COMPUTER DATA BANKS AND SECURITY CONTROLS

W. H. Ware

March 1970

P-4329

COMPUTER DATA BANKS AND SECURITY CONTROLS

W. H. Ware

Resumé

There is not today an established "data bank industry", nor are the technical risks of operating computer-based data banks widely understood. Moreover, the financial incentive of the data bank operator favors the user of the bank rather than the private individual. For these reasons, it is argued that strong government intervention and control is necessary to protect the privacy and reputation of individuals. The state-of-art for information safeguards in computer systems does not permit a handbook approach to the subject; only general principles and guidelines can be stated. Suggestions are made for controls that can protect information within the computer and govern its divulgence to authorized users. Ideas are proposed for the role of the government agency in the matter, its interface with and control over operators of data banks, and inferentially, the need for relevant legislation. The individual's position is defined, and his expectations identified. Views are expressed on the responsibility and liability of data bank operators and data sources relative to such things as identification of legitimate users of the bank, divulgence of information, and accuracy and completeness of information.

COMPUTER DATA BANKS AND SECURITY CONTROLS

W. H. Ware*

The RAND Corporation, Santa Monica, California

TORONTO CONFERENCE - MAY 1970

The issue is to safeguard the privacy and reputation of the individual by guaranteeing that information about him in computer files is not revealed indiscriminately. This implies that information contained in computer files must be protected and released only to authorized users. In discussing this matter, my point of view will be that we must accept data banks as desirable and as serving a useful purpose for society. Certainly, existing computer data banks on criminal activities contribute to better law enforcement; existing credit reference data banks assist society by making it more difficult to pass bad checks which, in turn, means a financial saving to society. Thus, I will not discuss the fundamental question of whether data banks are desirable or not; but I will consider what can be done to protect and control dissemination of information in them. This discussion cannot be a how-to-do-it handbook; the general state-of-the-art is not yet to that point. Rather, this paper is a collection of ideas for consideration and debate.

Perhaps the largest and best organized systems for protecting information are those devised by governments to safeguard national defense information and national secrets. Since such systems have been in existence for a long time, it will be instructive to consider them both for insight and as a framework for thinking about the problem. They are usually established through legislation and/or executive order, and typically include the following features:

1) Classes of information are defined whose divulgence is considered a threat to the interests of the country. The threat is defined for each class.
2) Procedures are defined for establishing that individuals are trustworthy to receive information.
3) The principle is established that an individual receives information only if it is necessary for performance of his job.
4) Procedures are created both for controlling dissemination and for protecting information, no matter in what form it may be recorded; e.g., documents, magnetic tapes, etc.
5) Penalties are defined for deliberately revealing information to unauthorized recipients.

Data banks may not need as full a treatment as accorded national defense data. However, we should consider the implications for data banks of those features that have been found desirable to protect national secrets.

Consider what is probably the most hazardous circumstance: a data bank that receives information from specified sources; that is maintained in a computer at a central location; and that serves users who are remotely connected to the computer by communication lines. The several parts of the problem are as follows:

1) Information that is inserted into the data bank must either be known to be correct, or have some level of confidence attached to it.
2) Information, once within the data bank, must be revealed only to individuals authorized to receive it.
3) Identity of individuals requesting information from the data bank must be established before the information can be released.

First, it must be understood that the problem is a system one, which must be attacked from a system engineering point of view in the broadest sense. If handled in a bits-and-pieces fashion, the finest of safeguards in one part of the system can easily be circumvented by loopholes elsewhere.

Consider item 2 first. The protection of information within a computer system has already received attention in the context of national defense information, and a few such systems are operating with appropriate security safeguards. There is, therefore, an initial set of ideas that have been formulated and to which technical attention has been given.

To outline what is necessary, consider a data bank of information that is within a computer system; is considered to be valid; and is to be dispensed to users on request. There are five points that need attention:

1) Obviously, physical protection must be afforded the computing central and demountable storage media. All the safeguards in the world will be to no avail if magnetic tapes or magnetic disks can be stolen or copied by unauthorized persons.

2) Ideally, the communications should be protected by some form of encryption or physical protection of the circuits. Practically, this may not be essential because the amount of data on any one communication line will probably be small. On the other hand, wiretaps are very easy, and penetrating the system through its communication circuits is a serious threat that the system designer must consider.

3) A multi-user system (especially one which is accessed through a remote console) is, in effect, a timesharing system that itself must have appropriate computer hardware safeguards. There may be needed--depending on the precise details of the application--bounds registers to segment the memory, an interrupt system to control activities within the computer system, memory protect features,

and two internal modes of machine operations, one of which is privileged to the monitor program.

4) Software safeguards must also be provided. There must be a mechanism to control user access to the files. There must be audit trails to keep track of what users are doing and what data each has asked for. There must be mechanisms for alerting operations personnel to unusual situations, especially marginal conditions in the hardware or malfunctions of the software. There must be mechanisms for the system to self-test itself to guarantee continuity of the hardware and software safeguards.

5) The system's administrative and management controls must be security conscious, and include such things as:

   - o Provisions for monitoring and controlling the action of system operators;
   - o Procedures for loading certified copies of the software and verifying that it did correctly load;
   - o Procedures for controlling movement of and physical access to demountable files.

These five points have received considerable technical attention because each is relevant to currently operating systems that contain defense classified information; also,

each in some measure is an aspect of resource-sharing systems in general.

Notice in passing that certifying that the software is correct, completely designed, and contains no unanticipated paths through it is a major technical problem. It is one thing to establish that software will do what it is supposed to do. It is quite a different and more difficult thing to prove that it does not do what it is not supposed to do, especially when hardware malfunctions or unusual user actions occur. Once it is certified to be correct, the software must be guarded against unauthorized changes.

Now let us discuss points 1 and 3, which are larger issues of concern to this conference. These are intimately connected with directly protecting the privacy and reputation of an individual.

Given the initial assumption that data banks serve useful purposes for the public, are cost effective, and will be in existence, it follows that each individual wants to make certain that: 1) information in the bank about himself is correct; 2) information is divulged only to those who will use it in his interest or to his benefit; and, 3) he has recourse for damages in the event the users or operators of the data bank willfully or negligently mishandle the information.

Even though technical safeguards can help enforce these principles, I feel that ultimately they will have to be

passed in law. Therefore, government intervention is necessary. It follows that enforcement and monitoring of the law will be necessary, and I would, at least in the near term, center that responsibility in a governmental regulatory body. This may seem a strong position, but I will support it later. Furthermore, I would rather begin too strongly and weaken controls as experience shows it possible, than recover from awkward oversights after the fact.

Before an owner and operator of a data bank could be licensed, so to speak, I would ask that he demonstrate to an appropriate regulatory body such things as the following:

1) The nature and purposes of his data bank; the use to which the data will be put; and the general class of customers it will serve.
2) Precise identification and description of the data sources on which it will draw, and the checks that will be applied to validate the information from the sources.
3) A complete description of the safeguards in the system (physical, hardware, software, communication, personnel, and administrative/management) that protect information and control its divulgence.
4) A complete description of the procedural safeguards (software or manual) to edit source information for errors, to assure posting information to correct dossiers, to resolve ambiguity in identification of an individual, to treat information of doubtful

validity, and to establish confidence levels on information derived or inferred from fragmentary data.

5) A complete description of the audit processes incorporated in the system, and the audit information that will be made available for periodic review.

6) The mechanism whereby an individual can review his dossier and the sources from which the dossier was compiled, and challenge its contents and correct errors.

7) The tests and inspections that he has performed on the system to assure that it does operate properly, and especially that the software has been verified completely designed.

It is obvious from my position that I feel that a government agency not only must carefully investigate proposals for data bank business, but that it must also audit such business from time to time to assure continuity of safe and legal operation. The job of the regulatory agency is partly highly technical, and relevant expertise must be available. It is clear that the closeness and depth of inspections and investigations must depend on the nature of the data bank. For one which contains information that cannot seriously harm an individual, governmental intervention can be minimal; but for one which contains very extensive dossiers on individuals, the control must be

correspondingly greater. Since governments themselves are talking of establishing data banks, what I have implied for private operators should apply to government agencies.

Next, let us turn our attention to the users that the system serves. First, what determines who they are. Most data banks will sell services; thus, the nature of the bank and the aggressiveness of its marketing will tend to identify the user group. The operator must accept prime responsibility for certifying that his users are as they represent themselves. It would seem desirable to require a business man to present the usual credentials of his business status (e.g., business licenses, offices, staff, equipment, etc.) before being accepted as a customer of the system. The provisions of communications' secrecy acts would seem to be applicable since users will receive information as a privileged communique and should therefore be liable for willful or negligent transfer to other parties. If the user is another data bank, the operator of the first data bank must take additional safeguards. Audit trails must be maintained so that he knows where copies of any or all parts of data exist in computer files, and he must accept responsibility for updating or correcting such copies promptly and responsively. Conversely, if he receives data from another data bank, he must keep audit information so that original sources can be identified at a later date. This could be crucial in the event of damage suits in which the

operator's liability should be shared with data sources, be they other data banks, individuals, other businesses, or government sources. I believe that data sources cannot be anonymous and thus immune from legal action; they must accept responsibility for carelessness or negligence.

From the individual's point of view, there must be appropriate legislation enabling a person who has been maligned or damaged because of the activities of a data bank to take prompt legal action, and to seek redress against the users, operators, or data sources. There are special situations where the individual probably should have a legal, court-created document certifying that some action has been taken. For example, consider the person who has been arrested and accused of a felony; later, however, he is acquitted. This fact may well find its way into his credit reference file and he should have some positive confirmation from the data bank that his arrest experience has been expunged from all copies of his credit file.

This may all seem overwhelming and too much, but I have tried to explore the worst-case situations. I have tried to suggest some kinds of information safeguards that could be implemented, and may have to be done. Certainly, there is no general recipe that will *a priori* describe the controls relevant to every data bank. Depending upon the data it deals with, the completeness of records on each individual, the users it serves, the threat from subversive

penetration, etc., specific protective mechanisms and procedures will have to be evolved.

There are some general observations that are relevant to my position that strong governmental controls are desirable. Presently, there is no "data bank industry" as there is an automobile industry or a motion picture industry. There are no trade organizations; and thus, self-policing is not likely. Furthermore, the financial incentives of the data bank operator favor the user. It is from the user that the operator derives his revenue, and the individual is hard put to cause the operator serious financial damage. Business is unavoidably profit oriented, so there is no substantial intrinsic motivation for the operator to surround his data bank with a complete set of information safeguards. Moreover, an operator may be technically ignorant of the risks in his system, or unaware of the ease with which it can be penetrated.

Finally, consider what can happen if data banks proliferate widely and without control. We see all around us situations that were recognized after the fact and are now out of control and harmful to society; the many kinds of pollution are a prominent example. Protection of the individual's privacy and reputation is already recognized as essential to society's health; I would rather not have data banks become the problem that pollution has. Thus, my view is that we should vigorously and aggressively formulate appropriate

safeguards, mechanisms, and legislation. Let's try to be ahead of the situation before it is too late.