95385

PRIVACY PROTECTION AND TECHNOLOGY

Robert G. McKenzie
Audit Manager
Logistics and Communications Division
U.S. General Accounting Office

Information, while intangible, represents one of our Nation's most important and valuable assets. For without information, National goals cannot be established or programs created or administered. The Federal Government -- indeed governments at all levels function on the basis of information. It is precisely because of the criticality of information to the governmental process, that agencies have over the years requested -- no demanded -- more and more information from the public and have pressed into use the most advanced technology for its processing and storage. The use of computer technology and the concentration of information has given rise to a growing public and private concern over the potential for misuse and the invasion of privacy of the individual citizen.

The capability of computers to store vast amounts of readily-usable information has given the privacy issue new dimensions. For example, Dr. Grosch stated in June 1974 the following:

Presented at GAO Briefing on Automatic Data Processing, Feb. 13, 1979.

"We can store three trillion binary digits, a five hundred word dossier for every man, woman, and child in the United States, in a commercially available machine small enough to go in an elevator * * *. Only the enormous expense of setting up data banks in the first instance holds us back from recording everything about everybody and keeping it forever." 1/

Considering the amounts of information about individuals currently maintained by various Government agencies, such as the Internal Revenue Service, Social Security Administration, Veterans Administration, and others - it is obvious that large data banks of personal information exist today. Only the current inability to centrally access all of this information precludes the use of these sources to establish comprehensive individual dossiers.

Central access to information can be made possible through various methods, such as computer networking or physically consolidating data bases at a single computer facility. The first attempt to centralize Government-held computerized information was made in the mid-1960s with the proposal of research organizations to establish a National Data Center for the systematic collection of economic microdata. This proposal was supported by the then Bureau of the Budget, but met with concern over the potential for a large concentration of data which, if misused, could result in an invasion of individual privacy.

A special subcommittee of the House Committee on Government Operations was formed to investigate the National Data Center proposal. Its concerns were expressed in the following statement:

> "What we are looking for is a sense of balance. We do not want to deprive ourselves of the rewards of science * * *. We would like to know just what information would be stored in a National Data Center; who would have access to it; who would control the computers; and most importantly how confidentiality and individual privacy would be protected * * *." 2/

The Congressional response to the proposed National Data Center was summarized in a 1968 report by the House Committee on Government Operations. 3/ The Committee concluded that the data center concept posed serious problems regarding the collection, use, and security of personal information. It strongly advised against establishing a National Data Center until the technical feasibility of protecting automated files could be fully explored and privacy guaranteed.

The Joint GSA-Agriculture Computer Acquisition Project, which ultimately became known and received notoriety as the Federal Information Network or FEDNET, had a different objective, but met similar opposition. There was widespread concern when the Congress learned of the project because it

had not been fully informed of plans for a project of its size and because of implications that the project could be expanded to link computers in the Government. This in turn could pose a serious threat to the privacy of individuals involved in any Government operation or program. As a result, the scope of the project was reduced in July 1974 by canceling the telecommunications network and GSA's primary and optional data processing installations.

GAO issued a report to the Congress in June 1975 which identified deficiencies in the agencies' procurement planning, including the determination of data processing, communications, and privacy-security requirements. 4/ As a result of a Congressional limitation on spending, in October 1975, Agriculture's planned procurement and the request for proposals were withdrawn.

More recently, the Internal Revenue Service's proposed new system came under scrutiny by both the Congress and the administration. This proposed system, which was known as the Tax Administration System or TAS, called for extensive use of interactive online processing and the decentralization of the tax account master files from the IRS National Computer Center to 10 existing service centers. The network would have employed over 8,000 interactive terminals and approximately 5,000 direct data entry terminals.

GAO reviewed the proposal and concluded that the proposed Tax Administration System could provide a high level of protection for taxpayer information if the system was properly designed and implemented, and if the weaknesses in safeguards cited in our report were corrected. 5/ Nevertheless, the project was terminated early last year with privacy as one of the major issues.

I could cite other cases, such as the bid protests that have been upheld because of privacy-security issues. 6/ and 7/ The concerns over privacy and related security issues have had an effect on computer acquisitions in the past, but in our opinion - the full impact has yet to be felt. Let me explain why.

The various civil agencies are just now beginning to address their security requirements at a level above their basic physical security needs. This is evidenced by such activities as the Social Security Administration's symposium on privacy and security which was held last April; the National Conference on Fraud, Abuse, and Error convened by HEW in December; and the recent Automation Security Workshop conducted by the Army. Such meetings have drawn attention to some of the problems associated with protection of information and have served as a forum for the dissemination

of real and potential solutions. However, there is still much to be done.

Most agencies have yet to implement an effective security program. In a recent GAO report, we noted an absence of top management involvement, with a resultant lack of organizational structures, policies, planning and procedures which are necessary for the funding, development and implementation of effective security programs. 8/

Although many agencies are engaged in various aspects of computer security, we found numerous weaknesses including:

--absence of risk management techniques to select proper security safeguards,

--lack of procedures for monitoring and reporting on computer security effectiveness,

--noncompliance with existing procedures,

--a lack of effective security training,

--limited involvement on the part of internal audit, and

--specific weaknesses in physical, technical, and administrative safeguards for protecting agency data.

As long as these deficiencies exist, Federal agencies have no assurance that their computer resources and data are properly secured or adequately protected.

The Office of Management and Budget, the National Bureau of Standards, and the General Services Administration have issued various guidance for use by the Federal agencies in managing their computer activities. However, far more definitive guidance needs to be promulgated in the area of privacy-security. The Privacy Protection Study Commission stated in Appendix 5 to their report:

> "Setting forth broad public-policy objectives while allowing for various implementation alternatives and strategies does, however, create a need for reasonably definitive guidance to operating personal on what constitutes acceptable levels of performance in certain areas.
>
> * * * * *
>
> "The problem yet to be addressed in any broad and effective way, at either the State or Federal level, is how to translate the broad social goals of privacy and fair information practice legislation into precise steps which computer scientists and managers of automated systems may follow in order to achieve acceptable levels of performance." 9/

Once such precise steps have been defined -- once the criteria have been established -- then and only then will uniform programs and procedures emerge that will provide appropriate levels of protection for personal and other sensitive information.

In our report to the Congress on "Challenges of Protecting Personal Information in an Expanding Federal

Computer Network Environment" 10/ we recommended that the Director of OMB take the necessary actions to expeditiously provide the Federal agencies with comprehensive guidelines that:

--contain the definitions and criteria necessary to permit an assessment of their security requirements,

--provide the methodology to be used in conducting such assessments,

--identify the physical, administrative, and technical safeguards that should be applied in satisfying their security requirements, and

--specify the means to justify the associated cost.

I believe that the impact such guidelines could have on future procurements of computer hardware, software and services and on the manner in which they are used, is obvious. However, the question today is, what type of system can be obtained now that will provide a high level of protection for personal or other sensitive information? The report I cited discusses some of the technology that can be used to provide a high level of protection in a particular environment -- shared computer networks. (Incidentally, to my knowledge this report represents the only audit report ever

reviewed by such a prestigious scientific journal as Scientific American.) Before I discuss this technology, I would like to set the stage by discussing some of the threats to computerized data and a few of the system vulnerabilities.

The need for physical security against such hazards as fire, sabotage, and theft is well known and the subject of another GAO report. 11/ The National Bureau of Standards publication, "Guidelines for Automatic Data Processing Physical Security and Risk Management" 12/ and their document on "Automatic Data Processing Risk Assessment" 13/ should aid agencies in assessing their physical security and developing effective physical security programs. However, providing only physical security is no longer adequate for information protection. TRW Systems, Inc., in a study on computer system security, pointed out the following:

> "Third-generation computers introduced new capabilities that involved the concurrent processing on many jobs, extensive sharing of computer resources, and the use of remote terminals. While these new capabilities brought benefits of substantially lower cost, sharing of large data bases, and remote use of computers, they also introduced a complex security problem. With concurrent sharing of a computer system, the opportunity is present for inadvertent, accidental, or malicious acquisition of information by a user who has no right of access." 14/

In examining the risk to personal and other sensitive information maintained on data processing systems, it appears that the threats stem from two sources: First - authorized, but untrustworthy or dishonest users, and second - malicious penetrators. The untrustworthy user has authorized access to the data of interest, while the malicious penetrator does not. The penetrator may be an employee of the organization or an outside party.

The problem of untrustworthy or dishonest employees represents the major threat to personal or sensitive information contained in any system of records. The potential for misuse of information by individuals in positions of trust is not unique to automated data processing systems -- the problem exists in manual systems as well. Nevertheless, the concentration of data in computer systems increases the magnitude of the risk over non-computerized systems.

Protection against untrustworthy or dishonest employees is indeed difficult. However, the risk can be substantially reduced through proper application of well-designed managerial controls, which include: segregation of employee duties, personnel screening, activity monitoring, and effective auditing. These and other managerial controls have

been afforded extensive coverage in literature published over the years by universities, professional societies, and Government.

Malicious penetrators present a different threat than untrustworthy employees in that they must circumvent technical security measures. In order to place the threat from this source in perspective, it is necessary to understand how penetrators would achieve their objective and what skills they must possess.

Our study of the views of experts in the field indicates that skilled individuals generally penetrate a system by using an operating system function in a way unanticipated by designers, or by exploiting some anomalous behavior of the operating system. They are frequently aided by the fact that designers of operating systems have assumed that users will not deliberately attempt to force a malfunction of the system.

Penetrators may achieve their objectives by various methods, including (1) acquiring by any method a list of user identifiers and corresponding passwords or other confirmatory information needed to gain access to the computer system, or (2) obtaining supervisory (executive or master) control of the computer system. A number of means have been

found to do this. For example, in one version of an operating system, registers are shared between the operating system and the user's application programs. In this particular case, the operating system, in releasing a register to the user's program, uses a storage location provided by the user to load the register before turning control over to the user's program. This is accomplished without the operating system checking to ensure that the storage location is within the user's assigned area. Consequently, the operating system will load the register with eight consecutive words of memory from any location specified by the user. This flaw could be exploited to set up a search through all of the computer's memory for the password of the executive user (i.e., the master operator) which, when found, would permit the penetrator to masquerade as the executive user and have extraordinary privileges.

Using the first method, the penetrator can masquerade as any of the authorized users, while use of the second method gives him direct access and control of any file or program in the system.

In order for penetrators to accomplish their objective by either method, it is necessary that they be (1) at least moderately skilled in programming, (2) expend time and effort

to understand rather complex operating systems, and (3) have knowledge of the limitations that occur in the design and implementation of the systems. Such knowledge suggest to penetrators where to look for possible errors and design flaws. If they have access to system documentation, their ability is considerably enhanced.

Against such individuals, contemporary computer operating systems frequently fail to provide adequate protection for personal or sensitive information. The question is: Why?

It appears that this weakness is rooted in at least two causes. First, most operating systems that are available today were designed originally at a time when security issues were not being fully considered. As a result, the security elements were normally scattered throughout the operating system in a variety of apparently unrelated ways. In such systems, there is no assurance that all of the security-related parts have been examined and tested for flaws. In fact, known flaws exist in several commercially available operating systems currently in use. 15/

Second, there are no comprehensive criteria for security to guide those designing and implementing operating systems for computers. Compounding the difficulty is the fact that

security is usually stated in negative terms such as, "Data should not be accessible in an unauthorized way." Requirements that can be used in design and implementation must translate such negative statements into positive criteria which specify how a system should react under various conditions.

From my discussion, it would appear that computer systems are extremely vulnerable; and indeed they are, but there are ways to reduce the risk. If you recall, I mention earlier that in order for a penetrator to accomplish his objective - he must be at least moderately skilled in programming. The fact of the matter is, that a system or application programmer can do more damage to a system with less chance of being caught than almost any other person involved in data processing. It is, therefore, necessary to isolate the system from the programmer in order to provide any degree of security. While current research in the technical community is directed to the development of operating systems and mechanisms that will provide protection from skilled programmer attacks, there is no consensus on its achievement in the immediate future. Today, it is possible to attain a high level of data security by (1) reducing the threat from those individuals with the technical training

necessary to circumvent security safeguards and (2) segregating sensitive data and its processing from all other data, hence the adoption of a policy of isolation.

An isolation policy can be applied in either of two ways: (1) by isolating the system from the threat or (2) isolating sensitive data within the system. Let us examine the former method first.

Generally, the risk of a successful penetration increases with the capability provided to users of the system. Most multiuser teleprocessing systems attempt to provide the user with maximum capability under the premise that this makes the system more desirable and useful. Such systems can be highly vulnerable to penetration.

In order to significantly reduce the risk, the users' capability must be sharply curtailed. This can be done by permitting the terminal users to process transactions while removing their programming capability. Such a system--termed a transaction system--can, if properly designed and implemented, effectively isolate the system from the threat posed by the programmer.

An airline reservation system is an example of a transaction system. The terminal operator can enter, change, and retrieve data according to a limited number of command codes.

Each command code performs a specific function in relation to the information entered and the data maintained on the system. For example, one command code assigned to a reservation clerk may cause all available flights between two cities to be displayed, while another may reserve a seat on a particular flight.

The users' capability in a transaction system can be further reduced through the use of employee and terminal profiles. Such profiles can restrict the command codes and terminals an employee can use to only those necessary to perform specifically assigned duties. For example, a cargo clerk and the computer terminals located in the air freight department may be denied the use of command codes necessary to access passenger reservation information.

While this limits terminal users to transaction processing, it is also necessary that programs and their modifications be placed on the system under highly controlled conditions. Here it is necessary to isolate programmers from the system by requiring all programs and program changes to be submitted to an independent test and evaluation group. This group, which is a buffer between the application and system programmer and the operational programs, controls the programming function

by reviewing, validating, and approving all programs and program changes placed on the system. Where it is impractical to establish a formal and independent test and evaluation group, such as in a small organization or where the programming function is relatively small, mandatory peer review can provide a measure of control.

This approach provides a high level of protection to personal information by isolating the system from the programmer and reducing the risk by restricting the user to only those functions necessary to process authorized transactions. Where users are presented with only the functionality of one or more transaction systems, the security of such systems can be developed without necessarily relying on security features and mechanisms supplied by a vendor. Therefore, the security of a transaction system is dependent upon the adequacy of the design, operating procedures, and program testing.

Where user requirements demand the flexibility of normal programming capability, the policy of isolation requires effective separation of sensitive data and its processing from all other data. To accomplish this, isolation mechanisms must be present that cannot be bypassed by users exercising normal user-programming control of the

system. In this context, user-programming control extends to all of the supervisory, monitor, or operating system programs executed on behalf of a user's program.

Some of the basic architectural approaches, which are within the current state-of-the-art and provide this type of isolation include (1) virtual machine systems, (2) descriptor based systems, and (3) systems utilizing the kernal concept.

Virtual machine systems create an isolated environment through techniques which have the effect of creating for each user a complete system dedicated solely to the user's purpose. The software that creates this environment is generally known as a "virtual machine monitor." The monitor consists primarily of programs that provide (1) interpretive execution of privileged instructions, (2) minimal controls to initiate and discontinue virtual machines, and (3) the controls to cause several virtual machines to function in a single set of hardware.

The monitor permits each user to functionally have an operating system restricted to the individual user. Ignoring cost considerations, each user could have a unique operating system and thus completely close off any possibility of interaction between any two users of the system.

The following illustrates how conventional systems and virtual machine systems differ:

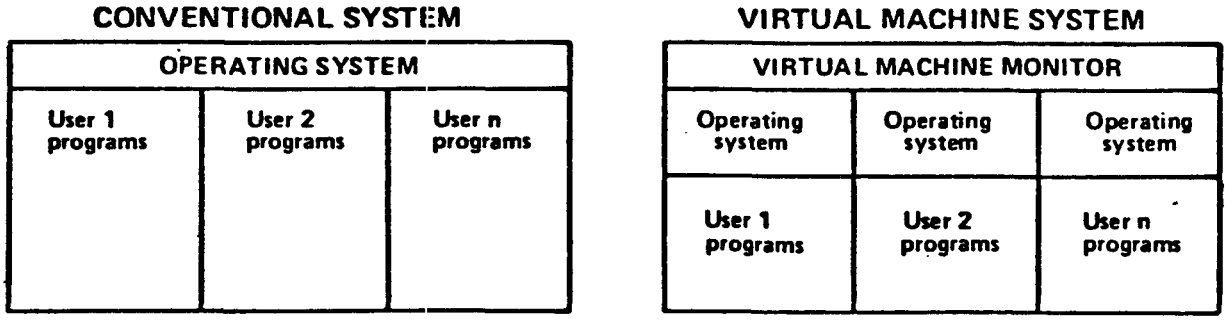

FIGURE 1. VIRTUAL MACHINE VIS-A-VIS CONVENTIONAL SYSTEMS

From a security viewpoint, a well-designed virtual machine system provides protection from a malicious programmer by isolating the operating system of two or more users. It also reduces the need to be concerned with the security-worthiness of an existing operating system because the operating system can be considered as belonging to a single user.

Due to the limited functions the monitor performs, it can be quite small and relatively simple compared to typical operating systems. Therefore, it is theoretically possible to subject the monitor to thorough testing and validation of design.

Virtual machine technology is available today and can be applied to many existing systems with only minor hardware modifications. However, the use of this technology has several disadvantages that could limit the circumstances in which the virtual machine concept is applicable. Two major disadvantages are that (1) the overhead burden associated with virtual machine systems can add materially to the system's operational cost and (2) the virtual machine approach does not adequately provide for high-volume sharing of data, or computer programs, among users.

Another approach to isolating users is to use descriptor architecture to provide each user with a totally independent address space.

An absolute address within a computer is a specific designation assigned to a storage location by the machine designer. Indirect addressing is a method of computer cross referencing in which one memory location contains information as to the absolute address of an object or where such an address can ultimately be found. A descriptor can be described as a computer word that acts as a form of an extended indirect address.

When a descriptor is referenced by a computer program, information contained in the descriptor is interpreted in

hardware to control the completion of the reference. It is, therefore, possible to represent an object's protection requirements in its descriptor and be assured that there will be automatic hardware controlled validation.

The major benefit from use of a descriptor-controlled approach is the ability to control sharing of programs or data by including the object to be shared as a descriptor in the sharing program with the descriptor containing the protection information as to how the object may be referenced --such as read only for execution, read-only, write, append, etc.

The following diagram is a simplification of how controlled sharing can be accomplished in descriptor-based systems. As indicated, each user program can (1) execute the operating system service functions and its own code within the addressing context established by the descriptor table, (2) can call on the operating system resource management functions, and (3) read and write its own data. Common library programs can also be shared among different programs as can data. With a descriptor capability, a variety of systems can be developed that will provide a high level of data protection.

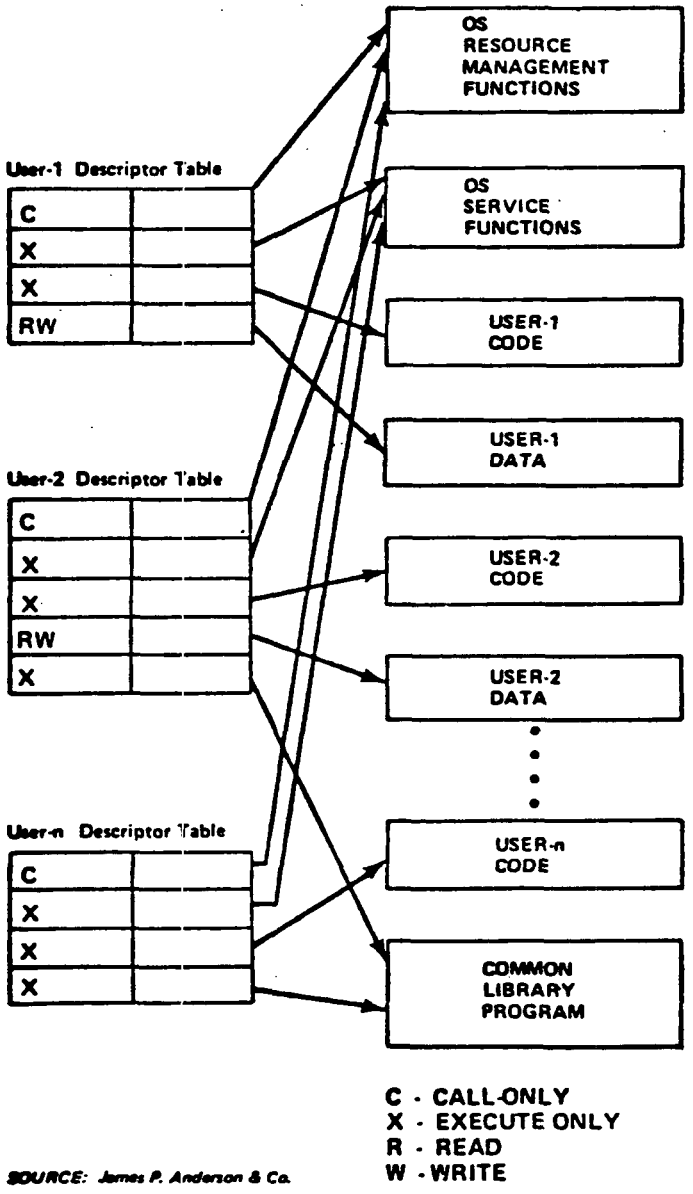

FIGURE 2. CONTROLLED SHARING IN A DESCRIPTOR-BASED SYSTEM

Another approach employs what has been termed "protected domains." Most third-generation computer systems support two domains--a privileged supervisor state and a non-privileged

problem state. In the privileged state, access rights are defined for such functions as scheduling and allocating the system's resources. In the non-privileged or problem state, the central processing unit cannot execute input/output and other privileged instructions. While these two protection domain mechanisms could theoretically provide a basis for security against deliberate subversion of the system, in practice, the problems of securing a computer system are so complex that many researchers have concluded that more sophisticated protection mechanisms are needed.

One approach that has been taken is increasing the number of protection domains, thus adding additional barriers that must be circumvented for a successful penetration. A three-domain approach has been used to structure the system into three environments--the user environment, the operating system environment, and the kernel environment--as shown in the following illustration.

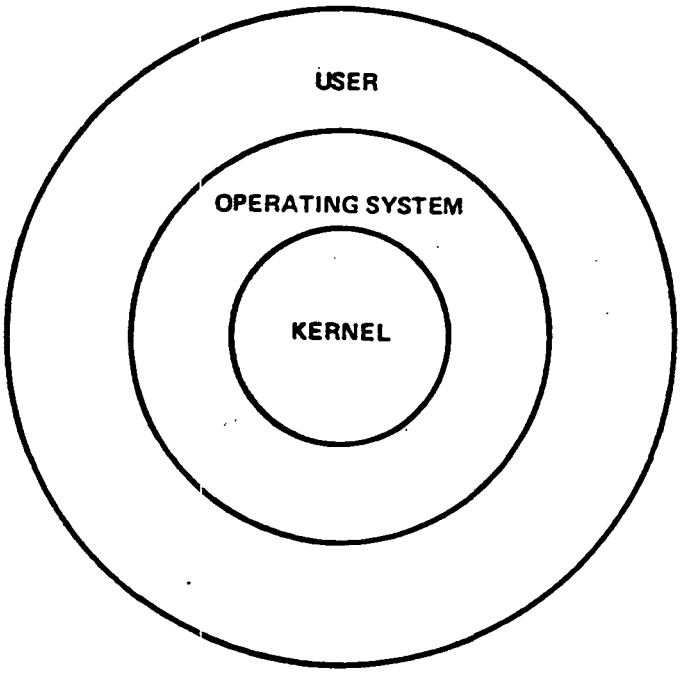

FIGURE 3. THE KERNEL CONCEPT

To be effective, all security-sensitive elements of an operating system must be located in the security kernel. The kernel is placed at the very highest level of protection, or innermost section of the system, while other functions of the operating system are placed at lower levels. Under this approach, the proper functioning of protection is not dependent upon the behavior of the outer layers (or less protected levels) of the system.

A methodology exists for proving the security worthiness of operating systems employing the kernel concept but has not been applied to large systems. Research is continuing in this area at several locations.

Operating systems have been developed with security as one of the objectives and have attempted to achieve this objective by creating a foundation, through good design, for establishing the reliability of existing security features. In these systems, the security-sensitive elements are normally scattered throughout the system as opposed to the grouping of such elements as discussed above in the kernel concept. The reliability of such systems can be demonstrated only through detailed study and use of penetration techniques. Unfortunately, successful penetration proves the presence of protection failures, but does not prove that all flaws have been detected. Nevertheless, these operating systems are purported to offer a higher level of protection than other such generally available systems.

CONCLUSIONS

Absolute security is in fact unobtainable. However, the proper use of current technology can provide a high level of protection for personal and sensitive information. It seems logical that further progress toward more secure hardware and software will be accelerated to the extent that management recognizes their security needs and places such demands upon the computer industry.

REFERENCES

1. DATA SECURITY AND DATA PROCESSING, VOL. 3, PART 1, "STATE OF ILLINOIS: EXECUTIVE OVERVIEW," IBM CORPORATION: WHITE PLAINS, N.Y., 1974, P. 7.

2. HEARINGS BEFORE A SPECIAL SUBCOMMITTEE ON INVASION OF PRIVACY, HOUSE COMMITTEE ON GOVERNMENT OPERATIONS, 89TH CONG., 2D SESS., (1966), P. 3.

3. HOUSE COMMITTEE ON GOVERNMENT OPERATIONS, REPORT: PRIVACY AND THE NATIONAL DATA BANK CONCEPT, 90TH CONG., 2D SESS., H. REPT. NO. 1842, (1968), P. 8.

4. REPORT OF THE COMPTROLLER GENERAL OF THE UNITED STATES, "IMPROVED PLANNING--A MUST BEFORE A DEPARTMENT-WIDE AUTOMATIC DATA PROCESSING SYSTEM IS ACQUIRED FOR THE DEPARTMENT OF AGRICULTURE," (LCD-75-108), JUNE 3, 1975.

5. REPORT OF THE COMPTROLLER GENERAL OF THE UNITED STATES, "SAFEGUARDING TAXPAYER INFORMATION--AN EVALUATION OF THE PROPOSED COMPUTERIZED TAX ADMINISTRATION SYSTEM," (LCD-76-115), JAN. 17, 1977.

6. COMPTROLLER GENERAL DECISION: PRC COMPUTER CENTER, INC., ET AL., 55 COMP. GEN. 60 (1975).

7. COMPTROLLER GENERAL DECISION: COMPUTER NETWORK CORPORATION ET AL., 56 COMP. GEN. 245 (1977): AS MODIFIED BY COMPUTER NETWORK CORPORATION, ET AL., 56 COMP. GEN. 694 (1977).

8. REPORT OF THE COMPTROLLER GENERAL OF THE UNITED STATES, "AUTOMATED SYSTEM SECURITY--FEDERAL AGENCIES SHOULD STRENGTHEN SAFEGUARDS OVER PERSONAL AND OTHER SENSITIVE DATA," (LCD-78-123), JAN. 23, 1979.

9. REPORT OF THE PRIVACY PROTECTION STUDY COMMISSION, APPENDIX 5, "TECHNOLOGY AND PRIVACY," JULY 1977.

10. REPORT OF THE COMPTROLLER GENERAL OF THE UNITED STATES, "CHALLENGES OF PROTECTING PERSONAL INFORMATION IN AN EXPANDING FEDERAL COMPUTER NETWORK ENVIRONMENT," (LCD-76-102), APRIL 28, 1978.

11. REPORT OF THE COMPTROLLER GENERAL OF THE UNITED STATES, "MANAGERS NEED TO PROVIDE BETTER PROTECTION FOR FEDERAL AUTOMATIC DATA PROCESSING FACILITIES," (FGMSD-76-40), MAY 10, 1976.

12. U.S. DEPARTMENT OF COMMERCE, NATIONAL BUREAU OF STANDARDS, "GUIDELINES FOR AUTOMATIC DATA PROCESSING PHYSICAL SECURITY AND RISK MANAGEMENT," FEDERAL INFORMATION PROCESSING STANDARDS PUBLICATION 31, JUNE 1974.

13. U.S. DEPARTMENT OF COMMERCE, NATIONAL BUREAU OF STANDARDS, "AUTOMATIC DATA PROCESSING RISK ASSESSMENT," NBSIR 77-1228, MARCH 1977.

14. RICHARD B. BLUE, SR. AND GERALD E. SHORT, "COMPUTER SYSTEM SECURITY TECHNOLOGY AND OPERATIONAL EXPERIENCE," TRW SYSTEMS, INC.: REDONDO BEACH, CA, (REPORT NO. TRW-SS-74-15), MARCH 1974.

15. R.P. ABBOTT, ET AL., "SECURITY ANALYSIS AND ENHANCEMENTS OF COMPUTER OPERATING SYSTEMS," U.S. DEPARTMENT OF COMMERCE, NATIONAL BUREAU OF STANDARDS, (NBSIR-76-1041), APRIL 1976.