THE PROBLEM

Information on the Web is usually available to a global audience that may well have *no* knowledge of SLAC, the institution, much less its environment and mores. Instead of being used by a relatively homogenous set of people who know each other, Web documents may generally be perused by anyone who has access to the Internet from anyplace in the world. There is no SLAC wall around our Web information.

This burgeoning accessibility may lead to mis-interpretation of information (*e.g.*, using logging data as if they were a polished report) or even abuse of information (like planning robberies around staff vacation schedules).

SLAC must address:

1) What information should be made available to the world and what must be kept private--and among whom.

2) How privacy can be maintained where necessary without burdensome procedures that inhibit unnecessarily the distribution of useful information to SLAC, its collaborators, sibling labs, and others in the world.

ANALYSIS

It is important to make information that can be usefully shared available through WWW in a timely fashion to SLAC and its various collaborators.

No one person or group knows all the material that falls into this category. We depend on others around the Lab to supply information from their own areas of expertise. We need to encourage them to contribute what they own to the Web and maintain it. Decentralized authority is the philosophy upon which WWW is built and has resulted so far in its spreading very effectively.

However, we also need to be concerned about issues of privacy. Information may be irrelevant, embarrassing, misunderstood, or even dangerous if read by the wrong audience.

Access to information on WWW may be granted variously. When considering making files available, document owners should initially consider the questions and answers below. These do not limit all relevant inquiry since only the document owners, who know their material best, can form the most appropriate set of questions.

2 methods
- IP address
- non-Kerberized password

1) Should the SLAC information be made available on the Web at all?

   Given current WWW technology, restriction to a particular group at SLAC is impossible to implement securely, so SLAC information that must have such limited access should not be put on the Web at all.

   Enterprise data like personnel, financial, and salary information and drafts of papers in progress generally fall into this category. Serious consideration should also be given before installing preliminary hardware and software evaluations, vacation schedules, and some pager numbers.

2) Should the SLAC information be made available only to people logging in from the SLAC.Stanford.EDU domain?

   Restriction can currently be handled securely by limiting access to users whose Internet addresses match a portion of SLAC's Internet address (IP number string).

   Accelerator Operations logs, *The Interaction Point,* the Stores catalogue, and problem reports like PROBTRAK are common examples of this category.

3) Should the SLAC information be made available only to SLAC.Stanford.EDU folks and collaborators who are not/cannot login to SLAC hosts?

   Given current WWW technology, restriction to a group of collaborators at SLAC and elsewhere is impossible to implement securely. Casual browsing may be discouraged by parsimonious use of passwords for particular files or file hierarchies, but their use creates an additional load on those responsible for Web information and software.

   Some experiment planning and discussion documents fit into this category.

4) Should the SLAC information be made available globally to users of WWW?

   The *Beam Line*, SLAC Pubs, preprints, the "white pages" of the SLAC phone book, and SCS documents for users are common examples of this category.

done partly through "Welcome Page" separation

Ultimately, group leaders or directors or their designates (hereinafter collectively referred to as "group leaders") are responsible for the appropriateness of SLAC information put on the Web. Traditionally, responsibilities like these have been delegated to computing czars.

THE PROPOSAL

WWW Documents: As called for in the "Final Report of VM-Phaseout Committee," Lab management will encourage SLAC information developers to contribute their knowledge to WWW to help restore "the collegial atmosphere which existed in SLAC computing in the early 80's,...." Part of this effort must be to create an institutional culture that *proactively* considers privacy issues along with WWW page creation. Specifically:

1) People making their first information available to WWW must read a short "WWW Information Contributor's Guide" and confer with their group leaders about the appropriateness of their proposed pages, discussing privacy issues explicitly.

   These include questions about the consequences of making their pages globally available. Will publicizing this information hurt someone? Infringe on the safety of the Lab? Betray plans of the Lab before their time? Make the Lab look unprofessional? Disseminate research results prematurely? Present working documents without necessary context? Broadcast incomplete drafts? And is this information appropriate for reading by some but not all people on the Web?

2) For subsequent contributions, the authors must consider privacy issues themselves for each document they add to the Web. If they have any questions, they should discuss problematical points with their group leaders. Page owners shall be held accountable by their group leaders, both for the information they supply to WWW and to whom it is accessible.

3) After the fact, group leaders may cause any information to be removed or access restricted as seems necessary. If time allows, any changes shall be discussed ahead of time with the authors; otherwise, afterwards.

WWW Servers: All Web servers must be authorized by group leaders. Each server should have a strong justification like SLD's (heavy use and specially tailored code for Oracle access). Web servers can present serious security issues if incorrectly configured.

SCS will provide a document on risks, standards, conventions, procedures, and software for installing and maintaining WWW servers. Before setting a server up, the owner must register centrally each one. The exact procedure is still to be determined but will include assignment of a unique name according to standard conventions.

IMPLEMENTATION

To implement the proposal the following steps should be taken:

1) Develop a common view among Lab management, WWW czars, and other responsible parties at SLAC on the types of privacy restrictions we need for what kinds of information.

2) Document the consensus in a "WWW Information Contributor's Guide" that treats privacy considerations along with other standards, conventions, procedures, and software for creating pages for the SLAC Web.

3) Publicize the model with a presentation to the key managers, a UNIX Journal Club seminar, an article in *The Interaction Point,* discussions at the (anticipated) Web Users' Group, *etc.* Make the document available through WWW itself.

CONCLUSION

This proposal tries to strike a balance between supplying WWW pages in a timely and minimally onerous fashion, while meeting SLAC's needs for the privacy of a portion of the information it puts on the Web. The proposal is patterned partly after the successful way U-disk file contributions were handled in VM at SLAC.

The model is a collaborative work in progress. As the technology improves in server code, tools, *etc.*, it may well prove beneficial to iterate the model to take advantage of these advances.

SLAC Director's Office

ALL HANDS MEMO

DATE: January 10, 1996

FROM: Burton Richter

SUBJECT: Use of SLAC Information Resources

Working at SLAC we have certain responsibilities imposed on us by SLAC, Stanford University, and the U. S. Department of Energy regarding the use of information resources. As Director of the Laboratory, I want to remind everyone of these responsibiliies and the serious consequences to us as indviduals if we do not adhere to the regulations which govern us.

SLAC information resources are government property and, as such, are subject to "appropriate use" requirements found in federal law and the provisions of the SLAC contract pertaining to the proper use, protection, accountability and disposition of government property. These laws and policies apply to all data-communication and telecommunication facilities and services (including, but not limited to, e-mail, telephones, voice mail, faxes, personal computers, workstations, networking services, mainframes, minicomputers, and associated peripherals and software), whether for administration, research, teaching or other purposes.

Any member of the SLAC community who, without authorization, threatens the access and sharing of information is engaging in unethical and unacceptable conduct. Such unethical conduct includes destroying, altering, dismantling or damaging SLAC information technologies, properties, or facilities; or interfering with access or use of these SLAC resources.

The following is an outline of some of the responsibilities of those who use information resources and is intended to ensure that users: (1) respect the integrity of data-communications and telecommunications-based information resources; (2) respect software copyrights and licenses; (3) use these instrumentalities and resources only in support of their official duties; and (4) respect the privacy of other users.

A WWW Security page is under development (http://www.slac.stanford.edu/comp/security /security.html), which will contain relevant new information as it develops.

General

- Suspected misuse of SLAC computing facilities must be reported at once. This is an affirmative duty. The SCS Help Desk (Ext. HELP [4357]) can refer you to the proper authority for making such a report.

- SLAC information resources and networks -- including browsing the World Wide Web (WWW) -- may only be used for work related to SLAC business.

- SLAC is an open environment as befits a research institution. In this type of environment it is appropriate for users to share files and have access to data or tools belonging to another user in order to accomplish one's work. For this reason, normally only files which contain sensitive data and need to be secure are read protected. For information concerning the protection of such data from unauthorized use, contact the SCS Help Desk (Ext. HELP [4357]).

Computer Hardware and Software

- SLAC computer accounts are intended for use only by the individual assigned to that account. Each account holder is responsible for the resources used by that account and for taking necessary precautions to prevent others from using the account. Users shall not seek to gain unauthorized access to information resources or to enable unauthorized access.

- Passwords must be chosen with care and not divulged to anyone. They must be at least six characters in length and should not be names or dictionary words that can be easily guessed. Some words are prohibited by the system because studies have shown that such words do not conform to this standard. If one suspects that a password has been compromised, it should be changed immediately. In any event, general users' passwords should be changed at least once every nine months and privileged users' passwords should be changed at least once every six months.

- Computer users shall not interfere with the intended use of SLAC's information resources or without authorization: destroy, alter, dismantle, disfigure, prevent rightful access to, or otherwise interfere with the integrity of computer-based information and/or information resources.

- Unauthorized copying of proprietary software is strictly prohibited.

Electronic Mail

- Since electronic communications (e-mail, news groups, contributions, etc.) pertain to SLAC business, their content reflects on SLAC's image. Therefore, it is important that such items be professional and not personal in nature. Use of SLAC's electronic communication facilities to send fraudulent, harassing, obscene, threatening, racial, sexual or other unlawful messages is prohibited and illegal, as is use of SLAC information resources for lobbying of any kind.

- Electronic communication files such as e-mail are not intended for general dissemination. Unauthorized perusal of such files is not permitted. Conversely, sensitive data should not be sent as e-mail because no means of transmission available at SLAC is completely secure. Users need to be aware that e-mail is not private or confidential, that it may be monitored at any time, and that it may be discoverable in a legal proceeding.

- *SLAC's contract with the government makes virtually all information on SLAC computer systems, including e-mail, available to the government. It is not SLAC policy to routinely monitor e-mail. However, SLAC cannot and does not guarantee the privacy of e-mail communications.*

SLAC requires that members of its community act in accordance with these responsibilities, SLAC policy, University policy, and relevant laws and contractual obligations. In order to assure all relevant parties that no misuse of resources occurs, SLAC reserves the right to sample file contents at any time. Improper use of computing facilities may lead to disciplinary actions up to and including termination and/or legal action.

WJ/pa/bjf

S L A C M E M O R A N D U M 21 Sep 1994

To: Appendix to WWW Wizards report to C. Dickens

From: WWW Wizards Committee (ADDIS)

Subject: Brief background info on Web at SLAC

HISTORY:

1. The World-Wide-Web (WWW) was developed at CERN by a group led by Tim Berners-Lee. Paul Kunz spotted its potential for providing easy access to the SPIRES databases during one of his trips to CERN in 1991. He and Terry Hung immediately installed a browser and server on a NeXT machine and also on SLACVM and called it to the attention of the library. Subsequently George Crane wrote a WWW interface to SPIRES which was quickly added to the then CERN front page. This early link to SPIRES was influential in quickly increasing the use of the Web in the particle physics community. (Currently, searches to SPIRES-HEP number over 70,000/month and is still the most used portion of the Web at SLAC.)

2. A ad-hoc group calling themselves the WWWizards started meeting late in 1991 under the sponsorship of the library and more informally, SCS and later SLD. The Wizards have been a totally volunteer effort and have considered themselves a 'working group' rather than a committee, meeting fairly regularly each fortnight.

MILESTONES - several events were important not only to the Web at SLAC but in laying the foundations for the explosion of Web use which we have seen in the last year.

1. Tony Johnson of SLD developed a full-screen X browser which he called MidasWWW. (The first X browser had been a very handsome but buggy 'Viola' whose developer, Pei, had gone on to other projects.) MidasWWW could display all sorts of graphics, including compressed Postscript. This opened the way for the SLAC Library to start providing postscript versions of the electronic bulletin board papers actually linked to records in the SPIRES-HEP database. (As of Sept 1994, over 9000 full-text preprints were available to the particle physics community via WWW-SPIRES.)

2. Tonyj received a development grant from the ill-fated Texas Research Comm. for further development of MidasWWW in connection with SSCL experiments.

The development of XMosaic followed in the footsteps of MidasWWW incorporated many of Tonyj's innovative ideas into the family of Mosaic browsers supported by the National Center for Supercomputing at Urbana.

3. Addis, Tonyj, and others visited Paul Ginsparg at LANL, the developer of the highly successful physics preprint bulletin boards (electronic preprint archives) and introduced him to the potential of the Web as a means of providing access to the xxx.lanl.gov server. Tonyj installed MidasWWW on a Sun workstation in Ginsparg's office and subsequently, the Web server at XXX has become one of the busiest in the Particle Physics Community.

4. Bebo White put in place servers on other platforms, installed a Gopher server for linkage to WWW, and in so far as time permitted, started to maintain a number of the pieces of the Web at SLAC. Bebo also was invited to CERN for a month to work directly with the Web group there.

5. Winters put in place an 'official' SLAC home page (in reality a complex of pages) and a mechanism for handling the exacting process of organizing, maintaining and changing this material in an orderly, timely fashion. Since that time, almost 300 significant changes have been made to the home page in response to expressed needs of the SLAC community. Over 685 hypertext links must be maintained to insure the integrity of the home page complex.

6. Several Web Wizards give talks and demonstrations (Addis, Johnson, Kreitz, White, and Winters) to various conferences as well as SLAC groups Currently, talks about the SLAC/SPIRES/WWW full text preprint system are scheduled for Am. Chem Soc., Am. Phys. Soc. and Astron.Astrophys. Meetings in the coming months.

7. The Committee on Future Computing at SLAC suggests WWW as a lingua franca for tying together the SLAC community on diverse computing platforms.

PRESENT STATE OF THE WEB AT SLAC:

At this time, Sept 1994, the Web looks as follows at SLAC:

1. SERVERS -At least five and probably more 'servers' (needed by information providers only) are running at SLAC on various platforms:

   a. SLACVM - The original CERN VM server, now an antique, is still the main server for the SLAC 'Home Page' and by necessity, all the SPIRES searching via WWW.

Statistics: currently averages 6300 requests/day of which about 50% are direct SPIRES searches.

Comments: This server was the source of pride in the early days of Web when SLAC was a leader and innovator in providing various services. Now, this server is an embarrassment since it is outdated code which doesn't work with many of the standard browsers (i.e. Mac-Mosaic). The moratorium on SLACVM development has meant that we have fallen further and further behind other facilities. Many places provide links to our SPIRES-HEP searching and several have started issuing warnings to their users about the out-datedness of this server, with instructions for special steps which users must take to successfully use it.

b. Unix - Modern servers are running on several Unix machines.

Statistics: None

c. VMS - SLD runs a server on SLACVX to handle their special needs for shift scheduling and other projects.

Statistics: ?

d. NEXT - Two CERN servers (one latest version) run on LIBNEXT in the Reason cluster (maintained by George Irwin). The intention is to run the latest CERN server at both ports 80 and 5080 to serve the full-text postscript preprint files.

Statistics: ~650 document requests/day

a. Browsers:

FUTURES:

Areas needing attention:

1. Technical - most urgent. support issues

2. Home page issues
   Policy issues, Winters may want advisory committee???
   Need glitzy version of page (Rene Donaldson?) for 'public' access

3. Security and privacy

4. User education, i.e. WWW User group

Fermi National Accelerator Laboratory

FERMILAB-Conf-95/056

Collaborating Over the Web: Libraries and Laboratories or The Librarian and the Webmaster

Paula Garrett

Fisher Library, University of Sydney
Sydney, Australia

David J. Ritchie

Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois 60510

March 1995

Submitted to *AusWeb95, the first Australian WWW Conference,*
Ballina, New South Wales, Australia, April 29-May 2, 1995

Operated by Universities Research Association Inc. under Contract No. DE-AC02-76CHO3000 with the United States Department of Energy

Collaborating over the Web: Libraries and Laboratories*

or

The Librarian and the Webmaster

Paula Garrett[1], Fisher Library, University of Sydney, NSW 2006, Sydney, Australia.
Phone +61 2 351 3257 Fax: +61 2 351 2890

Email: p.garrett@library.usyd.edu.au

David J. Ritchie[2], Fermi National Accelerator Laboratory, Batavia, Illinois 60510 USA.
Phone 1 708 840 3940 Fax: 1 708 840 2783

Email: ritchie@fnal.gov

Paper submitted to AusWeb95, the first Australian WWW conference.
Ballina, New South Wales, Australia, April 29 - May 2, 1995.

Keywords: World Wide Web, High Energy Physics, Collaboration

Introduction

Fermilab [HREF1] has been using World Wide Web [HREF2] (WWW) servers and browsers for the past several years. The purpose has been to allow information sharing within internationally dispersed teams of large collaborations of high energy physics experimenters and to make possible the easy dissemination of information by support groups who work with those collaborations.[3] That is, the purpose has been in fact that for which WWW was created at CERN [HREF3] beginning in March 1989. See (Berners-Lee *et. al.*, 1992 and 1994).

In early 1992, Ruth Pordes and Jonathan Streets of the Fermilab Computing Division's [HREF4] Online Systems Department [HREF5] (OLS) were considering the problem of providing information about online data acquisition systems to high energy physics experimenters. Seeing the WWW presentation to Artificial Intelligence in High Energy Physics (IHEP'92) at La Londe, France in February 1992, Streets recommended WWW as being "the best thing around," and OLS decided to adopt it.[4]

A server to provide online data acquisition system documents [HREF6] was set up by Tim Berners-Lee [HREF7], CERN , and Jonathan Streets, Fermilab, on a visit by Berners-Lee to Fermilab in July 1992. Tim Berners-Lee went on to visit NCSA [HREF8], the National Center for Supercomputing Applications in Urbana, Illinois beginning the collaboration whose result was NCSA Mosaic for Mac [HREF9], PC [HREF10], and X (Unix) [HREF11].

That first server ran on a VAX/VMS system and used DCL (Digital Command Language) in its implementation. In response to browser requests, it extracted full text documents from a previously existing Fermilab-written document database[5]. This 1992 technology would be considered fairly primitive by today's standards. Using a revised implementation, the server continues in operation today.

July 1992 through January 1994 saw the creation of an estimated ten to fifteen servers by various groups associated with Fermilab. Some, such as E781 [HREF12], were created by experimenters to help organize the dispersal of information within their far-flung collaborations; others, such as the Sloan Digital Sky Survey [HREF13] were created by Fermilab support groups to help with the development of software and hardware intended for a particular project.

In February 1994, it became clear that WWW was gaining a real presence at the laboratory and that an actual Fermilab Home Page [HREF14] was needed to bring together the servers created by the "early adopters" as well as set the stage for further use of the web technology. Initially, a process involving consultation with a broad set of relevant groups was imagined. However, the impending announcement of evidence for the top quark [HREF15] made it imperative to have the home page up quickly[6]. Instead, a small team drawn from the Fermilab Directorate and Computing Division was charged to accomplish the task. Beginning with a concept authored by Judy Jackson, Fermilab Directorate, and Liz Quigg, Fermilab Computing Division, an implementation was put in place by the date of the announcement on April 28, 1994. The Laboratory continues to add additional links to its

home page and to update it with significant announcements–the most notable one being the announcement of the discovery of the top quark [HREF16] posted on March 2, 1995.

Just prior to the April 1994 date, the Fermilab Library, knowing the intense interest of its readers in the scientific papers concerning the evidence of the CDF [HREF17] and D0 [HREF18] collaboration, set up a Fermilab Library web server [HREF19] in order to provide references to the top quark papers. The Library continues to extend and enhance its server with reference offerings as well as information about the Library.

In the months since, the Laboratory and Library servers have evolved in distinctly different ways. An effective collaboration between the librarian and the webmaster has occurred. Some information has been listed on both servers. In other cases, the choice has been made not to list the item on one or the other. A certain intentional complementarity and dissonance of approach has evolved. Looking back, we have come to realize that this process has been valuable and stimulating to the evolution of both servers and may hold a key to the way other organizations may wish to orchestrate their web offerings.

In the remainder of this paper, we describe the architecture of the Laboratory and Library servers, discuss the goals, the implications of those goals, the drivers of structure and content, and the effect of these factors on the architecture. Comparisons are made. The implications for Libraries within organizations are described. The application to large Libraries which begin to have multiple servers within their own organization (e.g., one in reference, one in cataloging, one in circulation) is made. The broader implications are discussed.

Laboratory Web Architecture

The laboratory web architecture had to be created in a way that succinctly grandfathered the existing servers. The individuals who created these servers were actively using the web to further significant national and international collaborations, such as the Sloan Digital Sky Survey [HREF20] and the DART [HREF21] (Data Acquisition for fixed target Real Time) projects. In addition, they possessed much of the knowledge and lore concerning the server and browser software. Thus, whatever the "architectural rules," the rules had to let these collaborations go forward and not require wholesale reorganizations of servers.

Goals and Implications

There were a number of goals in creating a laboratory web architecture and in setting up the Fermilab home page. Years of experience with the laboratory were very helpful in suggesting matters to consider.

Goals for the overall organization were:

1. Whatever the scheme, it was crucial that it be easily adaptable to new situations.
2. It had to work in a decentralized fashion (basically in order to achieve Goal 1); yet it needed to provide laboratory management with the necessary control.
3. The decision processes involved in listing new servers had to be quick in the ordinary "same as three previous servers" situations but allow for careful consideration in unusual, totally new types of server listing requests.

Goals for the home page itself were:

1. As the laboratory's "electronic front door," the home page should be presentable and "spiffy".
2. The home page should provide information appropriate for an audience with a wide and varied background. It should do this without impeding the use of WWW for the scientific, support, and technical collaboration purposes for which it was invented.
3. The home page should point to other Fermilab WWW servers for the benefit of outsiders.
4. The home page should point to a limited number of non-Fermilab servers of high interest to Fermilab insiders.

Some implications of these goals were:

1. A high performance computer platform with high uptime was chosen. Login access was restricted to maintainers only. The existing net news server turned out to be the appropriate choice.

2. An effort was made to structure the home page so that the "tourist" sections about Science of High Energy Physics [HREF22] and the Laboratory as a Research Institution [HREF23] would draw the visitor naturally into material appropriate for them while researchers and technical personnel could quickly move past these sections into more technical material (in the Fermilab At Work [HREF24] section).

3. The pages, particularly in the Fermilab At Work [HREF25] section, have evolved into something of an information directory. (This is in some ways analogous to the "community information" catalogs that are often provided by libraries along with their automation systems.)

4. The list of pointers to high interest, non-Fermilab servers has been limited to information very focused on the laboratory's mission. It has not become a "meta-page," one with extensive and continually growing lists of references to other servers. Instead, the referenced servers are those that contain information associated with the laboratory's mission in a very mainstream way.

Content and Structural Drivers

There were a number of drivers of content and structure in the construction of the Fermilab Home Page during February - April 1994.

The role of the Laboratory Directorate had to be recognized:

The role of the Laboratory Directorate in shaping the presentation of the laboratory to the external community had to be recognized. That this should be a consideration with respect to World Wide Web seems obvious now because of the web's clear public presence.

A year before the setting up of the Fermilab Home Page, however, the web was simply a device for technical communications between an experiment's collaborators[7] with little, if any, visibility to the external community. The growing public presence of the web meant that creators of web servers had to accommodate themselves to the Directorate's beginning to play a role and the Directorate had to educate itself about the importance of the web as a significant vehicle for presentation of the laboratory to the public.

Result: An individual within the Directorate who already had certain responsibilities for media relations was designated as Editor of the top level page and those pages designed for the general public. An individual from the Computing Division was designated Editor of the technical pages– these being Fermilab At Work [HREF26], and its related links.

A relatively flexible procedure involving e-mail consultation between the two Editors and the webmaster was set up to coordinate proposals for changing the home page. See the United States Department of Energy Televideo Service [HREF27] home page for a *much* more structured approach in this area.

The role of the web in the life of experiments and their support had to be recognized:

By April 1994, the web was inextricably linked into the way a number of experiments conducted their collaborative activities. Experiments had clearly found the web to be an effective communication tool to further the research process. Support groups came to similar conclusions. It was very important to foster this use of the web–after all, not only was the support of collaborative high energy physics research the reason the web was invented but it was (and is) one of the main purposes underlying Fermilab.

Experiment web pages truly describe "work in progress"–unpublished, unrefereed material inappropriate to be cited outside the collaboration. Support groups similarly use the web for provisional matters.

The concept of a "work in progress" area became clear. What made this concept particularly useful was that "work in progress" was a sensible disclaimer generally understood by many–not only individuals involved in research but also those associated with funding agencies supporting that research.

Result: The Fermilab At Work [HREF28] section assumed a prominent role in the web offering.

Fermilab is very much a "laboratory without walls" and has operated in this fashion since its beginning in 1967. Those who have contemplated the "library without walls" concept will understand the unique challenges such a structure brings. A member of the Fermilab research community may be anywhere–on the laboratory site, at the college just down the street or at a university in Australia.

It would have been impractical and administratively very costly to restrict access to the "work in progress" collection of pages at an overall level. Indeed, it would have had a chilling effect on collaborative work. Restriction by browser IP address would not have worked because of varied individual researcher affinities. Usernames and passwords would not have worked because of the dynamic way individuals move in and out of collaborations.

Collaborations that wish to isolate the more sensitive areas, such as minutes of collaboration meetings, particularly controversial results, etc., *do* make use of access restrictions. It is easy for specific collaborations to employ these techniques because they, with their knowledge of the collaboration, are able to manage access much more easily.

The role of the home page in presenting information to the public had to be recognized:

In April 1994, the web was rapidly taking on a public presence. Fermilab's major funding agency, the United States Department of Energy [HREF29] was setting up a home page. The White House [HREF30] was rumored to be coming up with one.

Result: As with any public information channel, it was important to have appropriate structures set up to allow the laboratory speak clearly and with one voice. The designation of the editor, based out of the Directorate, addressed that matter. By design, the "public voice" of the laboratory could be presented in the top level page and the pages designed for the general public while work in progress could be described in the Fermilab At Work [HREF31] section.

The relationship between the home page and technical publication had to be clarified:

The status of information presented on the home page (and related pages) had to be clarified with respect to whether or not its presentation constituted technical publication.

Result: The matter was deferred. No area on the home page was *formally* designated as a publication area.

There are clearly documents within the Fermilab At Work [HREF32] section that are nearing publication. However, they are regarded as either (a) work in progress, (b) in support of that work, or (c) material presented in the spirit of inviting "community refereeing," additional collaborative work or co-development. A particularly good example of the latter is the Fermilab Software Tools Program [HREF33].

To address the deferral, a collaboration between the Fermilab Publications Office, Library, and Computing Division was set up to work through the issues. A number of items (varied input formats, the need for patent and copyright review as well as peer review) make this a complex task. Once accomplished, a Publications Fileserver will be put in place. Authors will be able to submit their papers and readers will be able to obtain publications–all over the web.

The varied view points of those browsing home page had to be addressed:

Result: Within the top level page and the pages designed for the general public, this was addressed by including material at different levels of technical sophistication.

Within the Fermilab At Work [HREF34] page and related pages, the approach of providing different access points was followed–borrowing here from library organizational techniques.

Many people simply want to access what is new. Others know the laboratory organization chart and wish to find their material via that route. Others still know the name of the experiment or large scale project. Others simply know the general category of information. Finally, others want a very focused guide to external web resources with some "jumping-off" points to broader "meta-page" guides.

All these access methods are provided but without elaborate page structure. Duplication under each access method, if it serves the reader, is allowed. Since screen real estate is at a premium and short load time is very desirable (this is, after all, a list for those *at work*.), simplicity is the rule: the structure is that of a simple list indented as appropriate.

Effect of Drivers

Fermilab top level page and immediate subsidiary pages:

The effects were:

1. The home page tended to be more static and less easy to change.
2. There were more severe strictures on form and content of the top level and second level pages.
3. There was more sign-off required to make a change to any of these pages.

These were intentional and desired. A technology which achieves prominence as a communications channel to a large general audience needs to be managed carefully and in a fashion that is consistent with the organization's overall direction. In addition, as people have come to depend upon the pages for information as part of their regular everyday activities, stability and familiarity have become important–particularly at the top page levels.

Collaboration and Support Group pages:

The main effect was that collaborations and support groups were asked to designate a "local webmaster" as a prerequisite to having their web server listed as a link off the Fermilab home page (e.g., in the Fermilab Experiments & Projects portion of the Fermilab At Work [HREF35] page).

This "local webmaster" is expected to be a contact with whom the Fermilab webmaster could discuss technical issues (e.g., server down, change of server node, etc.). The local webmaster is also expected to make sure that the server content is consistent with various proper computing usage policies. A distinction is made between collaboration/support-group "institutional" pages and home pages for individuals. The "local webmaster" is held responsible for the content of the "institutional" pages. The individuals are held to conformance with proper computing usage policies in the content of their home pages. When setting up the individual's home page, the local webmaster is expected to remind individuals about these policies which they have seen before at the time they received their computer account.

An element of "line management responsibility" was put in place by requiring the local webmaster's supervisor (or official spokesperson in the case of collaborations) sign a form designating the particular individual as local webmaster. In addition, a supervisor is strongly encouraged to obtain the blessing of their Division or Section head in order to make sure there is a clear understanding at a high level about what is being done.

Consequences:

With respect to the top level pages, the consequences have been favorable. The structure has held up over the year since the Fermilab Home Page was created and seems to have served its purposes well.

With respect to the collaboration/support group pages and servers, there has been a modest amount of comment about the sign-off bureaucracy. This has been weathered. The delegation of responsibility for content to an accountable, recognizable line management chain has given management (and certainly the webmaster) a certain comfort level in linking independent servers to the Fermilab Home Page.

At least in one circumstance, the arrangement worked well when a particular content was questioned. Thus, the procedure appears to work and, most importantly, the collaborative use of the web and web servers at Fermilab appears to be healthy, on-going, and adopted in greater and greater numbers.

Library Web Architecture

Overview:

With a staff of six the Fermilab Library is responsible for the information and research needs of some 700 scientists and another 2,000 laboratory employees. There is a collection of 15,000 book volumes, subscriptions to 250 journals, and a weekly receipt of approximately 200 preprints. Preprints, research papers which are circulated among the high energy physics community worldwide, are the most important part of the Library's collection. Preprints in paper format are diminishing. They were one of the first sources to move to full-text online format.

When a library's major collection changes formats, focus on access must shift. The typical Fermilab Library reader is highly literate in electronic media and in turn has high expectations in regard to information access and delivery. At the same time preprints were migrating to such places as the xxx e-Print archive [HREF36] at Los Alamos National Laboratory [HREF37] and to any of a number of other preprint bulletin boards [HREF38] as well as onto WWW servers of specific research individuals and groups, more and more information sources were also becoming available via the web. Through its own single web server [HREF39], the Library was able to direct its readers not only to WWW resources, but also to gopher, ftp and telnet sites. WWW has become a common denominator in this high energy physics setting as a way of making active references to information sources.

Goals and their Implications

Immediacy of access to information resources:

Since many documents of interest to the high energy physics community are now being generated on web servers, web access is often the quickest and most efficient means of retrieval. The choice of a web server over a gopher server or ftp access was based on the following:

1. the natural affinity between the web and high energy physics research
2. the ease of server implementation
3. the growing web infrastructure in the Laboratory
4. the flexibility of the URL structure allowing for connections not only to other WWW servers, but also to gopher, telnet, ftp sites.

Relevancy, both in content and format, to reader and laboratory information needs:

In order to achieve and maintain relevancy to reader needs the Library's web offerings are fluid in design and change according to reader demand, both actual and anticipated.

With its own server, the Library is free to make links, change links, break links, and change the design of the hierarchy of resources–all in conjunction with the ever-changing reader needs. As the maintainer of the Library's web server, the head librarian incorporates new web resources discovered by means of listservs, library literature, library and laboratory colleagues. Most of these are placed under the "Hot Links" [HREF40] section of the Library's home page.

Centralization and marketing of Library resources:

The Library home page brings access to diverse resources together in one place. It is a "clickable" library brochure which, moreover, can always be current. Since automating the Library's catalog in 1990, the idea had been to customize the automation system to be the platform for jumping off to other databases and resources.[8] The Library home page has in fact become that "jumping off" point. The automation system, the principal repository of bibliographic and holdings information, is now coming to be thought of as just another information source among those provided by the Library. The web is what is bringing them all together.

Empowerment of the Library staff to author their own web information sources:

The Library server was installed on a Macintosh already in place in the Library. This was familiar territory for the librarian which increased the probability of the server's success. In addition, the experience the librarian gained in writing HTML documents and navigating the web made it possible for the Library to become one of the organizations in the Laboratory that could provide information, guidance, and mentoring on WWW, HTML, and URL's. Most of the Library's web offerings are links to other servers rather than original documents, but this of course reflects the referral nature of library services.

Content and Structural Drivers

Reader Interest:

Distinctive from the Laboratory web server, the factors driving the content and structure of the Library server are much more closely tied to reader interests and needs. Setting up of the Library web server in April 1994 was motivated by reader need to gain access to two preprints which presented evidence of the existence of the top quark, the (then) last, remaining, yet to be discovered quark in the Standard Model theory linking all sub-atomic constituents.

Links to the full-text of these two preprints were the first ones the librarian added to the Library server's "Hot Links." Also handling most of the reference queries, the librarian increasingly turned to WWW first in tracking down government documents such as Presidential press releases and science policy statements requested by the Laboratory's Director and other readers. Good URL "finds" generated from this work were often times incorporated into the Library's "Hot Links" [HREF41] section.

Organizing the web:

When found on the web, links were quickly incorporated into the Library's server often "bumping down" or replacing older links. For example, links to the "top quark evidence" preprints were moved down the Hot Links list as reader interest waned. Now those first links have been removed altogether and replaced by links to the top quark discovery papers [HREF42] by the CDF and D0 collaborations. Again, when significant documents concerning governmental support of science were released, links were added. Two examples here are the U.S. White House report, "Science in the National Interest," [HREF43] and the Drell Report [HREF44] issued by the U.S. Department of Energy on the future of high energy physics. As these documents became less topical, links were removed.

The frequency of the changes is an element of the chosen architecture for the Library's web server: reader-driven change.

Hardware/Software Availability:

As previously noted, the Library server runs on existing Mac hardware. The MacHTTP server and MacMosaic browser software was downloaded from the National Center for Supercomputing Applications [HREF45] (NCSA) in Urbana, Illinois. The easy installation obviated lengthy searches for staff resources that would have been necessary had the software been more difficult to install.

Implications for Libraries and Library Administrators

Involvement of the Library Staff in the Web Architecture

Make Sure there is an Understanding of Basic Concepts:

With a conceptual understanding of the fundamentals of web architecture, library staff can build the confidence needed to maximize their use and in turn the readers' use of web sources. Involvement in the initial period of setting up web server and browser, as well as many other technological tools, provides a sense of "owning" that technology which often motivates future participation in such projects.

Encourage Interaction between Librarian Staff and Computing Professionals:

In the spirit of collaboration, interaction between library staff and computer professionals can foster exchange of ideas and mutual professional respect. The potential for a successful web service is heightened when technology is combined with understandings of usage patterns and needed content. The library staff know their web users–their readers–and their ways of getting from one topic to another. The library staff need to participate in designing the links and other architecture for the web server so that readers will find the linkages natural.

Support the Provision of Continuous Training and Access to Developments in Technology:

Support from library administrators of continuous training programs for library staff along with the provision of hardware, software, network upgrades, etc., is imperative. Library staff will have to have the tools, the training, and the time to experiment with expanding technology and electronic information sources.

New Face of Library Services

Library services have a new face brought on by the new technological capabilities. The pace of development is so fast that one must plan while implementing and implement while planning.

The focus in this paper has been Reference Services but changes are occuring in other service areas as well. It is entirely possible to imagine multiple web servers within a large library–some devoted to circulation (e.g., to support reader access to his/her own circulation records), technical services (e.g., to provide information on recently cataloged materials), and reference (e.g., to provide access to "Hot Topics" pages maintained by individual reference specialists for their particular specialty fields). Indeed, a number of libraries already are making innovative uses of web technology along the lines described.

See for example, Innovative Internet Applications in Libraries [HREF46] maintained by Ken Middleton of Todd Library, Middle Tennessee State University and also Fisher Library at the University of Sydney [HREF47].

When a library begins to have multiple servers with multiple collections of pages maintained independently by varied individuals across the institution, the web organization issues begin to resemble those for a large laboratory, such as Fermilab. It may be that the solutions described here will also apply.

New Role for Libraries within Institutions and Communities

The work of librarians as information specialists has suddenly become popular within our culture. However, the general populace does not necessarily associate librarians with the rise of electronic information. The need for consciousness raising about the library profession has never been greater. The doors are open to librarians to step in as their institution's or community's mentors for Internet access and tools such as WWW.

The ease of gathering and authoring via the web empowers librarians with a greater independence and flexibility. As a result, they will be able to change their library's public interface at a pace more in step with reader needs instead of being so highly dependent on vendors, computer professionals, or others.

The Next Phase of Library Automation

With web technology, systems centered around a single vendor or single information provider give way to a reader-centered collection of many systems and information resources accessible over the network. The unique aspect about the web development has been the *ease* of access along with the *quickness* of switching from one information source to another.

The web's "Uniform Resource Locator" (URL) concept is key to making this ability to switch quickly possible as is, of course, the underlying presence of that uniformly accessible, totally interconnected, network of networks, which is the Internet. Together, the two provide a seamless interface to these varied information resources, bringing on the next phase of library automation by creating what one might call the information marketplace.

Of course, the information marketplace has existed for a long time–the pivotal aspect of this new information marketplace is that the point of purchase has changed. Formerly, the decision to go with this or that commerical database vendor was likely made in the library administrator's office once a year at contract time. Now, the decision is made by the individual reader as he or she chooses this or that network information resource on possibly a minute to minute basis.

Currently, in the "spirit of the internet," there are many "vendors" supplying information for free. Given the ease of authoring and making documents available on the web, this may continue for some time. Still, the demand for authoritative documents or those with other special characteristics accessible in the same way as the free material is giving rise to vendors who provide access to information resources on the network for a price.

Vendors naturally work to have their resources made essential and placed at the center. Special browsers, special servers, or other software that use the underlying web and network transmission rules are all possible. Regardless of what vendors devise for providing access to their information over and above the basic Mosaic browser kinds of capabilities, libraries must require conformance to standards and mandate interoperability (e.g., require that information servers from various vendors work with a wide selection of information browsers).

Not only will this force competition into the information marketplace, encouraging high quality and low price, but it will also allow the library to stay "light on its feet" and very adaptable to the highly diverse, growing numbers of information resources on the Internet.

By so doing, libraries will insure that they, in the next phase of automation (which in fact is now occuring), can continue to be reader-centered rather than system-centered and thereby be of maximum benefit to their reader communities.

Broader Implications

The web architecture within an organization is shaped by many different drivers. The institution's top-level web server is likely to be very structured while the library's web structure, particularly in the reference area, needs to be flexible, dynamic, and responsive to reader interest. The democratization of information authoring and access which the web supports makes it possible to bring many more people into the process of information distribution and access.

Conclusion

The advent of World Wide Web has brought together diverse, formerly distant parts of the Laboratory. Suddenly, press officers, publications specialists, patent lawyers, librarians, and researchers are finding the focus of their work to be just a mouse click away from one another. In collaboration, the Library and the Laboratory have devised a web architecture which makes information accessible to readers while respecting the policies of funding agencies, the sensibilities of authors, and the concerns of other information stakeholders.

References

T. Berners-Lee, R. Cailliau, "World Wide Web," *Proceedings of 10th Conference Computing in High Energy Physics*, Annecy, France, September 21-September 25 1992, 69.

T. Berners-Lee, R. Cailliau, A. Luotonen, H. Frystyk Nielsen, A. Secret, "The World-Wide Web," *Communications of the ACM* **37**, (1994) 76.

Hypertext References

HREF1 http://www.fnal.gov - Fermilab

HREF2 http://www.w3.org/hypertext/WWW/TheProject - WWW

HREF3 http://www.cern.ch - CERN

HREF4 http://fnnews.fnal.gov/cd/ - Computing Division

HREF5 Online Systems Department - http://www-ols.fnal.gov:8000

HREF6 http://fnala.fnal.gov:8000/docdb - Server to provide online data acquisition system documents

HREF7 http://www.w3.org/hypertext/WWW/People/Berners-Lee-Bio.html - Tim Berners-Lee

HREF8 http://www.ncsa.uiuc.edu/ - NCSA

HREF9 http://www.ncsa.uiuc.edu/SDG/Software/Brochure/MacSoftDesc.html#MacMosaic - NCSA MacMosaic

HREF10 http://www.ncsa.uiuc.edu/SDG/Software/Brochure/PCSoftDesc.html#WinMosaicNCSA - WinMosaic

HREF11 http://www.ncsa.uiuc.edu/SDG/Software/Brochure/UNIXSoftDesc.html#XMosaic - NCSA XMosaic

HREF12 http://fn781a.fnal.gov - E781

HREF13 http://www-sdss.fnal.gov:8000/ - The Sloan Digital Sky Survey

HREF14 http://www.fnal.gov - Fermilab

HREF15 http://www.fnal.gov/top_news_release94.html - Announcement of evidence for the top quark

HREF16 http://www.fnal.gov/top95/top_news_release.html - Announcement of evidence for the top quark

HREF17 http://www-cdf.fnal.gov - CDF

HREF18 http://d0sgi0.fnal.gov - D0

HREF19 http://libmc1.fnal.gov - Fermilab Library web server

HREF20 http://www-sdss.fnal.gov:8000/ - The Sloan Digital Sky Survey

HREF21 http://www-dart.fnal.gov:8000/ - DART

HREF22 http://www.fnal.gov/hep_descript.html - Science of High Energy Physics

HREF23 http://www.fnal.gov/fermilab_intro.html - Laboratory as a Research Institution

HREF24 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF25 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF26 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF27 http://www.hr.doe.gov/doevideo/doedtsad.html - U. S. Department of Energy Televideo Service

HREF28 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF29 http://www.doe.gov - U. S. Department of Energy

HREF30 http://www.whitehouse.gov - The White House

HREF31 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF32 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF33 http://www-fermitools.fnal.gov - Fermilab Software Tools Program

HREF34 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF35 http://www.fnal.gov/fermilab_at_work.html - Fermilab At Work

HREF36 http://xxx.lanl.gov/xxx - e-Print archive

HREF37 http://www.lanl.gov/ - Los Alamos National Laboratory

HREF38 http://libmc1.fnal.gov/Hot.html - relevant preprint bulletin boards

HREF39 http://libmc1.fnal.gov - the Fermilab Library's web server

HREF40 http://libmc1.fnal.gov/Hot.html - "Hot Links" on the Fermilab Library web server

HREF41 http://libmc1.fnal.gov/Hot.html - "Hot Links" on the Fermilab Library web server

HREF42 http://www.fnal.gov/top95/top_discovery.html - links to the top quark discovery papers

HREF43 http://docs.whitehouse.gov/white-house-publications/1994/08/1994-08-03-science-in-the-national-interest-policy-statement.text - "Science in the National Interest"

HREF44 http://www.hep.net/documents/drell/full_report.html - Drell Report

HREF45 http://www.ncsa.uiuc.edu - National Center for Supercomputing Applications

HREF46 http://frank.mtsu.edu/~kmiddlet/libweb/innovate.html - Innovative Internet Applications in Libraries [HREF46]

HREF47 http://www.library.usyd.edu.au - Fisher Library at the University of Sydney

Footnotes

---

* Note: Double underlined words in the paper version of the document indicate the hypertext references in the network-resident version of the paper submitted to the conference. For the network-resident version, see the URL: http://www.scu.edu.au/ausweb95/papers/libraries/garrett/

For information more generally about the conference, see the URL: http://www.scu.edu.au/ausweb95/

This document was prepared with MicroSoft Word and converted to HTML for the network-resident version using rtftohtml. For further information, see ftp://ftp.cray.com/src/WWWstuff/RTF/rtftohtml_overview.html.

1 Head Librarian, Fermi National Accelerator Laboratory, January 1987 - March 1995.

2 Webmaster, Fermi National Accelerator Laboratory.

3 T. Berners-Lee, Presentation given at Fermilab, July 1992 (unpublished).

4 J. Streets, private communication.

5 Penelope Constanta-Fanourakis, Guide to Using the Document Database - DOCDB, PN 336, March 29, 1988, (unpublished).

6 The top quark is the last to be discovered quark in the Standard Model theory linking all sub-atomic constituents. Researchers have searched for it in physics experiments for some eighteen years. In April 1994, evidence for it was announced by the Fermilab CDF and D0 Collaborations. In March 1995, CDF and D0 announced its discovery.

For example, in March 1993, one experiment's private unpublished meeting minutes requested collaborators to store material in the experiment's VAX Notes electronic "conference," a vendor-specific, platform-specific, system somewhat like a bulletin board. By June 1993, the experiment's minutes had announced V1.0 of the experiment's "Documentation System," a WWW Server and implied the demise of VAX Notes. The sense in the announcement was that merely of setting up another channel of collaboration communication–of little relevance to the laboratory Directorate and the laboratory's public posture.

8 P. Garrett and D. Ritichie, "Fermilab Library Projects," FERMILAB-TM-1667 (1990) and P. Garrett and D. Ritchie, "Fermilab Library Directions," FERMILAB-TM-1668 (1990).

SLAC MEMORANDUM
Research Division

TO: Pat Kreitz
Library MS 82

FROM: David Leith

DATE: October 27, 1995

SUBJECT: WWW Coordinating Committee

Thank you for agreeing to serve as a member of the World Wide Web Coordinating Committee. As the first Web site in the United States, our Web accomplishments are extensive. It is now time, in this evolutionary process, to provide SLAC Web authors, groups, and committees with policy guidance, standards, and further support.

As the charge below reflects, the Associate Directors' Committee on Computing and I expect this committee to be responsible for developing and recommending policies and standards to ensure the information provided on SLAC Web pages is appropriate and consistent with Stanford University and DOE policy, presents an acceptable public image, and works effectively to meet the Laboratory's goals and its needs for a variety of approaches to Web communication. While recognizing the need for more structure at the policy level and for a group which is responsible for the appropriateness and accuracy of SLAC's Web presence, the purpose in creating this committee is to support the appropriate uses and further the creative development of the Web at SLAC.

The charge and committee membership are below. Pat Kreitz, the committee chair, will be contacting you soon with an agenda and to arrange meeting details.

CHARGE:
The World Wide Web Coordinating Committee (WWWCC) is charged with the following purpose:

To develop and recommend policies and standards to ensure the information provided on SLAC WWW pages is appropriate and consistent with Stanford University and DOE policy, presents an acceptable public image, and meets the Laboratory's goals.

The committee should establish a structure for the organization and presentation of WWW pages which will provide tailored tools to meet the variety of purposes that WWW serves at SLAC, working closely with existing Web committees and groups as

appropriate to provide them with policy guidance, standards, and support. The policies, standards, and oversight levels should be appropriate to each level of SLAC WWW pages -- public, technical, and individual. The WWW Coordinating Committee will identify gaps in SLAC's WWW information and make recommendations for improved coverage in under- represented areas of the lab. Finally, the WWWCC will assist and oversee the WWW Public Page Committee in identifying gaps in Web information for the general public and in encouraging the creation of needed pages or electronic documents.

The committee will report to the Associate Directors Committee on Computing (ADCoC), and the committee's report will be incorporated into SLAC's computing strategy document.
The committee should consult with the ADCoC on its progress and to refine the direction of its work.

Membership of the committee:

Lisa Dunn
~~Katherine Cantwell~~, SSRL
x3191, MS 69, cantwell@sse*
Andrea Chan, PEP-II
x3524, MS 17, achan@sse
Les Cottrell, Central Computing
x2523, MS 97, cottrell@sse
Jenny Huber, SLUO and SLAC Experimental Groups
x4169, MS 94, huber@sse
Pat Kreitz, Technical Information Services
x4385, MS 82, pkreitz@sse
Ruth McDunn, ES&H
x3054, MS 84, isis@sse
Sharon Minton, BSD
x4458, MS 02, sminton@sse
Jon Rosell, Public Information Officer, Directorate
x2605, MS 80, jerosell@sse
David Whittum, Technical Division
x2302, MS 26, whittum@sse
Steve Williams, Research Division
x2276, MS 80, steve@sse
Tony Johnson, ex-officio member and chair of WWW Technical Committee
x2278, MS 71, tony_johnson@sse

*SSC = Slac.Stanford.edu

Stanford Linear Accelerator Center
SLAC Library
(415) 926-2411

Memorandum

To:: ADCoC
From: WWWCC, Pat Kreitz
Subject: PROPOSED Web Support Coordinators Group
Date: 5/9/96

The Web Coordinating Committee proposes that the ADCoC formally establish a Web Support Coordinator organization modeled on the Apple Coordinators group. Each Division could determine the structure and number of Support Coordinators needed, but ought to have at least one formally-appointed Web Support Coordinator with sufficient technical expertise to manage the Division's Web AFS space, advise authors, monitor security and other technical issues, and assist the Division's Web Coordinating Committee representative in ensuring compliance with SLAC Web policies. If one or more groups within a Division have sufficient need, the Web Support Coordinator structure could be extended as the Division chooses.

Use of the SLAC Web has gained sufficient momentum that formal technical support is now necessary at a more local level than provided by the WWW Technical Committee and the SLAC Web Users' Group (SWUG). Both the Technical Committee and SWUG support this proposal and would be better able to pursue some of their objectives if they had a formally-acknowledged Web technical support structure at the divisional or group level. This concept has been discussed with a number of very active Web authors (proto Support Coordinators). They have supported it in concept and have supplied names of active Web technical experts to help form the attached list of potential Web Support Coordinators.

If more than one Web Support Coordinator is appointed for a Division or Group, it is important that the Associate Director clarify lines of responsibility and expectations. It is also critical to the success of this proposal that the Support Coordinators should work closely together to ensure a smooth management of a division or group's Web presence and to work cooperatively with the existing WWW committee structure. At least one Web Support Coordinator from each group/division would be that group/division's official representative to SWUG, perhaps eventually forming the core of a SWUG steering committee.

Web Support Coordinators must have, or be willing to develop, sufficient knowledge to provide the technical management and day to day support of the Web within their area of responsibility. Their responsibilities would include:

- work with their divisional WWWCC representative to understand, communicate to users, and implement SLAC-wide and any specific Division/Group policies, standards, procedures and/or practices, particularly in the areas of computing guidelines, copyright law, privacy, security, etc.;
- plan and manage the division/group Unix/AFS Web space effectively;
- represent their division/group (or ensure a substitute) at SWUG meetings and work closely with the WWW Technical Committee on relevant issues;
- provide technical training and support to 'local' Web authors and users;
- evaluate, advocate, and (perhaps) maintain Web development and maintenance tools needed by their users.

WWW Coordinating Committee
Membership:

Chair: P. Kreitz
Andrea Chan, PEP-II & BaBar
Les Cottrell, Central
Jenny Huber, SLUO and SLAC Exp. Groups
Ruth McDunn, ES&H
Sharon Minton, BSD
P.A. Moore, Directorate
David Whittum, Technical Division
Steve Williams, Research Division
Tony Johnson, Chair of WWW Technical Committee

Web Support Coordinators
Potential Membership List:

PEPII/BABAR
Tanya Boysen
Andra Chan
Steve Meyer
George Crane

BSD
Sharon Minton
Sarah Jones

DIRECTORATE
Diana Gregory
Kathryn Henniss

ES&H
Ruth McDunn

SSRL
Lisa Dunn
Alan Winston
Heinz-Dieter Nuhn

TECH. DIV.
David Whittum TSP
Tor Raubenheimer TSP
Bob Boeninger MD
Barry Prentiss MD
Rob Richards MD
Gregory Sherwin CD

RESEARCH DIVISION
Sharon Jensen TH
Laurie Gennari TP
Harv Galic LI
Charlie Young EA
Jochen Swiening EB
Lilian DePorcel EB
Tom Glanzman EC
Perry Anthony EFD
Christian Bula EI
Andrew Lee EK
Gayana Shabad EK
Joan Winters SCS
Les Cottrell SCS
Ilse Vinson SCS
Karen Heidenreich SLD
Joe Perl SLD

DISCOVERY, CONFIDENTIALITY AND SECURITY: ISSUES IN COMPUTER USE

June 26-29, 1994

Richard Leurig
Montgomery College
Rockville, Maryland

I. INTRODUCTION

Confidentiality and security issues in computer use are rapidly becoming a major part of every higher education information technology and computing center. There is increased emphasis by corporations on security, resulting in movement within institutions to offer courses dealing with security. The incidents of violations of confidentiality and security is increasing based on complaints expressed in e-mail messages exchanged by faculty, administrators and students discussing their experiences. This increased awareness will accelerate institutions interest in legal, and moral, issues.

Understanding the basics of computing confidentiality and security is essential to deal with the legal and operational issues. Unfortunately, this understanding usually takes place after a problem event, and only then does the need for responsibility of keeping up with rapidly changing developments become apparent to the institution. Legal departments, attorneys, information technology and computing are involved when complaints are received, but often these areas are not involved in preparing policies and procedures to avoid problems. Lack of understanding of the basic issues and assignment of responsibilities are a major cause of lack of involvement, as well as the need to staff the functions connected with computer confidentiality and security inside the organization.

The various aspects of confidentiality and security include virus protection, breaking into other users e-mails, hacking and cracking at computer systems and networks, legal and policy issues related to new forms of encryption, freedom of speech, and harassment issues. The areas are so broad and wide ranging that it is difficult to focus on the subject, but the part affecting most institutions is electronic mail (e-mail). This session is using e-mail to highlight some of the issues and to discuss some basic concepts, identify present and upcoming security issues and suggest recommendations for keeping current on the subject. This part of the session presents more of the technical and operational aspects, and the other part presents more of the legal and discovery issues.

II. MORE NETWORKS = MORE USERS = MORE PROBLEMS

Networks are expanding at a rate far beyond most projections a few years ago. The most significant increase is in local e-mail within companies and institutions. There is difficulty in determining the numbers of users of local e-mail as this depends on staffing levels and availability of networks, but most institutions have internal e-mail or will have it by the year 2000. The number of users is significant in dealing with security issues as it is easy for the local users to tie into local e-mail, or to a network provider or to Internet. To deal with the size of the problem, it is an easy projection to imagine that every working individual will use at least one e-mail provider within the next six years.

A. For higher education institutions, most faculty, administrators and administrative staff will be connected to at least one network, with many connected to two or three. For an institution of 500 staff, this could be at least 1,000 types of connects with hundreds of thousands of messages a year. The function represents a major training problem for proper use of computing and e-mails, distribution of policies and procedures, security monitoring and possible confidentiality and potential security problems.

Messages will be sent and received within the institution and externally to other institutions, governments, corporations and world-wide communications. The institution e-mails will have expanded beyond the institution walls and have the potential of national and global legal issues affecting the individuals, departments, institutions, boards, alumni and potential students. Basic security and policies are essential at the first level of user involvement.

B. Growth of networks increases the potential for "hackers" (those who try to break computer code and security because of the need to test their knowledge of hardware and software capabilities, usually without intentional harm), and "crackers" (those who try to do harm by cracking the system with intent to spread virus or corrupt the system). The inter-relationships of networks allow rapid dissemination of confidential and perhaps damaging material. By the time breaches in confidentiality and security are found, the damage may already be past repair.

C. **Local Area Network** (LAN) e-mail systems are usually internal systems and seldom are meant to go outside the institution. However, the confidentiality and security issues basic to this system need to be in place to protect the institution and staff, and also to be the basis for determining policy and procedure for other networks. The size of this network depends on the institution, but probably includes all faculty, administrators and administrative support staff by the year 2000.

D. **Commercial providers** of networks offer e-mail in addition to various forums and discussion groups. Institutional use of commercial providers usually is to contact someone not on local e-mail or Internet, or to take part in discussion groups or to find information on the forums. The largest of the commercial providers is CompuServe, and many institutions use the technical forums with some dial-in for e-mails. America OnLine, Prodigy, Genie, Delphi and other providers are all competing for their place on the Information Highway. These providers generally attempt to control and police use of their networks for meeting their criteria and standards, and can sever the contract with users who do not follow their rules.

E. **Internet** is the largest network and represents the idea of the "Information Highway" concept now embraced by corporations, government and commercial providers. Being on the "Net" is a goal of many individuals, even though many have little knowledge of the capabilities and responsibilities, or of potential personal issues.

Internet encompasses more than e-mail as there are also discussion groups, news groups, files for access and downloading and the ability to connect (telnet) to other computer systems. Software such as "Gopher" provide access to information such as libraries and reference material, weather, calendars of events and other local items of interest. Some government agencies, like the State Libraries of Maryland, are providing free access to this information. Many higher education institutions provide all of the Internet services for free; however, commercial providers and some institutions have fees for e-mail and to download files.

Although originally started by the National Science Foundation and primarily used by higher education, Internet now has commercial and other government users. There has been some recent attempts to permit advertising and there will be other future attempts at expansion and limitations. For now, the Internet is open with almost no restrictions as to content, intent or use. From a legal, policy, and operational viewpoint, Internet represents the largest problems because of the network design and initial objectives of being open and not controlled.

The number of Internet users is almost impossible to determine since this is a system of networks rather than users, and each network is responsible for recording and assigning users. In May 1994 according to the Computer Emergency Response Team (CERT), the number of host computers (users) on the global network has risen from 80,000 in 1988 to 2.2 million today, a 2,400% increase. [1] Internet size is usually measured in the number of participating networks rather than estimating users. One recent estimate is over 12,000 local networks connected to the Internet, and in May 1993, over 1,000 new networks joined the Internet. [2]

In twenty-five years the use of networks have created a major gap between legal and moral computing issues, and the ability of society and institutions to enact laws and policies in a timely manner affecting these issues. There is no indication this gap will narrow, and many users prefer networks to have as little control as possible. Unfortunately, many local network or commercial network users are not aware that networks other than Internet do have restrictions for content, use, distribution and copyrights.

III. "PRIVATE AND CONFIDENTIAL"

"Private and Confidential" was a method of writing or rubber stamping on envelopes to indicate this was a secure document intended only for the recipient. Computing e-mails have a difficult time showing this, although the new encryption methods are trying to replace this old rubber stamp on the envelope. Various confidential schemes have been tried and most do not seem to meet the task of beating hackers and crackers. Increases in networks and users have resulted in increased problems being reported of violating e-mails and networks.

A. Increased violations are caused not only from the increases in numbers of users, but also because of the easy access to multiple networks and interest in computer literate technical individuals experimenting or trying new methods. Some computer bulletin boards publish detailed instructions on how to enter networks and cause confusion, if not damage.

B. Some recent examples affecting what used to be called "private and confidential".

1. University of Michigan in April had messages sent to global bulletin boards containing racial jokes and threats. Authorities are pursuing criminal charges, but have not made arrests. [3]

2. Indictment of a student at Stephen F. Austin University and a student at the University of Illinois for threatening the life of President Clinton in separate e-mail messages. [4]

3. Campus library closed at University of Wisconsin at Stephens Point because of a bomb threat posted on computer e-mail. [4]

4. Indictment by a federal grand jury of a student at Massachusetts Institute of Technology for allegedly operating a bulletin board used for exchanging copies of copyrighted software. [4]

5. Half the students in a course at Dartmouth missed an examination because they were fooled into thinking the professor had canceled the test after reading a bogus e-mail message.

C. Many institutions are dealing with increased complaints and security violations. Conferences dealing with computing and information technology are have increased numbers of lectures and presentations on dealing with the entire issue of computing, network and e-mail use. There appears to be an increasing number of racial and sexual messages on higher education networks, resulting in questions of how this deals with harassment. Chain letters, anonymous complaints, and personal attacks have been on the increase in messages. It has become easier to obtain false identification numbers and easier to break into other users passwords and files.

It is difficult to tell if this is the same percentage of problems cause by the same percentage of users or by new offenders consisting of a larger percentage of all users. The increase in the size of networks and users masks the real statistics and the potential problem-makers. While probably only a small percent of the total users, the offenders will cause a major impact on staffing for security and for additional legal and policy considerations.

D. There are interesting legal issues, and moral issues, resulting from publications of material available on networks dealing with how to crack networks, identify passwords and create havoc. There also seems to be a problem of perception that this is not affecting anyone as there is usually not a face-to-face confrontation or any identification of the damaged parties. It is not "personal" and the computer is a piece of hardware, not a person. "Personal and Confidential" to many users is something in the past.

IV. SECURITY

Security and counter-measures have become a serious game between crackers and authorities. Security packages are being developed and new encryption methods are being tested. Many of these represent legal issues or potential legal issues, and some are being debated in Legislatures and Congress. This section describes some of the security issues, examples of how security can be breached and items that will be of future legal interest.

A. Basic Security measures include a statement of computer use and user obligations from the institution and at least a basic instruction in how to use basic security. The heart of basic security is the stand-alone computer at the workplace. Computers become violated usually through passwords, downloading corrupt software, or having system security holes or flaws.

1. Passwords remain the building block for good security.
   a. Change what was assigned
   b. Choose a good password. Many groups feel at least 75% of computer break-ins are a result of selecting a poor password. Some guidelines:
      More than six characters; Mixture of case and numbers
      Not easily typed on a keyboard (QWERTY)
      Not a word in the dictionary
      Not directly related to you, your job or your activities
      If you have the ability, assign passwords to your PC and selected software
   c. Change password often and don't record it in the computer or post it on the video

2. Don't be so trusting. Telling your passwords may reach someone else looking for openings. Other users can see you type passwords by looking over your shoulder. Lists of users and passwords are often kept by network administrators and may be printed.

3. Be suspicious if unusual events happen to your computer, such as an obvious slow down, or a change showing on a file that you did not change, or if a log-in showed and you were not there. There are tracking programs that can assist in finding unauthorized use.

4. If you feel the computer or your files have been violated, follow standard processes to determine if your files are secure. **The Whole Internet**, chapter 4 on "What's Allowed on the Internet" covers methods of computer security and what to do in case of compromise. [5]

B. Networks increase the possibilities of passwords distribution of name and numbers. Multiply your stand alone workstation by the number of networks you are connected to and your various passwords to see the total problem. If a password is found, hackers can trace to your workstation.

C. Think of where your data and your security codes can be located and take precautions:

Regular files, such as word processing (*.DOC) or text files (*.TXT)

Backup files (*.BAK)

Hidden files and other files with special attributes (check out attributes)

Deleted files that have not been "cleared" remain on disks (usually as $*.*)

Backup diskettes, tapes, LAN backups and network backups

Printers that are dedicated or attached to networks (and your wastebasket)

The screen you are looking at when working; even screen savers do not stop snoopers

D. Some examples of problems are shown on the attached computer screens.

Screen 1 shows "finger" useful in finding names; but can also be used for other information.

Screen 2 shows hidden and files with user information and a bad idea for a password file.

Screen 3 shows a password reference in an autoexecute file to start a computer.

Screen 4 shows a backup file and other file designations that are easy to forget

Screen 5 shows deleted files not yet cleared. These can be undeleted.

E. System holes and security holes are the subject of many bulletin boards and Internet discussion groups. These "holes" are areas found by hackers and crackers enabling them to enter into systems and find data for use in defeating security measures. One technique becoming standard is shown on the Firewall slide. The firewall filters movement of data between internal and external systems to make it more difficult for unauthorized entry or exit. The filter is a system of hardware and software.

V. CONCLUSIONS AND KEEPING CURRENT

The other section of this session has suggestions on internal process and procedures. The core of most conclusions and suggestions is to follow common sense approaches to informing users and the institution of potential risks and legal implications. Policies and procedures are an essential part of any security plan. Assignment of responsibilities are required. Policies, Procedures, Responsibility, Communication, are the essential tools to making security or any institutional initiative work and apply in the case of confidentiality and security.

The field of computer use and effects on confidentiality and security require keeping current. In addition to professional publications containing legal references and updates, there are publications and bulletin boards and networks containing the latest news and projects. Some of these are shown in the endnotes, but many are being developed, initiated and changed as is common on networks.

CompuServe has a security forum and other hardware and software forums contain security information. When on CompuServe, type "Find Security" for a list of the forums and information.

Internet has a number of references to security. A mail list for discussion of computer security is available through SECURITY@RUTGERS.EDU. The Computer Emergency Response Team (CERT) can be reached by e-mail on CERT@CERT.SEI.CMU.EDU. They also have papers and files reached by FTP CERT.SEI.CMU.EDU; CD PUB. Computer ethics discussions and papers can be found on FTP ARIEL.UNM.EDU; login ANONYMOUS; CD ETHICS. These are part of a good method of keeping current, and contain many discussions of systems security holes and password protections.

ENDNOTES

1. Joseph C. Panettieri, "Guardian of the Net" Information Week (May 23, 1994), describes the Computer Emergency Response Team (CERT) concerning computer crime on the Internet. Two useful numbers in case anything goes wrong is 412-268-7090 or access through Internet CERT@CERT.ORG.

2. Bennett Falk, The Internet Roadmap, (Sybex publishers), page 4 gives the latest statistics for May 1994 on the Internet. Other references to this book contain network etiquette and security considerations.

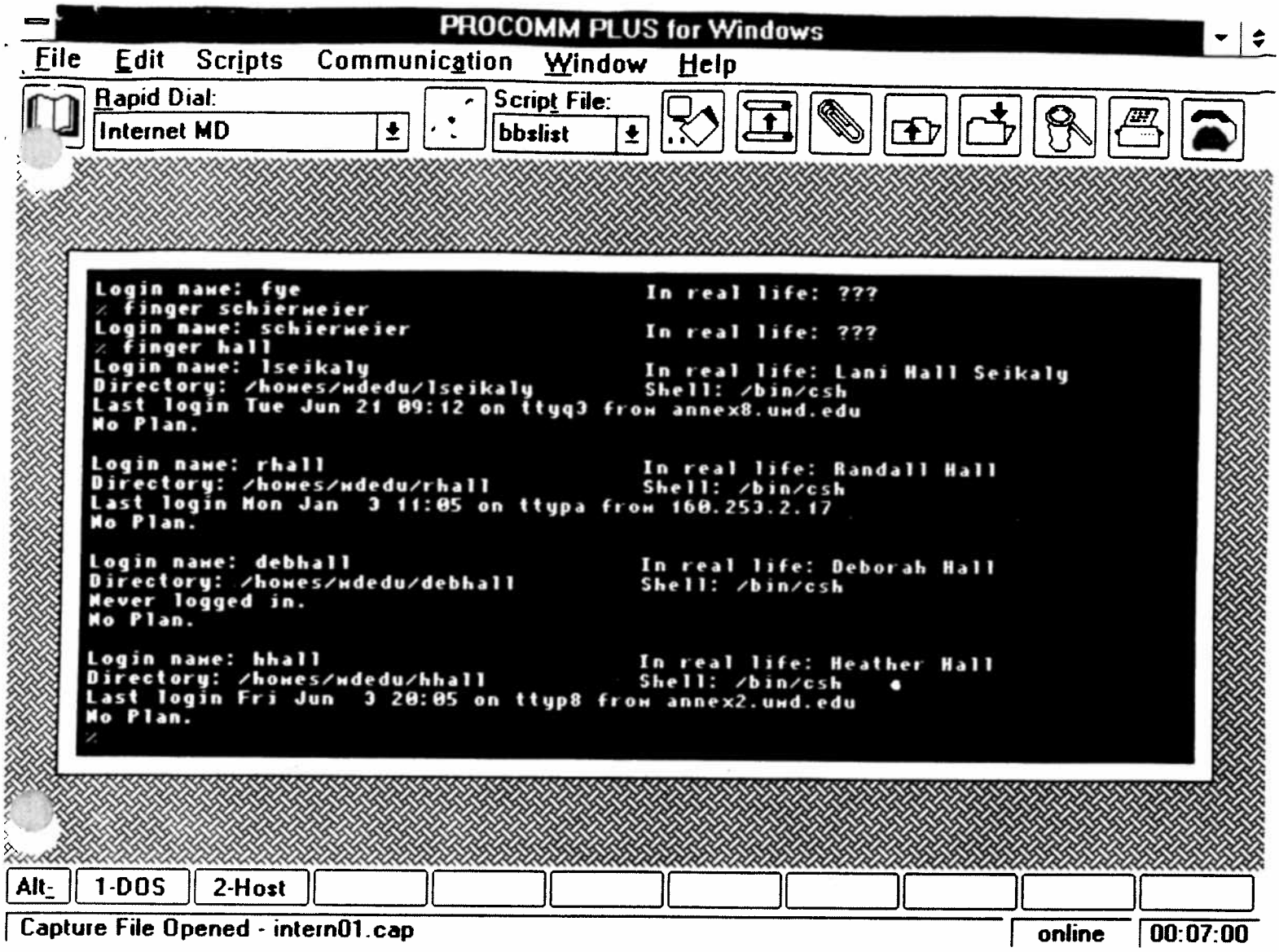
3. The Washington Post article, "Message Posted on Internet Spurs Probe; Jokes, threats directed at African Americans", April 26, 1994, is typical of many stories in newspapers related to network confidentiality and security.

4. Thomas J. DeLoughry, "Countering Mischief" (Chronicle of Higher Education) May 25, 1994, in the Information Technology section dealing with methods higher education is using to combat computer abuse.

5. Ed Krol, The Whole Internet, O'Reilly & Associates, Inc., Chapter Four "What's Allowed on the Internet" has discussions of security, legal implications and network ethics. This is an excellent book on describing the Internet, how it works and how to use the network.

My CompuServe number is 72245,1503.

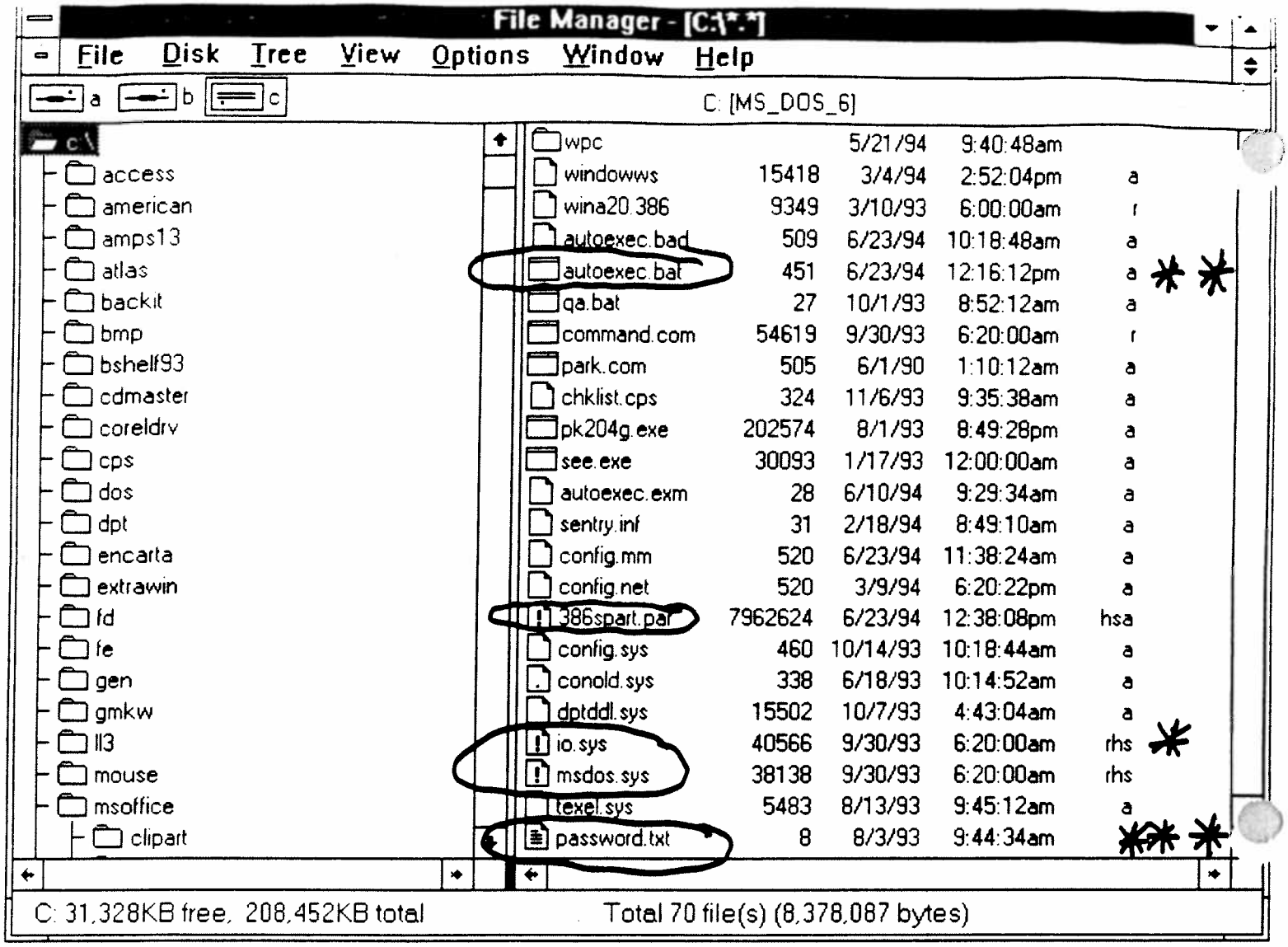
My Internet number is RLEURIG@UMD5.UMD.EDU



FINDING A NAME IS NOT HARD

NEXT IS "GUESSING"

COMPUTER SCREEN 1

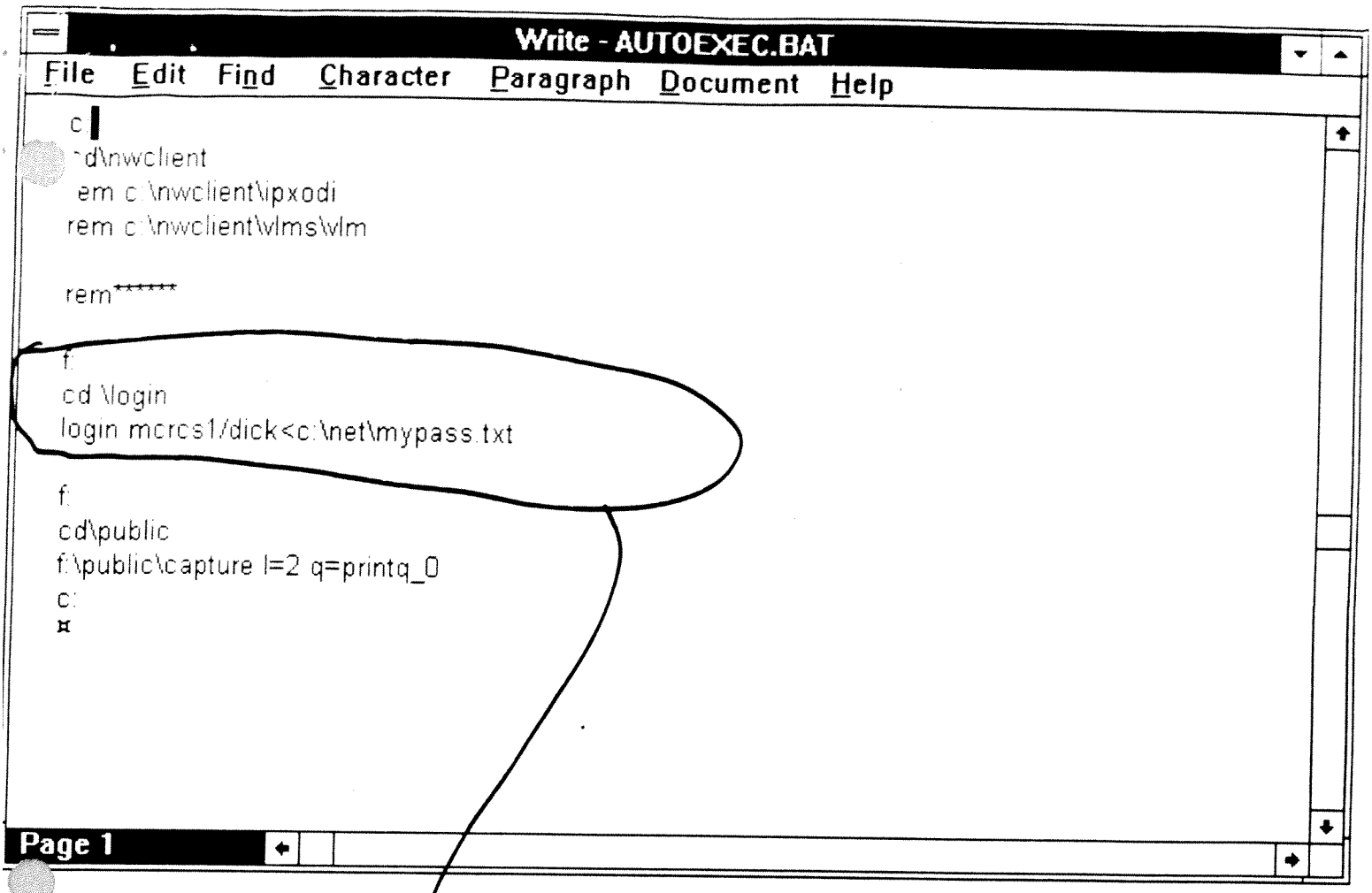


* HIDDEN FILES

** AUTOEXEC HAS USER INFORMATION

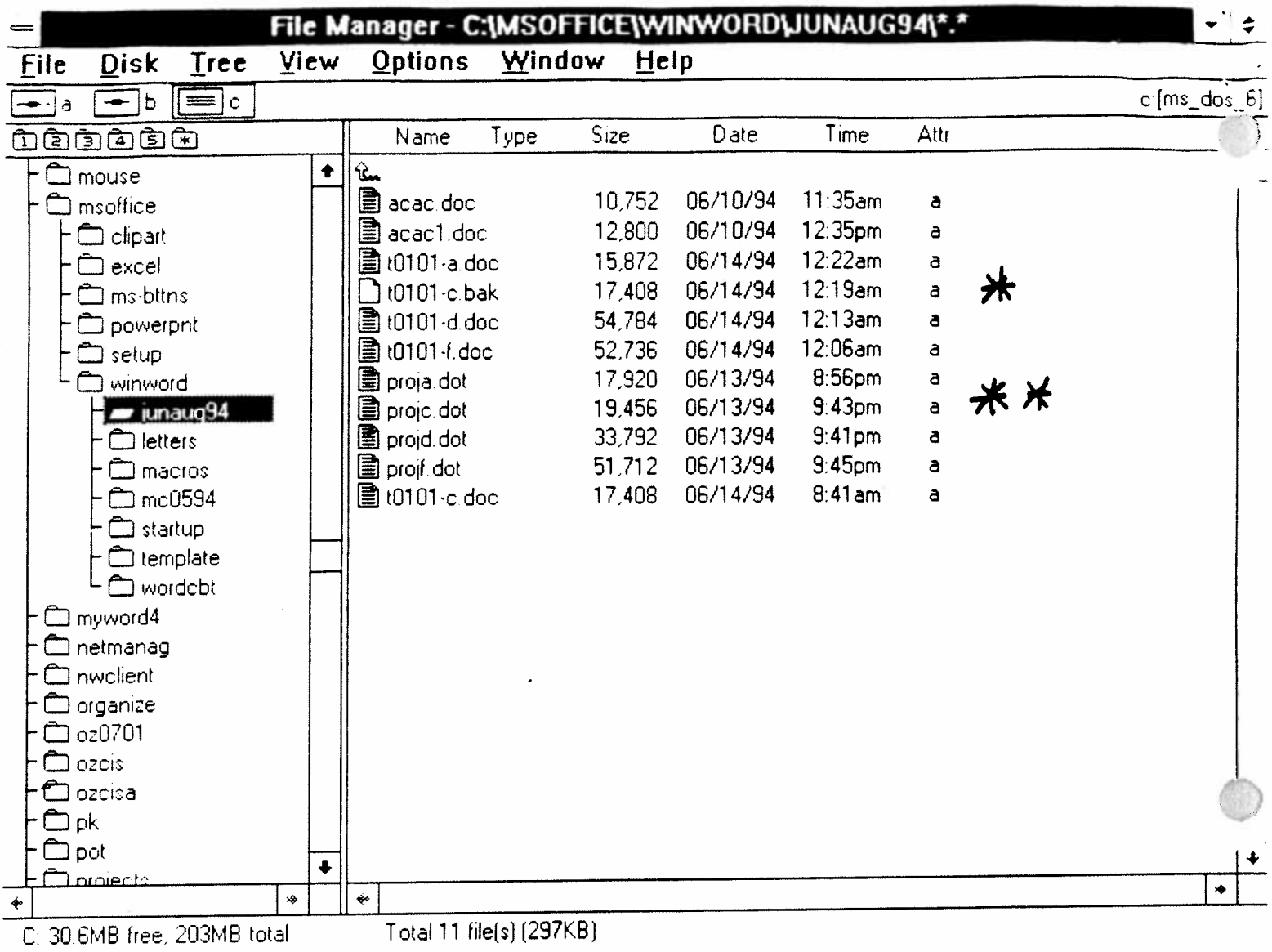
*** PASSWORD FILE!

COMPUTER SCREEN 2



USING PASSWORD TO SAVE TIME —

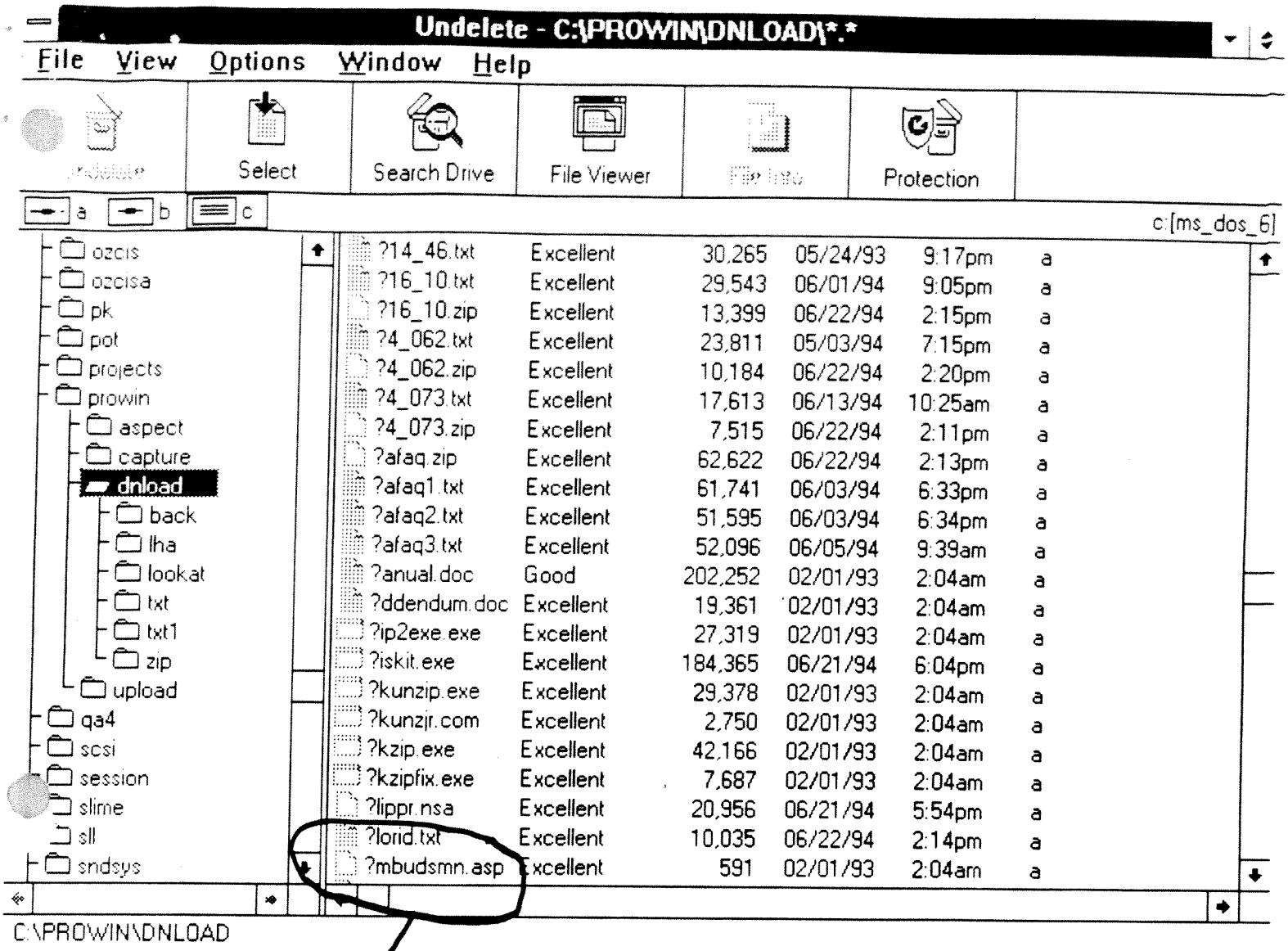
COMPUTER SCREEN 3



* BACK-UP FILE

** NOT SHOWN AS A DOCUMENT

COMPUTER SCREEN 4



DELETED, BUT NOT CLEARED