covered.

---

mailed to him by an undercover postal inspector. The Thomases refused to comment after the verdict. They remain free on \$20,000 bond to await sentencing, for which no date was set. Defense lawyer Richard Williams said his clients will appeal, arguing the jury was wrongly instructed on how to apply the Supreme Court's standard on obscenity. The trial raised questions of how to apply First Amendment free-speech protections to "cyberspace," the emerging community of millions of Americans who use computers and modems to share pictures and words on every imaginable topic. Williams argued unsuccessfully before trial that prosecutors sought out a city for the trial where a conservative jury might be found. "This case would never have gone to trial in California," he said. During the weeklong trial jurors were shown photographs carried over the Thomases bulletin board featuring scenes of bestiality and other sexual fetishes. Their conviction also covers videotapes they sent to Memphis via United Parcel Service. The videotapes were advertised over the bulletin board.



The Act gives individuals the right to find out about what information is recorded on computer relating to them. They also have the right to challenge the information and, if appropriate, claim compensation in some circumstances.

Personal data is defined by the Act as data which consists of information which relates to a living individual, who can be identified from that information alone or in conjunction with other information held by the data user. Some kinds of data are exempt from the Act, do not need to be registered, and the individual concerned does not have any right of access to the information.

Exempt data includes:

- information that the law requires to be made public
- information maintained to safeguard national security
- personal data held for domestic or recreational purposes
- information held for payroll, pensions, and accounts purposes, subject to certain conditions
- information relating to unincorporated members clubs
- mailing lists, subject to certain conditions

### **The Data Protection Principles**

- the information to be contained in personal data must be obtained fairly and lawfully, and the data must be processed fairly and lawfully
- personal data must be held only for one or more specified and lawful purposes
- personal data held for any purpose shall not be used or disclosed in a manner which is incompatible with the purpose
- personal data held for any purpose must be adequate, relevant and not excessive in relation to the purpose
- personal data must be accurate and, where necessary, kept up to date
- personal data held for any purpose must not be kept for longer than is necessary for that purpose
- an individual must be entitled at reasonable intervals, and without undue delay or expense, to be informed by any data user whether that user holds personal data relating to the individual. The individual is also entitled to have access to any such data and, where appropriate, to have the data corrected or erased

### **The Internationalisation of the Issues**

While this paper has concentrated mainly upon UK legal issues, it will be clear that the international nature of the WWW means that we may well need in the future to be cognisant of the laws of many more countries, and how these may affect us.

### **Further reading**

Cyberspace and the Law: Your Rights and Duties in the On-line World (US law)  
Edward A. Cavazos & Gavino Morin  
The MIT Press (1994), £17.95, 215pp, ISBN 0-262-53123-2

Publishing and Multimedia Law (UK law)  
Michael Henry  
Butterworths (1994), £75.00, 820pp, ISBN 0 406 03768 X

Information on the role of the Data Protection Registrar may be obtained from:

Office of the Data Protection Registrar  
Wycliffe House  
Water Lane  
Wilmslow  
Cheshire SK9 5AF  
Tel: 0625 535777

## Book Reviews

*The following two books were reviewed by Andrew Charlesworth, Lecturer in IT law and Director, Information Law and Technology Unit, University of Hull Law School. The reviews were conducted for the CTI Centre for Law's Law Technology Journal (1995).*

**Cyberspace and the Law: Your Rights and Duties in the On-line World**  
**Edward A. Cavazos & Gavino Morin**  
**The MIT Press (1994), £17.95, 215pp, ISBN 0-262-53123-2**

This book, aimed primarily at a non-legally trained American audience, attempts to provide a much needed guide to the legal problems and pitfalls which await the unwary who provide and use on-line services ranging from electronic mail through to bulletin boards. As such, it does not provide a great deal of analysis of the legal issues, preferring instead to concentrate upon the practical implications of breaches of various aspects of the law. While this lack of analysis is disappointing to the more enquiring reader, it does accord with the authors' stated aim of providing a concise guide to the law in this area.

After defining its terms of reference, notably the term "cyberspace" (with the usual obligatory nod towards William Gibson), the book turns to the legal issues, beginning with that of privacy. This is an issue that has long exercised American writers rather more than it does their UK counterparts. In the UK, the right to privacy, even defined in the broadest of terms, is a very nebulous concept, and the degree of protection in the US, hedged around though it may be by numerous exceptions, makes impressive comparative reading. This chapter also touches upon the topical issues of cryptography and the Clipper Chip debate, and while for the purpose of this book, the coverage of these topics is probably adequate, it would have been interesting and informative to have more analysis of US government action on these fronts, as the implications of US government moves to control encryption have important international implications. Indeed, the speed of developments in the area of privacy is such that the events described in this book are already dated, with the struggle over the right to privacy between state and individual having already moved to other battlefields.

The business transactions chapter provides a brief explanation of the law of contract and then applies this model to 'cyberspace'. This chapter specifically excludes Electronic Data Interchange contracts, and in doing so avoids discussing some of the more interesting and contentious legal developments which are arising in that area. Indeed, the chapter is fairly parochial in nature, showing no significant interest in transborder transactions. This is an aspect of 'cyberspace' which, with the increasing use of credit cards and the potential for 'digicash' (particularly with relation to transactions via the World Wide Web), must surely be a growth area. The main areas tackled are the issue of the postbox rule in contract, the problem of validating contracts which are in dispute, and the difficulties of defining 'in writing' and 'signature' in electronic transactions. Overall, this chapter suffers from the problem which afflicts other more legally oriented texts in this area which is that it spends most of its time explaining the law of contract, and relatively little time on the specific issues relating to 'cyberspace.' Its underlying theme, particularly with regard to the section on credit cards appears to be simply 'let's be careful out there', a laudable enough sentiment, but one which does not reflect any great, or indeed novel, analysis.

The intellectual property chapter, which refers almost exclusively to copyright, takes an equally 'nuts and bolts' approach to the law. This results in a fair summary of the law, but fails to give more than

a passing mention to the fact that the application of intellectual property law in 'cyberspace' might in fact be a highly contentious issue, whether one is dealing with computer program copyright or copyright in Playboy pictures. Given the suggestions for changes to IPR put forward in late 1994 by the US Information Policy Committee of the Information Infrastructure Task Force's Working Group on Intellectual Property Rights, and the reactions to those suggestions from business and academic commentators, it seems that this will remain a controversial area for some time to come. The issue of how what was essentially a printers monopoly in the 17th century has evolved into something that threatens the usefulness of cyberspace as an information source is perhaps a question that too few people have taken the time to consider.

However the final three chapters, dealing with free speech, adult materials and cybercrime respectively, provide rather more insight into their respective topics. This may well be because 'cyberspace' usage has so far tested the boundaries of the law in these areas rather more rigorously, and thus the authors are able to consider real-life examples instead of hypotheticals with regard to them. It may also be that the interests of the authors' are related more to such issues, rather than to those of contract and intellectual property, which would thus explain the workman like approach taken to the latter.

The appendices, which account for over one third of the book, contain primarily US statutory material, including some extremely useful lists of the state statutes concerning child pornography and computer crime.

In conclusion, it would be fair to say that this book does represent a welcome venture into the area we might call 'cyberlaw'. While it has been written for the layman, the law appears both concise and accurate, and it is an ideal introductory text for any examination of the interaction between the American legal system and the 'cyberspace' explorer. It is true that it does little more than state the relevant law, and offers no real challenge to conventional wisdom with regard to developments, potential or actual, in that law, but it does provide the reader with the necessary background material to explore and evaluate those developments. All in all, it is eminently accessible, being tightly written and thus easy to read and comprehend, and as such it would seem to be a useful and informative comparative work for UK law students interested in the area of IT and computer law.

**Publishing and Multimedia Law — Michael Henry**  
**Butterworths (1994), £75.00, 820pp, ISBN 0 406 03768 X**

This book is primarily aimed at the practitioner and publishing professionals markets, as may be deduced from its hefty asking price. It contains both a general guide to the law in the related areas of publishing and multimedia development, and a number of precedent documents. These may be used to construct user-specific documents for transactions such as assignment of copyright and permission for photograph use. This is however, subject to the author's caveat that correctly drafting and amending documents in this area may be a complex task and may require professional help. In order to make these documents more readily available, they are also supplied on PC disk. There is no indication as to whether similar provision for other computer systems, notably Macintosh, is readily available.

The main text covers a wide array of topics divided into twenty sections. Of these the first eight deal with various aspects of intellectual property rights, being primarily concerned with copyright and other associated rights, such as moral rights. Section 9 deals with potential areas of concern with regard to liability for content of publications, covering such topics as defamation, obscenity, blasphemy, and provisions with regard to sex discrimination and advertising standards. Section 10 considers further issues including contempt, reporting restrictions, and secrecy. It does so mainly with regard to the press, although the author is careful to note that a number of the issues may equally concern multimedia producers. Section 11 reverts to a discussion of exploitation of intellectual property rights. Section 12 proceeds to discuss distribution agreements, then in Section 13 the author goes back to intellectual property again to consider the issues of infringement and enforcement. Sections 14 and 15 consider what may happen to rights in the event of termination of contracts and insolvency respectively. Section 16 considers the potential effect of competition law,

while section 17 examines the role of the EU and Council of Europe with regard to the regulating of the publishing and multimedia sectors. Section 18 examines the implications of data protection laws and practice. Sections 19 and 20 specifically examines the issues surrounding current and future use of multimedia, again with a heavy emphasis upon intellectual property rights.

While the choice of material cannot be faulted, its arrangement does perhaps seem unusual. There may well be clear arguments for separating the multimedia sections from the rest, but it is hard to identify the reasoning for the somewhat disjointed nature of the sections concerned with intellectual property rights. It would surely be simpler for the reader to have those sections gathered together, rather than distributed throughout the book, particularly when at times they are separated by subjects which have no intellectual property content. There seems to be no particular logic behind that distribution, and it would be interesting to know why the author felt that it was necessary.

With that presentational caveat aside, however, the book is well written, with the subject matter of each topic being both clearly and concisely presented. As such, it is hard to fault it as a primary reference text in this area.

With regard to the supplied disk, as already noted, this contains all but one of the precedent documents.<sup>16</sup> The disk does not contain simple editable documents, but software programs which allow for the creation of such documents or which automatically load them into the Windows clipboard. The reason for this approach appears to be an attempt to solve the problem of the multiplicity of word processors currently available to the PC market.

The DOS installation allows for the creation of the precedent documents in ASCII, RTF, WordPerfect 5.1 and WordStar formats. Thus WordPerfect and WordStar users may select their own format, while other word processor users, such as Microsoft Word users (of which I am one), may use the translation software supplied with their package to access the document.

The Windows installation makes use of the Clipboard feature to enable users to paste documents into the word processor of their choice.

The instructions for loading and using the software are contained on a separate sheet, and are clear and comprehensive. The software may be loaded in either the DOS or Windows environments, and both installations appear to be swift and bug-free.<sup>17</sup> When installed, the software is both well designed and easy to use.

All in all, the software is well thought out and executed package which adopts a sensible approach to the supply of ready-to-adapt documentation for the PC market, and the author and publishers are to be commended for using it.

Together, the book and disk make a useful package for both the practitioner and multimedia professional. While the author is at pains to make clear that it should not be used to replace specialist legal advice, it is an excellent guide to some of the pitfalls which await the unwary in the areas of publishing and multimedia. Some measure of its use and usability may perhaps be gauged from the fact that it is one of the books recommended in the *Copyright Guidelines for the Teaching and Learning Technology Programme* handbook.<sup>18</sup> It will also certainly go a long way towards helping to answer the many questions which are now being raised by individuals both in academia and business concerning the legal issues surrounding the creation and use of World Wide Web servers and homepages. With regard to its use in teaching, while its price obviously precludes it from widespread student purchase, it would certainly be a useful addition to the reference section of University law libraries.

---

<sup>16</sup> The book packaging agreement, based on the Book Packagers Association standard contract for Packager/Publisher Agreements is omitted as nominal consideration must be made to the BPA for its use.

<sup>17</sup> The test system used was a Pentium P90 with 32MB RAM.

<sup>18</sup> Published by the TLTP in Autumn 1994.

## Legal Issues Discussed on Mailbase

Andrew Charlesworth has now set up a Mailbase list for discussion of legal issues. Some details are given in the text below:

### **Description of List:**

A list for the discussion of UK and European issues in Intellectual Property law, with special reference to the impact of information technology and the Internet.

It is hoped that the list can address some of the problems facing those in HE, and elsewhere, with regard to use of materials in CAL and WWW page development. It has been set up in response to the increasing number of requests I receive for information on IPR in those and other areas. It is not intended to use this list to discuss issues of US law for which the US list *cni-copyright* is the appropriate forum.

The list will be open subscription and unmoderated - although I reserve the right to moderate the list at a later date should I feel it necessary. Non-subscribers will not be able to contribute mail.

To subscribe to law-ipr:

Send the message

```
subscribe law-ipr {your firstname} {your lastname}
to
mailbase@mailbase.ac.uk
```

To send a message to the list use the address:

```
law-ipr@mailbase.ac.uk
```

If you have a query about the list use the address:

```
law-ipr-request@mailbase.ac.uk
```

To unsubscribe:

Send the message

```
unsubscribe law-ipr@mailbase.ac.uk
to
mailbase@mailbase.ac.uk
```

## **CCTA Collaborative Open Groups Initiative**

The CCTA COG may offer some support for legal issues. Information can be found on WWW at:

<http://www.open.gov.uk/cogs/>

There are also some discussion group lists, the text below is extracted from the WWW page.

### **CCTA Collaborative Open Groups Initiative**

The CCTA conference on Information Superhighways launched the idea of Collaborative Open Groups to discuss further the issues raised at the conference. Since that conference, there have been calls for additional COGs to be initiated.

In order to simplify maintenance of the COGs and correspondence between COG members we have set up an email listserver to handle subscription requests.

Note that email addresses given here and in pages about the COGs themselves are set up as browser "hotspots". If your browser supports mail (and not all do), then you may mail directly to the address given from within your browser.

All requests to join any of the COGs listed here should be addressed to [listserv@ccta.gov.uk](mailto:listserv@ccta.gov.uk). The subject field should be left blank. The message should contain the words `subscribe mail_id cog`. For example a person with the mail address [Fred.Smith@company.com.uk](mailto:Fred.Smith@company.com.uk) wishing to join the COG discussing legal issues would send the message:

```
subscribe Fred.Smith@company.com.uk legal
```

The listserver program will respond with an acknowledgement. Full details of how to use the listserver can be obtained by mailing the single word `help` as the message to the listserver. If you normally attach a signature to your mail messages, it would be helpful if you turned it off when mailing to the listserver.

If you have difficulty using the listserver, please send a mail message describing the problem to:

```
listserv-manager@ccta.gov.uk.
```

Please note that mail sent to a COG mail address will be forwarded to all members of that COG. Requests to join the COGs should not, therefore, be sent to the COG mail addresses themselves.

---

## JANET ACCEPTABLE USE — ALISTAIR CHALMERS

---

*Alistair was unable to be at the workshop and his paper was presented by Derek Law. This paper is based on a combination of Alistair's slides and Derek's comments.*

Acceptable use policy has several originating strands. The first is now only important in terms of constantly reminding the world at large - especially the commercial world - that JANET is not subsidised, but reflects the ability of higher education to achieve discounted prices through bulk purchase. This was accepted initially on the basis that such use would be to a restricted group.

Secondly, all networks have AUPs and there was a need to have some conformity with our peer networks.

Thirdly, we had to be able to satisfy funding bodies - and indeed the National Audit Office - that all the expenditure was to the benefit of higher education and that there was no cross subsidy of other groups whether in the public or private sector. The ability to prevent free loading remains a significant concern.

However, there is going to be a change of emphasis in acceptable use which will in essence open the network to all groups - provided that they pay the full cost of connection to the network - I repeat, to the network rather than the site. This does not imply any change in those who may connect as of right. The eligibility rules remain and are essentially those directly connected with the UK HE community.

Peer networking remains a problem area but we are attempting to put suitable arrangements in place. We know that traffic which cannot be directly connected to JANET has been able to get in the back door often via a transatlantic link. There are too many networks out there with too many diffuse purposes. We are therefore persevering with peer agreements with the larger network operators where sensible links to JANET can be made.

Let me now turn to the acceptable use policy. Many of you will have noticed in the weekend press the arrest of an American student who had posted a particularly nasty sexual fantasy to alt.sex\_stories. The fact that similar material is readily available in published paperback is neither here nor there. And so we have a draft to which you should pay attention since quite apart from the sort of juvenile pornography which goes around, universities dabble in all sorts of areas ranging from animal experiments, to comparative religion to studies of race which someone will find offensive. So, JANET may be used for any legal activity, except:

**Exclusion 1:**

"use of the network to propagate material that is grossly offensive, indecent or of a menacing nature, or material intended to misinform and thereby cause annoyance, inconvenience or needless anxiety to another"

This exclusion is the most difficult of all. The key words here are probably propagate, grossly and intended. It seems to me inescapable that sites must expect to police this kind of activity to some extent. In particular web servers, unlike say electronic mail, have a quasi-permanent existence which we may reasonably expect institutions to police. Institutions will have to move quite quickly, I believe, to put some regulation in place which protects everyone.

**Exclusion 2:**

"use ... to propagate unsolicited material either to organisations directly connected to other networks - e.g. commercial advertising and promotions; however it is acceptable to place such material on, for example, bulletin board systems connected to JANET, subject to the agreement of any management involved. Similarly, it is acceptable to publish material to an electronic mail list, provided that the material is related to the intended use of the list."

Exclusion 2 represents our first nudge towards the commercialisation of the network. Although billed as an exclusion, this actually allows inclusion of some forms of advertising, but under a carefully controlled regimen.

**Exclusion 3:** "deliberate unauthorised access to facilities or services accessible via JANET"

Exclusion 3 is fairly straightforward and is aimed at preventing things like non-BIDS sites connecting to that service.

**Exclusion 4:**

"deliberate activities with the following characteristics:

- wasting staff effort or networked resources, including time on end systems accessible via JANET and the effort of staff involved in the support of these systems;
- corrupting or destroying other users data....;
- violating the privacy of other users via the network;
- disrupting the work of other users via the network;
- using the network in a way that denies service to other users (for example deliberate overloading...);
- continuing to use an item of distributed software after UKERNA has requested that use cease because it is causing disruption to the correct functioning of the network;
- other misuse... such as the introduction of viruses."

Exclusion 4 is, I think a slight softening of the rules. You will of course all know that JANET may not be used for social chit-chat and that such listserv groups as Mornington Crescent (who do nothing but play that game) should not exist. This becomes wasting resources, a much looser concept which most of us can however comprehend.

Wasting staff time is probably the worst crime since there are so few of them and they really represent a valuable resource. And of course corrupting other users data is almost too terrible to contemplate - but in an increasingly competitive and irrational world it's no longer almost unthinkable.

This is followed by a list of other fairly obvious sins. Although here again the difference between network acceptable use and the law is interesting. Then of course at the end we have the catch-all.

These are all fairly general concerns so let me now particularise on the problems of web servers, where the major concern is much less to do with acceptable use and much more with the voracious



consumption of resources. This is especially true of international traffic where everyone seems to follow the same broad pathways, which are now becoming traffic jams.

The sub-committees of the JISC are now beginning to turn their attention both to the growth in this traffic and to the appearance of network hotspots. Bandwidth has been significantly increased this session, but we do not want to spend money on helping everyone to fool around making needless connections. We want to spend the money on research and on network content rather than on telecomms costs. There then seem to be three serious options and a small group is beginning to work on these. Charging for international access may have to come and now that charging for JANET itself has been approved this will not be difficult to push through. A similar effect could be achieved through rationing, but I at least am reluctant to see this happening when we are trying to stimulate all sorts of groups to maximise network use. Thirdly comes caching, whether nationally or locally, and this is something I would like to see developed strongly.

We may also need to develop some self regulatory mechanisms such as better self regulation and minimising the distribution of inessential graphics and sound. I hope that you will spend time tomorrow looking at this issue. The Web is a splendid and attractive development but it is beginning to develop and sprawl in a way which invites punitive controls.

---

## ISSC - SPIDER OR FLY? — DEREK LAW

---

*Derek then built on his presentation of JANET Acceptable Use to turn to his own paper.*

Let me then turn to this issue of local and national caching which might be taken as one description of the JISC's network services policy. I hope to persuade you that the JISC has a valuable role as the spider at the centre of the web rather than being its victim. The objective is clear, to create a central core of material which is centrally defined but meets user needs in all disciplines. The user will then have a limited need to search for materials outside the core. We will spend our resources on developing that core rather than on cataloguing anything that might ever be used on the Internet. In doing this we hope to provide a variant of Gresham's Law. While bad money may drive out good, we hope that quality assured data, available reliably and with excellent nationally prepared documentation will remove the need to use unknown data of unknown validity available intermittently and unreliably.

JISC has two sub-committees concerned with content provision. The first is ISSC (the Information Services Sub-Committee) which has had a working group called ANIR which has been looking at network resources and whose recommendations I will describe later. It also controls FIGIT, the Follett Implementation Group on IT which will be bringing a whole range of network resources to the community over the next year or so.

We are then looking at three major issues over the next year or so and have substantial budgets to deliver product.

### *Training*

JISC has some experience in this area, for example through the ITTI programme. The new chairman is particularly keen to see developments here, whether for completely new groups such as the Colleges or whether of new tools for existing groups. FIGIT will be putting perhaps a million pounds into a training programme over the next year or so and there is clearly scope for fitting new tools into this.

### *Content Provision*

Content provision is perhaps less relevant to your immediate concerns, except to say that we are very keen to spin or cache major resources in the UK. Clearly this cannot be too rigidly planned but we have done well in distributing resources thus far. The guidance of the technical advisory unit based at Kent will in future be important in determining the destination of resources.

### *Hotspots*

Hotspots are a worry. We are funding some experiments in major resource gateways, beginning with SOSIG, the Social Science gateway. There is an implication in supporting such services that we will ensure that the resources they point to will be properly supported and accessible without network degradation.

So, in support of the development of such services we have been developing criteria. These may seem fairly obvious but we think them none the less important. They all reflect the fact that our experience shows that the real costs of resource provision are ownership costs. We must expect that use will be high, typically in the hundreds of thousands of connections each month. We want to distribute the resources around sites and not simply create hotspots in one or two places. The services will have to contract to provide training and support. The documentation we have funded, for example for BIDS, shows what can be achieved. Nor is it good enough to suggest that it can all be left to on-line help. Finally there must be coverage of the whole community. This has two aspects. Firstly, we want to provide resources for the whole HE community. Mediaeval historians are as important as particle physicists. Equally a resource aimed at some segment of the community such as philosophers must be open and not restricted if it is to have central support.

Finally, then let me turn to the specific activities we have in hand at the moment. We have commissioned a study to look at setting up a UK equivalent of CNIDR, the clearing house for information on network tools. I hope that this will go out to tender before the summer and it will have amongst its aims the creation of a clear focus for those involved in the development or implementation of resource discovery tools in the UK.

The InterNIC Scout service is another US invention which we are considering moving to the UK again as a means of alerting and informing politicians as well as technicians to the developments which are of concern.

A structure is emerging for registering gopher servers and we want again to have a clear UK focus. It is the clear intention of the JISC that the UK should be an information producer as well as a consumer nation and this will require clear structures to be managed.

Standards is a permanent concern of the JISC and Professor Mike Tedd convenes a committee in this area. We want to be players as well as receivers and certainly to try and harmonise UK resource. Everything from URLs to the work of IETF and resolving problems of Netscape licensing are grist to this mill and AGOCC plays a notable part in work in this field.

Lastly, we are attempting if not to catalogue the Internet at least to provide guidance to the major resources, using tools such as SOSIG.

To summarise, the rapid emergence of World Wide Web technology as a captivating easy to use tool has perhaps taken us all by surprise. It is the job of the JISC to put some kind of order into the anarchy which is emerging and I hope that Alistair and I have demonstrated some of the issues which its emergence has highlighted. So can I close by stressing that it is the role of the JISC to act **for** the community and not **against** it - always remembering that the community is not just you, but the whole of higher education. I hope that positive recommendations will emerge from this workshop; we have some resources to do sensible things and help resolve or smooth the action issues you identify.



---

# EXPLOITING AND PROTECTING THE NETWORK

## PETER KEMP

---

UCISA have developed guidelines for "Exploiting and Protecting the Network". They note that:

"The explosive expansion of the Internet will lead to radical changes in the way our students and staff gather and use information. The teaching world is showing great interest in the use of World Wide Web technology both as an authoring tool and as a facilitator of distance education. There are also new opportunities to disseminate research findings, independently of the traditional publishing process. The underlying ethos of the approach to Internet use will be to encourage widespread use of the network for teaching, learning, scholarship and research. Students should be able to make their information available. Such free access, however, could lead to misuse of the international networks."

The UCISA Guidelines consider ways in which institutions can prevent abuse of the networks whilst empowering users to exploit networked information services to the full. The report makes a number of recommendations. These are:

- "• Institutions should institute training programmes incorporating guidelines on responsible use for students and staff.
- Institutions should consider their policy on the provision of networked information.
- Institutions should reconsider their internal regulatory policy in order to ensure that it is adequate in the current environment.
- Institutions should formally review their computer and network security policy.
- Institutions should ensure that all users are registered and that all usage is logged.
- Institutions should review their attitude to the monitoring of network traffic and ensure that their IT staff have sufficient authority to carry out any necessary investigations of misuse.
- Institutions should do their best to ensure they supply a NEWS feed from the Internet to their systems which is free of offensive material.
- Institutions should ensure that they appoint a Computer Emergency Response Team (CERT) contact who has adequate internal powers to take actions recommended by CERT.
- Institutions should ensure that their network operations provide adequate local support for CERT activities out-of-hours."

Note that these are extracted from the draft guidelines and that readers should access the latest version from NISS under the UCISA pages.



---

# APPROACHES TO WIDE AREA INDEXING

## MARTIJN KOSTER

---

There are a number of approaches which can be taken to indexing. These include:

- manual indexing
- robot assisted indexing
- automated distributed indexing

### Manual Indexing

Manual indexing includes both personal hotlists and public hotlists. Both have a number of major problems. Personal hotlists tend not to be up to date or to be comprehensive. They also tend to collect references which have become outdated. Public hotlists have a high signal to noise ratio and permission needs to be sought to update or remove information.

### Robot Assisted Indexing

An example of this is Lycos. On the positive side, these tools provide automatic indexing. However, they do have problems. They tend to overload the network and/or the host. They can also give the wrong impression regarding the resources to the person searching for information. They can also provide too much information. The indexing is centralised.

### Manual Distributed Indexing

An example of this is ALIWEB which is described as:

"ALIWEB is a system that automatically combines distributed WWW server descriptions into a single searchable database. ALIWEB basically does for the WWW what veronica does for gopher or Archie does for anonymous FTP. Because the original server descriptions are maintained by server administrators, the information is likely to be correct and up-to-date. It also uses a special format that makes the results look very concise."

Aliweb is a public service run by NEXOR. See

<http://web.nexor.co.uk/public/aliweb/aliweb.html>

Aliweb has a number of advantages. It is simple and cheap. It has high quality summarising. There are likely to be fewer stale references. It does still need manual effort though and uses centralised or mirrored searching.

### Automated Distributed Indexing

An example of this is Harvest.

see <http://harvest.cs.colorado.edu/>

Harvest is an integrated set of tools to gather, extract, organise, search, cache and replicate information across the internet. It is therefore designed to help users find information as well as helping in its management.

Harvest has a number of advantages, not the least of which is that it is available. It is automatic, extensible, scaleable. On the negative side it is complex. It has the potential to offer a general search interface, automated summarising and distributed searching.

## **In Summary**

We need to have indexing tools which are automated. The solutions need to be distributed ones. We need to adjust peoples' expectancies so that they understand the reality of the problems and available solutions. Harvest may offer a solution. We also need to accompany the use of tools with high level manual resources which complement what we can achieve automatically.



---

# THE PROVISION OF EFFECTIVE LOCAL, NATIONAL AND INTERNATIONAL WEB SERVICES

## NEIL SMITH

---

When considering an institutional Web strategy there are three points to be borne in mind.

1. Demands on the hardware
2. Using simple, powerful software
3. Maximising network efficiency

### **Demands on the Hardware**

The exact hardware requirements, and the organisation of that hardware will be based on local considerations. In general, the community feeling is that distributing work across several machines while retaining a central server for general facilities is the right approach. However, this does not always fit in with working practice. A number of guidelines can be laid down.

In general the process of serving documents on the web does not require a large CPU resource, unless the server in question is very popular. Exact figures cannot be provided as there are too many variables in each individual system.

Whilst CPU requirements are, in general low, network bandwidth requirements are nearly always high. The only way to avoid this is to have an unpopular server and only serve text.

The authoring of web documents is not particularly demanding of CPU and will often be performed on a separate remote machine.

Specific services being provided on the web, for example database searches, may require that the server be on a specific machine.

The level of trust that the providers of the web server have in the authors of the documents being served may dictate specific security precautions that have to be borne in mind. If this level of trust is low then some server facilities may not be available, for example CGI scripts.

### **Using Simple, Powerful Software**

As far as software is concerned, an official service should be based on established software. This limits the choice somewhat, but fortunately the software that is available is very good. Servers running on PCs and Macintosh machines are not considered as these machines are not suitable for most large organisation servers. The choice is limited to either CERN or Netscape Netsite.

Both of these products are easy to install and maintain.

In general terms each of the products supports a common set of capabilities.

The CERN server is freely available. The Netscape product is commercial.

In terms of secure communication the Netscape product is ahead of the CERN server. CERN includes no support for security.

Development work on the CERN server at CERN seems to have stopped. The work has apparently moved to INRIA in France, although there has been no news concerning future directions on the server. As the Netsite server is a commercial product we can be sure that it is tracking upcoming developments.

## **Maximising network efficiency**

A number of recommendations can be made in order to maximise network efficiency.

Every institution should be encouraged to install a cache at their local site. Even if this is very small (say 200Mb) it will keep copies of the most popular pages. This cache should be configured to take advantage of a national cache.

At the moment there is are two well known public caching proxies. One at HENSA Unix and one at Imperial College. All local caches and browsers should be configured to take advantage of one of these caches. In the future it is hoped that more official national caches will come on line. These will be distributed around the country and each institution should make use of the closest.

---

## TEACHING AND LEARNING

---

### Introduction to the Issues — Roger Rist

The WWW has been used at the Institute for Computer Based Learning (ICBL) at Heriot Watt University since 1993. The ICBL provides servers for TLTP (Teaching and Learning Technology Programme) and LTDI (SHEFC's Learning Technology Dissemination Initiative). They run a Computers in Teaching and Learning module as part of the MSc course in HCI. Various projects are also relevant to the topic of the role of the WWW in the provision of, and access to, resources for teaching and learning. These include ELF (the Electronic Learning Facilitator).

Tim Berners-Lee described WWW in the following way:

"World Wide Web is a wide area hypermedia information retrieval initiative aiming to give universal access to a large universe of documents. The project merges the techniques of information retrieval and hypertext to make an easy but powerful global information system"

WWW offers many possibilities for use in teaching and learning:

- information systems
- electronic publications
- virtual libraries and museums
- teaching and learning resources

### Information Systems

These can include:

- Campus Wide Information Systems
- virtual campuses, with clickable orientation maps
- provision of University prospectus information
- course aims and objectives
- lecture notes and lecturer's slides and handouts
- coursework information

## **Electronic Publications**

These are emerging and include:

- online journals, for example the Communications of the ACM is now available online
- Bioline, which provides free access to abstracts but subscribers to the articles pay a subscription fee
- Wired Magazine
- Times Higher, including advance information regarding job advertisements
- Encyclopaedia Britannica Online
- Daily Telegraph

## **Virtual Libraries and Museums**

Examples include:

- The Electronic Beowulf (British Library)
- Gutenberg Project
- Le WebLouvre
- Natural History Museum
- National Museum of Scotland Mosaic Project

## **Teaching and Learning Resources**

The WWW has the potential to offer access to, and information about, resources as well as being used for distribution and delivery.

Information is being made available via the WWW and includes information about projects, for example CTI, TLTP, ITTI as well as access to some of the deliverables from these initiatives.

Some projects are making demonstration or sample versions of teaching and learning resources available via WWW.

There are also a growing number of examples of using WWW for courseware delivery. Examples include:

- the frog dissection example
- control of remote robotic devices
- the ability to interact with simulations, for example in engineering, guided tours
- BioNet Glycolysis example

As a result, HTML is becoming a new authoring standard. This has implications for the extensions needed to support this community.

The WWW can also be used for assessment purposes. This might involve multiple choice questions with answers being automatically marked.

There are also possibilities for students conducting learning tasks to produce multimedia essays with links. Students can be provided with review exercises to conduct in relation to resources available on the network. These reviews provide useful information to future groups of students.

Another example of use of the WWW is the "Answer Web" based on Mark Ackerman's Answer Garden concept. This provides a structured hierarchy of FAQs (Frequently Asked Questions) with questions routed to an expert. The answers to queries are placed in the Answer Web. The garden does need to be weeded though from time to time but has the potential to provide a useful and growing web of knowledge.

The WWW has the potential to offer more than access to resources. Recent developments point the way to greater use of the WWW in collaborative work with use of whiteboard facilities and the integration of video conferencing within the WWW. The Open University Summer School used video conferencing in 1994 with classes of 16 conducting supported learning for two intensive weeks. A major question is whether such an experiment will scale. Other examples of course provision include free courses with global access (introduction to C++, the protein structures course at Birkbeck). There are also an increasing number of synchronous meetings through MOOs. A distance learning MBA exists which has 10,000 students with distance learning materials which are becoming available via WWW, optional software and no tutorial element. There are 80 examination centres worldwide for this.

As far as teaching and learning issues are concerned, technology is **not** the problem. The main issues are: access, accreditation, authority and attitudes.

#### *Access*

Before widescale use of the WWW for teaching and learning is possible we need to address access issues. Will all students be given access to WWW? Will staff and student from outside institutions be given access to resources available on WWW? Do we really want to share our best resources?

#### *Accreditation*

Are students really paying for computer resources or are they paying for more than that? Does it matter how they learn? If we use computer based resources a great deal, what distinguishes one institution from another and how do we get our competitive edge?

#### *Authority*

How do we work out who to trust in when using the internet?

#### *Attitudes*

How will staff react to the new opportunities presented by using online resources? There are some bad lessons from the past, we need to ensure we do not repeat these. But, can we afford to ignore the opportunities?

In conclusion, the WWW offers many exciting new opportunities for teaching and learning. Teachers and learners can obtain many benefits by seizing these opportunities.

## WWW as a Strategic Tool in Research and Teaching of Molecular Sciences — Henry S Rzepa

The World-Wide Web has its origins in promoting information exchange amongst the high energy physics community, but was rapidly adopted by molecular scientists, often before it was adopted by computer centres and support staff, and has now achieved a productive maturity in this area. During the period 1993-5, amongst the many accomplishments by UK scientists that can be listed are:

- An on-line Protein Structure Course, involving around 40 "consultants" and in excess of 100 registered students.
- The Seminal work at Sheffield University relating to the Periodic table.
- Chemistry and Biology Workshops organised at the first WWW meeting at CERN by UK scientists.
- Involvement by UK scientists in the First International Chemistry Conference based around WWW and E-mail, held during November 1994.
- The planning of a UK mainstream chemistry conference using WWW and based around the "Young Chemists Committee", in collaboration with the Society for Chemical Industry.
- A formal proposal to the IETF for appropriate chemical MIME types to be used with WWW, with further approaches to the IUPAC committees planned.
- An expression of interest under the FIGIT initiative to publish a chemistry journal, in collaboration with the Royal Society of Chemistry.
- The start of a Global Instructional Chemistry initiative, illustrated by the integration of WWW technologies into teaching courses at Imperial College.
- The appearance of two scholarly refereed papers in mainstream chemistry journals describing applications of WWW to the science, and the production of a major book on applications of the Internet to chemistry, to be published by the American Chemical Society.
- Invitations to UK scientists to give keynote talks in American and Europe on their WWW accomplishments.

In formulating institutional guidelines we have to create an environment where individual efforts by such scientists are not stifled by excessive centralised regulation and control, yet one in which the essential generic tools are well supported. We also have to plan how to react to the trends already visible at the Second WWW meeting in Chicago toward the commercialisation of WWW. At this meeting, one left with a feeling that the interests of science in particular and teaching in general are not at the top of the priority tree for commercial developers, and we are reaching the end of the road where hitherto immensely successful "spare time" departmental enthusiasts can exist in isolation. An efficient and permanent support infra-structure is essential. I would nevertheless argue strongly that such support structure should belong within individual departments as well as within a centralised support service. Experience has shown us that intimate subject knowledge is essential for innovation and that effective communication with subject experts is essential for progress.

Many issues will need to be addressed in the near future. These include

- Developing support infra-structure for Uniform Resource Names, for which standards are slowly emerging. As with DNS names, a central support service appears essential.
- Furthering the HTML dtd into versions 3 and 4, and in particular acquiring and/or developing appropriate parsing and setting tools which may not become immediately available from commercial sources. The BBEedit tools developed at York by Lindsay Davies are a shining example of such support.
- Developing the relationship between SGML, HTML and scholarly publishing, and ensuring effective liaison with appropriate FIGIT initiatives.

- Evaluating successor or alternative systems such as document cluster browsers (Hyper-G) and virtual reality browsers implementing VRML. The chemists are indeed at the forefront of such areas, being particularly active in virtual reality applications and the emerging field of chemistry collaborators. Here, the SuperJANET network will have a major role to play.
- Implementing Z39.50 aware and other robust indexing and searching systems, with particular regard to more specialist subject disciplines. For example, FreeWAIS-SF is a highly customisable indexing system for which specific configurations for specialist groups will need to be developed. In this category also fall tools for integrating mail and World-Wide Web, such as Hypermail. This utility currently does not support MIME enclosures and has not been developed for indexing.
- Supporting scholarly journals in molecular science via local and central library activities. Again, subject interests will be paramount, and it will be essential to involve active academics in such developments.
- A major growth area is envisaged for Internet "robots" which are sent out with specific tasks to perform. Many of these tasks will be subject specific, and such agents will need involvement from academics.
- Preparing publicity material for subject disciplines. Many practising scientists are not familiar with these areas, and they will want promotional material that has a strong subject content before they will come to accept this technology.
- Promoting the role of learned societies and industry is essential, and much still needs to be done to develop these links. Such links are again best fostered within departments by individuals.

In conclusion, I strongly believe that the development of World-Wide Web as a Strategic tool in UK higher education must be a collaborative venture between support agencies and academic departments. In particular, the consultative procedures between the two must be robust enough that duplication of effort is minimised and that key enthusiasts in departments feel appropriately consulted and supported.

## World Wide Web as a Tool for Teaching and Research in Visualization — Ken Brodli

My interest in the web is as a person who lectures and does research in graphics and visualization - that is, primarily as a user of the information.

Clearly there is no shortage of information. A quick search using the WebCrawler for the keyword 'visualization' returned some 850 URL's - and a further 100 when the 'z' in 'visualization' was replaced by 's'! This is too much information, and useless in an educational context. However help is at hand through the Scientific Visualization Weblet, which lists a manageable set of some 50 important URLs with a paragraph on each. This is maintained by NASA in the US. The URLs are structured in four groups according to the information provider: universities; government labs; commercial; and military.

But how useful is this material in education? Can a lecture course be created simply by recommending a particular traversal of the information?

In a recent exercise with my MSc visualization class at Leeds, I have set out to answer these questions. I asked the students to review the URLs in the weblet and report back with:

- a description of the URL, its content and origin;
- a measure of the quality of the information, and its 'up-to-date-ness'
- its relevance to a course in visualization
- at what point in their course should a student be pointed at that URL.

The results were interesting. First of all, the students really enjoyed the exercise and I am convinced it was a good educational experience for them. They found good background reading material on applications of visualization (for example in chemistry) and good descriptions of visualization products. They were stimulated by being able to read about the latest research - which was often reported in a less intimidating manner than many papers in journals. They enjoyed the colour images and videos.

But all was not ideal. The students found the material to have a promotional rather than educational bias - even research material has a slant which seems to be advertising the institution. Not all URLs are kept up-to-date. The 'here today, gone tomorrow' URLs were also a problem (not least when it came to marking!)

So the lectures cannot yet be replaced by the WWW tour - at least in visualization. A simple improvement one can make is the structuring of a visualization weblet - it would be nice to have it structured by information, rather than information provider, and a final year student at Leeds is currently helping me do this.

As a visualization researcher, the web offers three new opportunities. First, it is a novel publication medium whereby research results may be reported. Second, it is a source of information when tackling a new research topic, or when training a new research student - and indeed I have successfully used it in both contexts when beginning research in visualization of medical imaging using SPECT. Most exciting of all to the researcher is the possibility of using WWW as a means of collaboration - either between visualization researchers, or as a tool for enabling collaboration in visualization itself.



## **Global Network Academy — Alan Mills**

Alan Mills described the work of the Global Network Academy (GNA).

He specifically referenced the work on protein structure. The Principles of Protein Structure organised by Birkbeck College in collaboration with the Virtual School of Natural Sciences (VSNS) of the Globewide Network Academy (GNA) can be seen through:

<http://www.cryst.bbk.ac.uk/PPS/index.html>

The GNA are actively addressing the following:

- (virtual) student registration and administration
- course ethos and practice, e.g. learning contracts
- communications using Listservers, email, MOOs
- assessment and accreditation

It was recommended that we should liaise with them.



---

# **WWW - OPPORTUNITIES AND OBSTACLES FOR ADMINISTRATORS — PETER TINSON**

---

## **Introduction**

Development of the World Wide Web and the volume of information accessible through the Web has grown dramatically in the last two years. However, universities' administration departments have yet to take advantage of the new technology in any great way; this paper examines the potential for use of the Web by administrators and some of the obstacles that may impede such use. The author has not made great use of the Web himself and as such some of the ideas for use may already be available and some may not be possible to achieve.

## **Around the Administration Departments**

### **Research and Consultancy**

Research and Consultancy is one area where the Web could make a significant impact. Many of the funding bodies are already connected to the network and some are starting to distribute information electronically. However, the information on the funds available for research comes from a variety of sources and not from the funding bodies alone. Some, such as Refund from the University of Newcastle, are collections of funding information and more general information of interest to the Research offices. The information, however received, is usually photocopied and distributed to the academic staff across the campus. Academics wishing to submit a proposal for funding have to complete what is, in many cases, a standard form and a paper supporting their proposal.

It is clear that a great deal of time could be saved in the research offices if the information emanating from the funding bodies and subscription services such as Refund were available as Web pages for all to view. However, there are some problems with this approach. The first is that the research administration and academics alike are used to being given the information rather than having to go and look for it. Provision of the information on the Web implies a significant change in culture which may prove difficult to achieve. It is possible that if an information provider changes from distributing their information via electronic mail or by post to displaying it on a Web page, that they may need to advertise that new information has been posted in order to get the intended recipients used to looking for information rather than waiting for it to be given to them.

Some of the services which collate information are subscription services and as such should only be available to those that have paid. Before such services can be made available via the Web, there needs to be a method of ensuring that only those sites that have subscribed are able to view the information; those who have not should be able to see a page giving details of the information available and how to subscribe.

The use of the Web need not just be one way. Many of the funding bodies have standard forms for applications for funding; these could be mounted as a Web page for research offices to complete. The supporting paper could then be sent on either by electronic mail. However, when making an application detailed costings of the proposed research project have to be submitted. These should be verified by the research office and in many cases additional verification of the application by senior University Personnel is required. Therefore, whilst it may be possible to allow downloading of the appropriate forms by any individual, it is not permissible to allow applications to be submitted from anyone; the application must be received from a verified source.

Finally, if details of the successful applications were held on the Web, it would be a small step to attach any papers resulting from the research to be held in the same place as the original application and thus build up database of research papers. However, before this can be achieved there is a need for papers published in this way to be refereed ensuring that the publications are of a sufficient quality.

## **Finance**

One of the main issues in the Finance departments regarding the use of computing networks has been that of security; it is not generally acceptable for people to be able to view other's personal information and is certainly not acceptable for people to amend it. This means that if the Web is to be used to access central finance computers, there will need to be adequate security in place to prevent unauthorised access. The Finance officers are perhaps more wary of change than others and so even with appropriate security in place, the benefits of Web access to their data will have to be sold to them.

One area where the Web may be of use is the publication of annual accounts. Every organisation has to publish its accounts and often a considerable amount of money is spent producing and distributing them. Publication on the Web could reduce that cost considerably, although there would probably still be a need for printed copy.

## **Personnel**

The obvious area where the Web can be used by Personnel departments is recruitment. The Times Higher Educational Supplement vacancies are now published on the Web some three days before they appear in print. Some sites have experimented with putting their vacancies up on their campus servers. However, it is too early to say whether advertising on the Web will be successful or not. Certainly the Web does not reach everyone at the moment in the same way that a newspaper can and for the immediate future there will not be any significant shift to using the Web exclusively for recruitment. One of the problems is that the vacancies will have to be located in one place in order for there to be a major uptake in Web advertising. It is unlikely that potential employees will look at the vacancy pages for all institutions if they are just looking on the off chance. The collection of vacancies in publications such as the Times Higher Ed allows easier browsing. The Web page though would be useful for local recruitment. Again, a change in culture may be required; potential employees will have to look rather harder to locate the information rather than have it delivered to them.

## **Registry**

The Registry is one area where some use is already being made of the Web; many institutions have already made their prospectuses available and the benefits are clear. However, there are areas where a link with the host student records database would be useful, particularly in student recruitment. It should be possible for sites to display the availability of places through Web pages which, if coordinated at a central point, would give prospective undergraduates a clear picture of where the places are, what the entry requirements of the institutions with vacancies are, and so who to contact in order to try to secure a place.

## **MAC**

Although those implementing MAC software are split into four families, there seems to be little co-operation or exchange of information between the sites within those families. Part of the reason for this is that implementation plans are sometimes a closely guarded secret and so it is difficult to find out who is working on the same modules. Clearly it would be of some assistance to be easily able to find out which sites had implemented which modules so that some exchange of information could take place without having to seek out who had done what.

Another area where the Web could be of use is fault reporting. There are always likely to be problems with items of software as big as the MAC suite. The need to maintain a full list of reported problems and their solution is important as not all sites are running the most recent version. The list of problems is extensive and it is not easy to search through in order to see if the problem a site has just found is a new fault or one that has already been reported. Again, the organisation of this data into sections for each form would assist searching and so help prevent such duplication. The Web could also be used as an alternative mechanism for reporting faults.

## Meetings

One feature of administration offices is meetings. Considerable time is spent travelling to meetings to meet ones peers for only a few hours; clearly this is one area where the Web may be of some assistance in the short term. The JANET User Group of Administrators (JUGA) has experimented with electronic meetings, firstly with an electronic mail meeting and more recently with a video conference. The latter is probably the way forward but few sites currently have the facilities to act as a node for such a meeting. The electronic mail meeting may well be repeated but in a more structured fashion as it was difficult to follow the agenda with each participant responding to whichever item was of interest.

The Web offers the ability to hold meetings across the network. Separate pages would be available for each item on the agenda with each participant being able to view the comments made already and to add their own comments. Voting could take place on a separate page with the participant clicking on an appropriate box to indicate their response. However, there is a certain amount of setting up that is required before such a meeting can take place and this will need to be reasonably easy for an administrator to achieve if the Web is to be used in this way. There will need to be easy to set security before the meeting could begin as it is unlikely that the facilitator would want anyone other than the participants to contribute or even see the discussions taking place. The tools for this are not yet available to the author's knowledge.

## Conclusions

There is little doubt that there is great potential for administrators in using the Web both to access and exchange information and to electronically submit standard information to bodies such as research councils and government agencies. However before this potential is realised the Web itself needs to be sold to the administrators; its benefits highlighted and demonstrated.

It is important to establish friendly mechanisms to access the information on the Web. Any browser must be easy to use and it is desirable to have simple access to key information, either by gathering it all in one place as the JANET User Group for Administrators have done, or through refined searching where perhaps searching on any given term returns a limited subset of the information available on the network. Use of the Web by administration departments will be limited if it is difficult for users to spot items of interest if keyword searches return a large number of largely irrelevant items.

Similarly, there must also be easy mechanisms for loading information onto the Web. Forms based input may assist here but there is the need for wordprocessor to HTML convertors for a large and diverse range of word processors. Institutions computing services must also be prepared to assist in locating information in the local Web structure.

There is a need to ensure that secure information can only be accessed across the Web by those who have permission to see it, particularly with private meetings and subscription services such as Refund, and there needs to be a simple mechanism to allow administrators to set up their own sections and set appropriate security and restrictions.

In order to make effective use of the Web, there needs to be a change of culture - administrators will need to start seeking information rather than waiting to be given it. However, above all, the benefits of the Web need to be demonstrated and sold to key administrators before there will be any major take up.

### Protocols

#### HTML

The HTML language is developing:

HTML 1 - Tim Berners-Lee initial specification

HTML 2 - NCSA's extensions implemented by Mosaic (e.g. forms).

HTML 3 - Currently under discussion, although many new features are probably finalised (e.g. tables, maths)

HTML 4 - Ideas being discussed - such as inclusion of a safe, platform independent authoring language.

Note that ideas for VRML (Virtual Reality Markup Language) are currently being discussed. Further information is available at the URL:

<http://www.wired.com/vrml>

#### HTTP

HTTP 1.0 is being standardised. Ideas for HTTP/HG are currently being discussed. Enhancements to security (SHEN, SHHTTP) are also being discussed.

#### URIs

URIs comprise:

- URL — Uniform Resource Locator
- URN — Uniform Resource Name. An object may be replicated on WWW. A URN may be regarded as the name of a unique object. When a URN is selected a URN-URL resolver will find the "nearest" location of the object.
- URC — Uniform Resource Characteristic (or Citation). Meta-information about an object, such as keywords, copyright details, cache expiry dates, etc.

#### CCI

CCI (Common Client Interface) provides a mechanism for integrating applications on the client with WWW browsers. This may be an important technology for the delivery of distributed teaching services.

### Application Areas

The following application areas are likely to become important:

- Collaborative Systems — Such as WIT.
- MOOs and MUDS — For support of distributed teaching
- Ecash — For charging each other?
- Browser developments — Such as support for CCI and integration with collaborative systems

## **Organisational Issues**

Involvement with W3O (the World-Wide Web Organisation) is likely to be of importance to the UK academic community.

Involvement with IETF standardisation groups may be of importance.



---

## DISCUSSION SESSIONS

---

### Teaching, Learning and Research Issues

Various inputs were made to the discussions on the issues and people's experiences. These are points which emerged:

All Computers in Teaching Initiative (CTI) Centres have their own WWW page.

Considerable interest from librarians in acting as mediators between the information and the recipients — making it all easier to use and more useful. There are difficulties for librarians in identifying and collecting material. It is often ephemeral and it is hard to identify what it is intended for.

There are developments in industry and we should liaise where appropriate (and possible).

We need to liaise with learned societies and SIGs.

There is a need to have refereed journals available on the network and to have appropriate subscription mechanisms — this is being dealt with through FIGIT (Follett Implementation Group for IT).

There can be problems with students becoming disillusioned at the poor quality of material.

We need to ensure that materials are not "amateur" and to train people in their production.

We need to think about futureproofing the content — the WWW is the delivery vehicle and provides some structure which is likely to last and be portable. Will the WWW be the tool we use in 5 years?

In HE, we are in a great position to collaborate, and we should/must do this — saves time (effective, efficient).

We will not always use online experiments — there will still be chemistry labs.

How do things scale up? MOOs may be great fun at first but can they work with larger numbers?

Students can usefully help each other using online facilities.

How can we offer help to students lost in a sea of information?

Learning is a very personal thing — using WWW is just a part of a strategy for learning.

We need help — a support unit? — to get people off the ground in using the technology.

### Summary of Issues Presented at Plenary Session

The main issues which emerged from the discussions were:

How can the WWW be used effectively in a variety of learning and research contexts:

- training of staff, student and trainers
- training materials
- quality and relevance of materials

- copyright and intellectual ownership
- peer review
- liaison

The group considered that these could be assisted by a number of actions:

- 1 Various training events and seminars should be run. These should include courses for staff development, particularly as part of induction courses. There should also be subject based symposia.
- 2 There is a need to develop a quality code of practice for information providers.
- 3 An initiative to stimulate the production of Web-delivered courses should be mounted.
- 4 A project should be funded to identify good practice in Web-delivered courses.
- 5 The outcomes of this workshop and other events which follow on should be widely disseminated.

## Supporting Information Providers

### Need for a support infrastructure

Sharing and collaboration are vital, self-help initiatives crucial, but we need a national infrastructure to support those at the local level supporting their local information providers.

There was a recurring emphasis on the need for **human** resources - both at this national level and at the local level. Somehow the message must be got across to universities that using WWW is going to be an essential medium for accessing, delivering and exchanging information and of delivering teaching and learning, but that this cannot be done without new or diverted personnel and should not be regarded as free or a way of making savings.

### National coordinator

The group focused on the need for at the very least a national coordinator - someone whose job it was to keep an eye on the futures, set up initiatives to solve technical problems, collect in findings and **pass on information** in a coordinated way.

There was talk also of the need for a server (below), ad hoc task groups, and the value of meetings as a way to exchange information and reduce isolation.

### An Information Providers Web Server

This was seen as the most appropriate tool for a national coordinator to use for delivering information (which did not need to be generated centrally, of course, it could come from one site or be the product of a task group, etc).

The sorts of topics to be covered on the server reflected the subsequent discussion on what needed to be done to support providers:

- tools
- draft guidelines for information providers
- technical solutions (eg in areas such as indexing, caching, graphics...)
- documentation and training materials
- case studies
- news/updating information/awareness

### Tools/software

The group agreed that there needed to be a mechanism (a **small** group?) for identifying what tools were available, making recommendations, ensuring the tools were available in a number of UK sites, if appropriate passing information to CHEST for an agreement to be negotiated, and, in the case of tools on the 'wishlist', taking steps to ensure their development.

The tools (and guidelines) fell broadly into the following categories:

a) *to edit or convert documents to HTML*

The need for guidance and software to assist in the planning and structure of data was also discussed. (The group spent some time swapping information about tools members used and found a considerable degree of unanimity. It was felt this was typical of the sort of information that needed a structured route to the rest of the community.)

b) *specialist/more advanced tools*

The tools needed are more than simple HTML: there should be libraries of scripts to help run standard services, automated for generating local versions. In particular there needed to be more collaborative work in the areas of graphics and gateways to databases.

c) *quality tools*

eg validation software in a UK location, links verifiers, usage monitoring tools and procedures, etc. The quality tools issue was tied up with the need for more guidance on, for example, the use of meta information for indexing and for maintaining the quality of information.

d) *collaborative tools*

It was felt that the original purpose of the Web should not be lost sight of and there needed to be some exploration of its use as a collaborative platform.

e) *clients, helpers, servers*

And of course there needed to be ongoing review of, guidance on and availability of the key Web tools. Incidentally it was felt that the UK could contribute by offering to develop a better lynx interface.

## **Guidelines**

A lot of the discussion on tools threw up needs for guidelines and guidance. There were many other instances when the experience of sites could beneficially be passed on to others. For example, early in the session the group discussed the problem of computing service involvement in getting web servers up and running and how then to smoothly migrate to distributed management and ownership of them by others, such as academic departments, PR offices, etc. (We got sidetracked, for example, on how far computing services should write the first generation of information or home page in order to stimulate the department to take this on - with the accompanying danger of dated information. But we all agreed that finding a departmental 'champion' was crucial.)

## **Training**

Various members of the group were already running regular courses or workshops for information providers, which cover both the technical and organisational (rules, procedures, regulations etc) aspects. There should be links to these on the proposed Information Providers Web Server (see above).

In addition there should be a nationally funded set of course materials designed in such a way that they could be modified to fit local circumstances.

Educating the decision makers (senior academics and administrators) was seen as a vital activity. The group did not think it possible to produce a national presentation to be used locally, because of the wide range of local conditions, but suggested that good case studies would be helpful and should be included in the Information Providers Web Server.

## **A National Information Strategy**

A comment on the need of external organisations to obtain information across the UK on particular topics or from particular groups led to a discussion around the idea of a national data model: at both a local institutional and a national level we need to start planning an information strategy which identifies who 'owns' data, how it fits together, how to share our meta information presentation so that we can provide integrated searching.

## Management Issues

*Issue:*

We need ongoing access to appropriate legal advice

*Recommendation:*

Establish a mailbase list to provide up to date information and answers to questions

*Issue:*

Many issues come down to institutional policy decisions but it is useful to have information about others' views

*Recommendation:*

Set up an online mechanism for sharing guidelines and codes of practice.

*Issue:*

There is a need for a national source of information and advice on technical issues relating to WWW

*Recommendation:*

Set up a WWW support post to:

- gather information on use in UK HE and disseminate
- provide advice and support on WWW tools
- provide training on setting up and running servers etc.
- establish a register of "official" WWW servers — via a UK Home page?

*Issue:*

We need a contact point for WWW management issues at each institution

*Recommendation:*

each institution should nominate a WWW contact person

*Issue:*

There is a need for information on national/local caching, indexing, access control etc.

*Recommendation:*

Commission a technical study of these areas.

## Technical Issues

### Caching

HENSA cache is not as widely known as it might be (<http://unix.hensa.ac.uk/>). Need to know how to communicate with computer services people to let them know. What mailing lists to use? Is the Web-support list suitable?

Should we use local caches as well as HENSA? Yes, in short term. Longer term national caches and local ones if low bandwidth to site.

How do you decide what to cache? Automatic, depends on expiry times but can be manually affected. CERN server can be configured to check every time - good if pages are frequently updated or can explicitly not cache it. Good documentation with CERN server. With some browsers once you have specified use Proxy, it can be hard to get round. Can manually disable for specific documents. A check box would be better. System manager can use standard UNIX tools to remove cache if short on space.

If there are 5 caches between you, and the document you want is not in any, have to go through all 5 first before going direct. Hit rate is dropping as number of docs on WWW increases, and slowing access time down may cause people to not to use caches. Need national infrastructure with big disks for caching. Need measurements on this.

It might be good if JISC funded a caching strategy - both local and national. Local servers are do-able now, but central servers of chains of servers need investigation. HENSA cache is 2Gig used to have 60% hit rate, but dropping, probable less than 50% now. Disks are relatively cheap - go for 9Gig disk (£3,000) is sensible for local cache with relatively small machine.

Other scenario have multicast - ask does anyone have this file in cache. Problem if all caches answer together - problem with any multicast protocol. Something for the future.

Now - put in place physical and organisational infrastructure. Makes it easier to switch to new strategy.

Brian Kelly's handbook good starting point for new web managers, then look at CERN documentation. Does documentation exist about why it's a good thing to cache, why at your site etc.?

Role for funding on going research?

A cache is not going to be able to cache all transactions, especially CGI scripts which are being increasingly used, eg web crawlers which generate large numbers of hits. As well as general caches, identify the services we want to replicate in the UK - useful, good for bandwidth etc.

Which services to replicate? Lycos and HTML checker spring to mind. Find out about use in general academic community, this group not good sample. Need some kind of monitoring. Can't collect logs of individual machines, but if everyone uses caches can analyse cache log. Useful to have some kind of pattern tracking. Lycos - resource discovery - a necessary resource better candidate for replication. Some stats sampling on backbone traffic to get idea of network hotspots. This is already happening in Kent (JISC project checking response time - not actually where traffic is going).

But this data will change - may show one thing being used, we mirror this, but 3 months later something else is being used instead. Needs to be on going thing, continually reviewing top 10-20 services. Could have questionnaire/CGI script what services do you use - but tend to get low response to this kind of thing, and the people who do reply may not represent the majority of the users. Anything involving human effort is probably not going to work well, may be better to analyse cache logs. Caching does respond to what is being most used now. Works for plain text files, images etc.

Mirroring better for services, but problem keeping them upto date. Given choice caching better than mirroring, but active things, eg search engines, can't be cached therefore have to be mirrored. Could get arrangement to have search software mirrored eg. getting agreement with Lycos to mirror their engine.

Before web, in the UK we tended to use sites like HENSA and src.doc to get docs and files. These services are proven to be useful, and similar resources in the Web is what we need to support now. This type of data doesn't change too often, and lends itself to both mirroring and caching, but mirrors/caches may not be used if other sites in US have better search engines for example.

## Hardware

Imperial said they will limit number of web servers, does this mean everyone needs access to the server machine? No, have home pages on separate file store. Best leave this issues to operational staff, a local issue. Vendors generate lots of hype trying to sell powerful machines. What you need depends on your usage. Universities may want figures to say what kind of machine you need, but can't predict this. What we can do is provide guidelines on how to setup machines so that it will be easier to migrate to more powerful machines. Potentially a useful appendix to Brian Kelly's document.

If we have an infrastructure where caches and web masters are talking to each other then will get lots of support from within this. Want guidelines on server hardware and software from experienced web masters. Abstract concepts rather than how to install version N of software X.

## Indexing

W30 goal is to have server registry. UK HE community should do this. This registry would be useful. If we do have this what servers do you include? Every server or have some quality control? Have 2 registries one with all servers, one 'quality controlled'. Have list of every department in every university with server, and list with all the other servers also.

From user perspective may want to know eg the biology server of a particular university. This may not be possible, for example the biology server may be part of a main univeristy server. What they may want is information based on subject areas. Problem with this is who will administer it? Library community keen on this, subject librarians. Is there a requirement for both approaches (institutional and subject)? Institutional level may already exist in active maps that are available.

This is a valuable resource. May want to make this an official resource.

Why are people looking for this info. Do they want info on an institution or a particular subject. Both. eg sixth former would want to look at subject area first broken down by institution. Lots of ways of looking at it. Need search engine, but maps are also very useful.

Aliweb, gets description of site from someone at site. Good in theory, but may get inappropriate description. Need this info and need someone managing it.

Network managers of official university servers should complete meta data about server and pass it to infrastructure.

Indexing at lower levels. Harvest for example has some facilities for this. This indexing is something librarians and archivists do (but not on the network). Volume problem as well, what to do with all the info if you get it. It is believed they force their users to index - every server has to be registered and part of registration is agreement to index, but how to do this hasn't been decided.

Whois ++ similar template to aliweb has centroids - kind of text index - possible to generate centroids for web servers, then harvest those and use those as key into HE services. This is just research at the moment. Too early for any recommendations.

Danger with recommendations in indexing, that people may use them, and when new development comes along we will need another short term solution. i.e this is an ongoing situation.

What will the user want to do when searching for info. What do current search tools do. eg may want to search for server name, file name, title, something in header limited by domain etc. Now spider type engines do that sort of thing eg WWW Worm - do-able now and useful. Robots have some fundamental problems, but are liked. Worm, web crawler Lycos etc. are private bits of software. Harvest is more public and freely available and can be configured as a robot. Has advantage that you could give it database eg UK database and let it sort through it.

Would be more acceptable to have robot visit your site that was being funded to provide information to community.

Definite need for indexing.

One tool will never do everything, but can recommend a current tool to provide a service. Need somebody funded to monitor what is required and available. Need a single interface that is flexible enough to allow you to use it with different tools/search engines. Want to be able to direct user to a particular centrally funded site that will let them do the search they need - a general search engine that acts on the parameters the user provides. eg. want to be able to say I want to search on this term in these fields, and this then sends it off to the appropriate engine, rather than having to say I want to do this so I will have to use this particular engine.

Users want to search and get results with good signal/noise ratio. Don't want to build up peoples expectations unrealistically. May use librarians to help with searches, very used to refining searches on bibliographic databases. Need to teach people to use the tools correctly, and they will have realistic expectations.

Should be able to allow effective searching within the UK. Need to ask people to add bibliographic data to their document, but can't force them. Publications are now being accepted on the network, say that bibliographic (cataloging etc) data must be included to get the papers to be accepted.

But will still get lots of documents without cataloging info included. Even if do get all this info what do you do with it? Do you leave it distributed or try and collate it all. Technical problems do exist with this. May need to compartmentalise data eg in library wouldn't search for something in all books, only in a specific area. May be best to organise web searches in same way. Say 'this web server contains the following categories of info.'

Can any current search mechanisms address non-HTML documents?. Harvest can do this to some extent. With images may need to provide some text to go with them.

Should information providers be encouraged to put info into their documents? Depends how stable format is - must be general and stable enough. May be too soon to ask for this- wait for URNs etc to stabilise, then ask for these.

URNs are internet drafted and are beginning to be implemented. URC is not stable. URNs are an abstract name, so can move around where object actually is. URN is key into URC which is a collection of attributes and values, but what the attributes and values are has not been established. URCs would be cached, so would probably include some kind of expiry date.

CERN daemon can handle URNs and will cache URCs. (Not standard version of CERN server).

## **Recommendations**

*Now*



- Education about taking advantage of current caching facilities available and getting funding for this.
- Need recognition and support for current caches.
- Need document explaining how to setup local caches.
- Operational expertise documentation

*6 months*

- Setup national cache infrastructure
- Creation of Support group - a group running and coordinating specific services for the whole UKHE community to use. Web equivalents of services like mailbase, HENSA unix/micros archives etc.
- UK searching services/ registration of UK academic servers - coordinated by support group

*Later*

- Chaining caches
- Replication of search engines (HENSA already keep top 100 hundred of things accessed)