INFORMATION SECURITY FOR LIBRARIES

Gregory B. Newby
School of Information and Library Science
University of North Carolina at Chapel Hill
gbnewby@ils.unc.edu / 919-962-8064 (v) / 919-962-8071 (f)

Libraries have made significant investments in computer-based resources, training and services. However, such investments need to be protected from misuse or mistake by taking an active role in information security.

INTRODUCTION

By most accounts, the proliferation of the Internet and other computer technologies has been highly beneficial to libraries. Investment in everything from online databases and computing equipment to personnel and training is significant. Libraries need to have policies, protection measures and trained staff in place in order to safeguard their investments in computer and computer-related technologies, personnel and services.

This paper will address the topic of information security, making concrete recommendations for safeguarding information and information access tools. Instead of giving detailed instructions for security techniques, the emphasis here is on setting the agenda for the role of security in library environments.

Historically, formal training for librarians' use of information technology was in the relatively narrow specializations of library automation and online searching. Library automation training (e.g., Ross, 1984) was for library staff that would manage, evaluate, and sometimes design and implement technology systems in libraries. The OPAC (Online Public Access Catalog) was a centralized system based in the library or a regional office for circulation and holdings information, as well as other types of data (serials control, acquisitions, cataloging, etc.). Library automation training seldom mentioned any type of security for protection of data, privacy or equipment. Even fairly recent books on the use of computers in libraries, such as Ogg (1997), make almost no mention of security issues.

Today, library environments are increasingly reliant on computer technology. Many libraries of all sizes have discontinued use of card catalogs in favor of electronic versions – and many of the electronic versions previously accessible only via terminals within library buildings are now Web-accessible. Online searching of a plethora of databases and other information sources has become ubiquitous for the end user, rather than being restricted to librarians trained in online searching. Access to general-purpose microcomputers and software, as well as to the Internet, is offered in nearly all libraries of significant size.

It is the position of this paper that security training for librarians is extremely weak, both on the job and in educational institutions. One result is that opportunities for problems related to information security in libraries are likely in many library environments. Although some recent texts on library security address aspects of information and computer security (for example, Shuman, 1999), most do not.

In this paper, a pragmatic approach to addressing the information security needs of libraries is presented. Effective information security must involve active staff and active measures to minimize risk of damage, theft, subversion or sabotage. Following an overview of information security, sections discuss security personnel, privacy policy, the OPAC, public access workstations, and the library's Internet connection. A concluding section addresses the emerging role of security training for librarians.

OVERVIEW OF INFORMATION SECURITY

Information security is not simply computer security. Whereas computer security relates to securing computing systems against unwanted access and use, information security also includes issues such as information management, information privacy and data integrity. For example, information security in a library would include personnel security and policies, steps taken for effective backups, and the physical integrity of computing facilities.

According to a recent survey of executive recruiters, computer security experts are among the six most sought-after professionals for the corporate world (Radcliff, 1999). Yet, there are very few college courses addressing computer security. Those that do mostly emphasize the mathematics of encryption rather than hands-on information security and management.

Minimally, effective information security in libraries should include:

- Staff assigned to information security tasks
- Training all personnel in information security issues and procedures
- Specific policies dealing with information privacy, physical security of equipment, and computer security procedures
- Physical security plans
- Data integrity measures
- Levels of access to data or equipment, and monitoring for different types of access

These points are intended for all types of libraries – public, academic, corporate, and special libraries and collections. They are intended for libraries of all sizes, with all types of patrons, funding models and organizational structures. In a particular library, the investment in information services, computer equipment and personnel may be greater or smaller than in another library, but the need for effective information security exists in both.

WHO'S IN CHARGE OF SECURITY?

On November 2, 1988, a malicious and dangerous computer worm was released to the Internet by Robert T. Morris, causing thousands of computers to crash (Denning, 1990). One of the most important lessons learned in the aftermath of diagnosis, analysis and cleanup was that there was no easy way for computer network managers at one institution to find their counterparts at other institutions, even during an emergency. Furthermore, many institutions that were connected to the Internet (still in its infancy) had no person in charge of computer security, and little or no staff prepared to address security issues.

Today, most libraries are connected to the Internet, yet often without personnel specifically responsible for managing the security of the library network from intrusion or tampering. The estimated number of illegal intrusions into US government computers and networks exceeded 500,000 for 1998, with most intrusions originating from the Internet. By analogy, it may be assumed that attempts to infiltrate library computer systems and networks are frequent (even if they are not frequently detected). If a break-in occurred, would a library be in the same position as the Internet in 1988, with no clear plan for how to proceed, and nobody in charge?

The good news is that computers excel at record keeping. Network logs, firewalls, routers, packet sniffers and data integrity checkers are all capable of identifying illicit access to computer systems. However, even the most sophisticated automated system requires someone to identify that a security breach may have occurred and decide what action to take.

Recommendations for effective information security management include:

- All areas of information security risk must be assigned to specific personnel. For example, if there is a public-use computer workstation, someone needs to be responsible for its security. Similarly, someone needs to be responsible for the security of circulation data.

- Information security must occupy a non-zero amount of the personnel's time (e.g., a 5% or 10% allocation devoted to security might be appropriate for the person who manages a software loaning program in a library).

- Personnel responsible for security must actively seek information related to their areas. Reading vendor news releases, subscribing to security-related electronic mailing lists, or seeking security-related training may be appropriate. Alternatively, simply having regular intra-library information security discussion sessions may be sufficient.

- Develop an access level hierarchy for personnel. Insure that people with access to potentially sensitive data or systems are known, and that their access level will periodically be reviewed.

Information systems are like buildings: simply creating them is not enough. Ongoing maintenance is required in order to avoid inevitable decay due to interaction with the environment. To be diligent about the security of information systems, personnel with specific security-related job descriptions are a necessity.

WHAT IS YOUR PRIVACY GUARANTEE?

Libraries, especially public libraries, have an outstanding record of protecting the privacy of their patrons. The American Library Association's Intellectual Freedom Manual (ALA, 1996) assists librarians in defending the Library Bill of Rights. More recently, the ALA has taken a strong stand against the use of filtering software in libraries (see ALA, 1998), and specified guidelines for the freedom of computer and Internet use in libraries.

In spite of this record, there are two important security problems often not addressed in libraries. The first is the privacy offered for data that may be collected or collectable apart from circulation records. The second is the risk of penetration of library systems from outside parties who may access circulation or other data.

For the first case, consider Web browser software often found on library computers. Such software keeps a history list of sites visited, and keeps copies of recently visited Web pages in a directory on the computer. Effective information privacy would dictate a specific policy for such data: will it ever be analyzed? Should patrons take steps to erase the cache after they use a system? If a violation of library policy were suspected, would Web browsing history data be subject to search?

In some libraries, anonymity of access to computer tools may be guaranteed. In others, such as college libraries, display of a current identification card or signing out an access key may be required. Even if anonymity may be assumed, such as in a school library, a small user population might result in easy identification of individuals by library staff.

To turn to the second problem, consider that reasonable and prudent steps should be taken by libraries to insure that private data are kept private. The greatest risk might come from outside the library via an Internet connection, an unattended modem or from staff who abuse their access rights.

Recommendations are simple, but could be time consuming or difficult to implement:

- Maintain a comprehensive list of data that may be collected and the circumstances.
- For each type of data, what risks of misuse exist?
- Specify a policy for the collection of data and possible misuses.
- Identify personnel responsible for ensuring the policies are followed, and for remediation as needed.

For example, a library might require that patrons who wish to use a general-purpose computer first show their library card or ID. At that time, patrons should be informed what data will be kept from their session – will their use of the facility be logged? Will the amount of time be logged? Will different software packages, Internet sites or other records be kept, and if so will the data be linked to the patron's name? Finally, under what circumstances will any data collected be released, and to whom?

THE OPAC

Traditionally, the most sensitive data that libraries collect are circulation records. By necessity, these are linked to identifying information for individual patrons who borrow books or other materials. Online Public Access Catalogs (OPACs) are centralized systems that handle circulation and holdings information, as well as a variety of other data ranging from acquisitions budgets to cataloging modules.

OPACs are still with us. Two important changes to OPACs in the final decade of the 1900s have been consolidation in the OPAC industry and the expansion to Web- and Internet-based access models. For consolidation, the number of companies selling OPAC systems suitable for use in all but the smallest libraries has diminished rapidly. Fewer than a half-dozen companies would be suitable choices for, say, a large academic or public library.

Nearly all new OPAC systems are based on variants of the Unix operating system. Modern OPACs include functionality to make the holdings information searchable via a Web interface. Here lies the substantial security risk: Unix systems have many potential security flaws, and many well-known flaws have easy exploits available to any potential intruder.

Connecting a system with critical data to the Internet is a bad idea. On the Internet, tens of thousands of amateur (and professional) potential intruders may try to get access to the system. Even if the OPAC software itself is thought to be relatively free of security problems – a risky assumption to make – the underlying Unix operating system is almost definitely not.

Yet, this potential risk needs to be balanced by the desire to make OPAC services available to the outside world. Recommendations include:

- Only services needed should be running on the OPAC computer(s). Specifically, all Unix services (such as email, FTP, rlogin, telnet) that are not required for OPAC functions should be disabled.
- System logs must be kept, and analyzed regularly (daily or weekly) by staff. A logins record should be maintained; integrity checkers such as Tripwire should be used to spot illicit changes to the system software, and the system should be audited regularly for usernames, programs or data that are no longer used.

- Ideally, the OPAC should only communicate with authorized terminals. For example, computers located within library buildings. If outside (Internet) access is required, the ideal scenario is to have a duplicate of the holdings database (or other information, if needed) on a separate server. This way, if the duplicate server were compromised, the original data and services would be intact.

- Personnel must be specifically responsible for monitoring security updates from the OPAC vendor, as well as the underlying Unix system vendor.

- Regular attempts should be made to bypass OPAC security from both within the library (at computer stations) and outside the library via the Internet. Intrusion tools are widely publicized on resources such as the BUGTRAQ mailing list[1] and PacketStorm Security Web site[2].

Historically, OPAC security has relied on (a) obscurity, and (b) OPACs' relative inaccessibility. These factors have changed. The Unix systems that OPACs are based on have well-known security flaws, and flaws in the OPAC software are more likely to be found when the OPACs are accessible to the thousands of potential intruders on the Internet.

PUBLIC ACCESS WORKSTATIONS

Instead of "dumb" terminals that can only access OPAC services, libraries often use fully featured microcomputers. These PCs might use a Web browser as an interface to the OPAC's collection information, or they might use a telnet client. Additionally, the PCs could be used to run other software, or access Internet sites outside of the library.

There are several areas of possible risk, as well as various policy issues. Policies for what the computers may be used for, whether priority must be given to particular purposes (such as searching the OPAC), and what data may be gathered from the PCs users (mentioned above) must be developed.

Risks include illicit access to resources within the library, physical risk of theft or damage, and risk of illicit access to resources outside the library via the Internet. Policy issues have to do with equity of access (for which the ALA has specific guidelines, in both their Intellectual Freedom Manual and more recent 1996 "Bill of Rights in Cyberspace."), what services and facilities will be offered, and whether limitations on particular types of use will be made[3].

Computers within the library building might have elevated levels of access to information services. For example, a computer might be able to access a CD-ROM database or circulation records for a patron (e.g., to see what books you have charged out), or access an internal electronic message board for reference. Security problems for PCs connected to LANs are well known, and range from packet sniffing (by which usernames and passwords might be gathered) to weaknesses in the underlying PC operating system.

For theft or damage, the requirement is to consider any accessible item to be subject to tampering, theft, damage, sabotage or subversion. Loose Ethernet cables might be plugged into notebook computers brought to the library by potential intruders. Computer system units might be opened so that memory or other components may be stolen (especially in privacy-enhance computer use areas such as carrels). Floppy disk drives may be used to attempt to reboot a computer to a different operating system, or they may simply be used as a depository for chewing gum.

For libraries offering Internet access, an important risk is that this Internet access could be used for illicit purposes. Perhaps a disgruntled employee wants to send anonymous threatening email to her employer. Or a teenager steals a credit card to pay for access to pornography download sites. Or a computer expert uses his hacking skills to break into the FBI's computer network. In these cases, the library may be called on to try to catch the perpetrator. The goal for the librarian should be to make it difficult for the library computers to be utilized for criminal activities.

Recommendations span the range of possible misuses of library-based PCs:

- Perform a complete audit of software and hardware for each computer available. Insure there are no components that are not needed (e.g., additional software that might make illicit use easier).

- Secure wiring and the computer system itself. Assume that any possible illicit use will be attempted.

- To prevent packet sniffing (listening to data sent to other machines on the network), network *switches* should be used in place of network *hubs.* A switch provides a single network channel for each computer, removing the opportunity for sniffing.

- Assume that the software on the computers is subject to change, by either accident or malicious intent. One solution to this problem is to have the computer software refreshed from a LAN server at every reboot. Another is to perform nightly backups to spot

inconsistencies and flag problems. A third approach is to rely entirely on remote read-only servers for all software.

- Have personnel regularly attempt to bypass system security.

- Decide what level of external access to allow (see below), and what level of logging of access is appropriate.

For a medium-sized public library, expenditures for a public-use computer facility with Internet access could exceed tens of thousands of dollars annually, plus personnel costs. Taking steps to insure the equipment is secure is a sound policy.

YOUR INTERNET CONNECTION TO THE OUTSIDE WORLD

Connecting a library to the Internet means that the Internet can also connect to the library. Procedures for minimizing risk that potential intruders could steal, modify or delete information are well known, and focus on the use of Internet firewalls (e.g., Chapman and Zicky, 1995). A firewall can help insure that many exploits used to gain illicit access to data or systems will fail, such as IP spoofing attacks (where an outside computer masquerades as an inside computer to get elevated access privileges).

Libraries that host their own OPACs (versus those that use a regionally shared one) should consider an additional firewall for the OPAC itself, also perhaps for other particular resources such as cataloging workstations or circulation workstations. This would deliver a two-tiered approach to network security. At the first tier, a firewall system would help protect the library from unauthorized data. At the second tier, a separate firewall configuration would help protect particularly sensitive systems within the library from unauthorized use.

For example, it might be determined that many sorts of Internet data are allowed to go in and out of the library, but only data to particular ports would be allowed to reach the OPAC. The second level firewall would also prevent computers inside the library from sending illicit data to the OPAC.

Recommendations for the Internet connection include:

- Collect logging data from the firewall or router machines, and regularly examine the logs.

- Configure automatic notification (e.g., a pager message to the network administrator) in case of a network outage or serious break-in attempt.

- Physically or logically (via firewall) separate machines that are intended to be Internet-accessible from those that are not. For example, a dedicated CD-ROM searching station should not be able to access the Internet, even if it's next to a PC workstation that should.

- Treat any Internet connection as high risk. Personnel with any elevated systems access (e.g., the OPAC administrator) should not be permitted to access sensitive systems over the Internet – such as to login to the OPAC system and perform administrative functions. Minimally, if logins or other network traffic that might disclose internal messages (email), usernames or passwords are needed, the session should be encrypted. Encryption software for login sessions includes Secure Shell (ssh); encryption methods for email include S/MIME and PGP. Free or inexpensive solutions are available for such encryption.

The Internet is the greatest source of risk to information security in libraries. This is simply because there is a far greater number of potential intruders "out there" on the Internet then there ever could be within the library.

CONCLUSION

Information security includes personnel security, privacy, policy and computer security. Specific personnel must be assigned security-related tasks in order for any security system to be effective. Due to the continuing emergence of new security exploits, tools and techniques – coupled with the constant parade of software and hardware upgrades likely in most library environments – ongoing diligence is required to keep informed of security developments.

This paper has not attempted to cover every aspect of security-conscious systems administration and library management, only those issues that may be particularly under-appreciated or help lead to more rigorous thinking about security. For example, any information system should have good physical security, to prevent unauthorized access by both casual or planned attempt. Data backup and backup policies are a necessity for any data collection system, and hopefully are in place for all OPACs and related systems. More detailed treatment of security-conscious computer systems administration may be found in such sources as Garfinkle & Spafford (1996).

The traditional of areas focus for library security have had little to do with computer and information security. Disaster and contingency planning (George, 1994) should now include an emphasis on these systems as well.. A final notable issue omitted in the analysis above is the proper maintenance of non-computer based records, such as paper sign-in sheets for computer facilities, written library card applications, and the like. Clearly, such data need to be subject to appropriate policy and procedure.

Colleges and universities that train librarians have changed their emphasis from the relatively narrow world of OPACs and online searching to include the Internet, LAN and server administration, interface design, and programming. However, issues of security and privacy still tend to focus on issues such as risks from troubled patrons, book theft, and censorship (e.g., Chaney & MacDougall, 1992). When information security is addressed, it may be from the point of view of corporate information security management, rather than library environments (cf. Davies, 1992).

In the future, consideration of information security issues will likely be seen in basic courses in information ethics or technology. Currently, however, information security is often under-appreciated in libraries. It is recommended that steps be taken in all libraries to assess and minimize information security risks.

NOTES

1. BUGTRAQ is a mailing list for discussion of security exploits and solutions. To subscribe, send "subscribe BUGTRAQ Your Full Name" to listserv@securityfocus.com.

2. PacketStorm Security hosts a collection of security software. This resource is for people who want to test available methods of intrusion against a particular system. If library personnel are not using these types of programs regularly, a potential intruder may be! See http://packetstorm.securify.com/.

3. The ALA has taken a firm stance against limiting computer-based activities. Minow (1997) analyzes Web content filters in libraries, suggesting a possible role for zoning certain types of content in particular areas of the library.

REFERENCES

ALA (American Library Association). 1996. Intellectual Freedom Manual. Chicago: American Library Association.

ALA (American Library Association). 1998. "Access to Electronic Information, Services, and Networks: An Interpretation of the Library Bill of Rights." Chicago: American Library Association. Available online: http://www.ala.org/alaorg/oif/electacc.html.

Chaney, Michael & MacDougall, Alan F. (Eds.). 1992. Security and Crime Prevention in Libraries. Brookfield, Vermont: Ashgate Publishing Company.

Chapman, D. Brent & Zwicky, Elizabeth. 1995. Building Internet Firewalls. Sebastopol, California: O'Reilly and Associates.

Davies, J. Eric. 1992. "Computer Misuse." In Chapman & Zwicky, op cit.

Denning, Peter J. 1990. Computers Under Attack: Intruders, Worms, and Viruses. New York: ACM Press.

Garfinkel, Simson & Spafford, Gene. 1996. Practical UNIX & Internet Security, 2nd ed. Sebastopol, California: O'Reilly and Associates.

George, Susan C. 1994. "Emergency Planning and Management in College Libraries." Chicago: Association of College and Research Libraries.

Minow, Mary. 1997. "Filters and the Public Library: A Legal and Policy Analysis." First Monday 2:12. Chicago: University of Illinois at Chicago. Available online: http://www.firstmonday.dk/issues/issue2_12/minow/

Ogg, Harold C. 1997. Introduction to the use of computers in libraries. Medford, New Jersey: Information Today, Inc.

Radcliff, Deborah. 1999. "Job Seekers' Best Bets." Computer World. September 13, 1999, pp. 50-51.

Shumann, Bruce A. 1999. Library Security and Safety Handbook: Prevention, Policies, and Procedures. Chicago: American Library Association.