Information Technology and Dataveillance

Roger Clarke

Principal, Xamax Consultancy Pty Ltd, Canberra

Visiting Fellow, Department of Computer Science, Australian National University

Version of November 1987

© Association for Computing Machinery Inc., 1988

This paper was published in Commun. ACM 31,5 (May 1988) 498-512, and re-published in C. Dunlop and R. Kling (Eds.), 'Controversies in Computing', Academic Press, 1991

This document is at http://www.anu.edu.au/people/Roger.Clarke/DV/CACM88.html

This paper was reviewed by J. Fendrich in Computing Reviews

Abstract

The concept of 'dataveillance' is introduced, and defined as the systematic monitoring of people's actions or communications through the application of information technology. Dataveillance's origins are traced, and an explanation provided as to why it is becoming the dominant means of monitoring individuals and populations.

The paper identifies, classifies and describes the various dataveillance techniques. It then examines the benefits, and especially the dangers, arising from dataveillance. It considers the intrinsic and extrinsic controls that act to keep the application of dataveillance under control, and suggests some appropriate policy measures.

Contents

INTRODUCTION

Concern about freedom from tyranny is a trademark of democracy. Between 1920 and 1950, the anti- utopian novels of Zamyatin [78], Kafka, Huxley [21], and Orwell [45] unleashed a visionary, yet paranoiac "literature of alarm" (see, e.g., [5], [12], [15], [19], [32], [36], [49], [53], and [62].

Surveillance is one of the elements of tyranny. The word conjures up unpleasant visions of spies, repression of individuals, and suppression of ideas. Nevertheless, some classes of people, at least when they undertake some classes of activity, are deemed by

suppression of ideas. Nevertheless, some classes of people, at least when they undertake some classes of activity, are deemed by society to warrant surveillance. Few would contest that people reasonably suspected of terrorism and organized, violent crime are candidates for surveillance. Meanwhile, the growth in crimes against property has resulted in the widely acclaimed "neighborhood watch" movement.

The computer has been accused of harboring a potential for increased surveillance of the citizen by the state, and the consumer by the corporation. Most accusations have been vague, asserting that harm will result, rather than showing the mechanisms by which it will come about. Some have even claimed that the potential is already realized: "It is possible . . . to imagine what one might call a `central clearing house' for mass surveillance and control. without straining the limits of present- day [i.e., 1974] technology and organisational skills . . . . [A]ll major agencies would render unlimited assistance to one another. Information generated in the relationship between a client and any one system would automatically be available to any other system . . . . [T]he client's contact with one would have the effect of contact with all . . . [N]o favourable decision from any agency would be implemented while there remained a dispute between the client and another agency" [55, p. 319].

Apart from research by Kling [23, 24] and Laudon [27-31], there has been little discussion in the computing literature. Most of the important contributions have been by observers rather than practitioners of computing. particularly Rule [55-59]. Marx [33, 34], and Reichman and Marx [35, 50]. See also [2, 11, 52].

The purpose of this article is to make the work of such authors more readily accessible to computing practitioners and academics, to extend it somewhat, and to propose a framework for policy. It commences by clarifying the concept of *dataveillance*, and then describes the manner in which information technology (IT) is stimulating its development. Popular publications have tended to deal with the topic in colorful, at times even hysterical, fashion (see, e.g., [1], [4], [6], [7], [12], [13], [37], [64], [70], and [73]). To enable the problems to be appreciated and responded to rationally, I will attempt to deal with the topic in a more neutral and dispassionate manner. In particular, I explicitly reject the notion that surveillance is, of itself, evil or undesirable; its nature must be understood, and society' must decide the circumstances in which it should be used, and the safeguards that should he applied to it.

SURVEILLANCE

The Oxford Dictionary explained surveillance as "watch or guard kept over a person. etc., esp. over a suspected person, a prisoner, or the like; often spying, supervision; less commonly, supervision for the purpose of direction or control, superintendence" [46]. The oldest usage noted was in the "Committee of Surveillance" immediately after the French Revolution. *Webster's 3rd Edition* defines it as "1. close watch kept over one or more persons: continuous observation of a person or area (as to detect developments, movements or activities); 2. close and continuous observation for the purpose of direction, supervision or control" [71].

Rule uses the term for "any form of systematic attention to whether rules are obeyed, to who obeys and who does not, and how those who deviate can be located and sanctioned" [55, p. 40], and later as "the systematic collection and monitoring of personal data for the purpose of social control" [59, p. 47] (see also [20, p. 90]. In this article the following definition is used:

Surveillance is the systematic investigation or monitoring of the actions or communications of one or more persons. Its primary purpose is generally to collect information about them, their activities, or their associates. There may be a secondary intention to deter a whole population from undertaking some kinds of activity.

The basic form, physical surveillance, comprises watching and listening (visual and aural surveillance). Monitoring may be undertaken remotely in space, with the aid of image- amplification devices like field glasses, infrared binoculars, light amplifiers, and satellite cameras, and sound- amplification devices like directional microphones; and remotely in time, with the aid of image and sound- recording devices. Several kinds of communications surveillance are practiced. including mail covers and telephone interception. The popular term electronic surveillance refers to both augmentations to physical surveillance (such as directional microphones and audio bugs) and to communications surveillance, particularly telephone taps.

These forms of direct surveillance are commonly augmented by the collection of data from interviews with informants (such as neighbors. employers. work- mates, and bank managers). As the volume of information collected and maintained has increased, the record collections (or personal data systems) of organizations have become an increasingly important source.

***Dataveillance* is the systematic use of personal data systems in the investigation or monitoring of the actions or communications of one or more persons.**

The terms *personal surveillance* and *mass surveillance* are commonly used, but seldom defined. In this article the following definitions are used:

> **Personal surveillance** *is the surveillance of an identified person. In general, a specific reason exists for the investigation or monitoring.* **Mass surveillance** *is the surveillance of groups of people, usually large groups. In general, the reason for investigation or monitoring is to identify individuals who belong to some particular*

class of interest to the surveillance organization.

Personal surveillance is an important weapon in the fight against such social evils as terrorism and organized crime. It is used to collect evidence in civil cases. It is also a means of learning sufficiently embarrassing facts about a person to assist in discrediting him or her in the eyes of some other person or group, or buying his or her silence or agreement. At its most secret, it can deny the subject natural justice, and at its most open, it can be tantamount to coercion or blackmail.

Personal- surveillance activities are undertaken by "private investigators" for corporate and personal clients. The majority of these activities, however, are undertaken by staff employed by government agencies, including police, national security, customs, and telecommunications officials.

Mass surveillance is difficult to discuss dispassionately because of the impact on our culture of the anti- utopian novels, particularly 1984 (45]. Its primitive forms include guards on raised walkways and observation turrets. More recently, closed-circuit television has offered a characteristic that significantly enhances the effectiveness of surveillance: The subjects, even if they know they are subject to monitoring, cannot know precisely when the observer is actually watching.

"Modern techniques have made possible a new intensity of governmental control, and this possibility has been exploited very fully in totalitarian states . . [E]mphasis upon the value of the individual is even more necessary now than at any former time" [60, p. 35]. Such a sentiment could be expected from the contemporary civil libertarian lobby. In fact, the words predate the use of computers even in information management, let alone personal data management, having been written in 1949 by Bertrand Russell. Ubiquitous two- way television a la 1984 has not arrived, even though it is readily deliverable. It is unnecessary because dataveillance is technically and economically superior.

RELEVANT I.T. TRENDS

Computers were originally developed for their high speed computational capabilities. They subsequently spawned or stimulated a wide variety of related technologies and have been married with telecommunications. Data can now be captured, stored, processed, and accessed readily and economically, even when the facilities and their users are physically dispersed. Figure 1 identifies some particularly pertinent aspects of current developments in IT.

Figure 1: Relevant Components of IT Development

- Magnetic data-storage capabilities have improved immensely between 1965 and 1985, and optical storage is expected to have a significant impact in at least some application areas
- A rich assortment of input and output technologies has been developed to support the capture and dissemination of data
- Textual and conventional 'structurable' data have been dealt with successfully for some years by DBMS technology. The management of image and voice data is improving, and integrated data management, and conversion between the various forms, are now being addressed
- Inroads have been made into natural language understanding, at least in respect of the more formal usages of language by humans
- The hitherto numerical bias of computing technology is being augmented by symbolic manipulative capabilities. Many kinds of complex deterministic problems can now be tackled, and progress is being made in modelling probabilistic, 'fuzzy' and stochastic processes
- Significant improvement in telecommunications continue, particularly in speed, cost, reliability, robustness, security and standardisation

Apparently distinct technologies have drawn together very quickly, as in electronic funds transfer systems (EFTS) and their nephew EFT/POS (point of sale). Change is being wrought less by computers themselves than by amalgams of many interacting and mutually supporting technologies. Optical- storage technology may portend a new surge in such compound high-technology ventures.

IT crystal- ball gazing is fraught with danger. Nevertheless, discernible trends include the integration with EFTS of air- travel systems and telephone charging; road- traffic monitoring, including vehicle identification, closely integrated with ownership and driver's- license records; computerization and integration of court records, criminal records, fingerprint records, and criminal-investigation systems; integration of structured and textual data to support criminal investigation and national- security applications; computerization and integration of birth, death, and marriage records; and homes wired for reasons of employment, security, entertainment, and consumerism.

As a consequence of the centralizing tendency of early IT, a "data imperative" arose, with government agencies and private companies alike collecting ever more data. Rule interpreted this as commitment to the "efficiency criterion," whereby privacy concerns should be recognized, but not at the cost of administrative efficiency. IT led to increasingly information- intensive practices and increasingly fine- grained decision making [55, 59].

With the repeal of Grosch's law during the 1970s, economies of scale no longer apply to processing power. Other factors that are

militating against the old centralist notions are the systems software overheads of large- scale centralized processing; risks associated with single- site activities; standardization of local and site networking standards; fast- growing capabilities of network workstations and servers; decreasing cost and increasing portability and robustness of dense storage, as in the so- called "smart card"; established techniques of distributed DBMS; and emerging techniques of distributed operating systems. The once-obvious tendency of computers to centralize information, and hence power, is quickly giving way to the looser concepts of networking and dispersion.

THE NATIONAL DATA CENTER ISSUE

In the mid- 1960s the U.S. government considered creating a national data center. The prime motivation was stated to be the need for more coherent data management to support economic and sociological research. Such a data collection, however, had clear potential for supporting administrative decision making. A few people recognized the vital role of data dispersion: "One of the most practical of our present safeguards of privacy is the fragmented nature of present information. It is scattered in little bits and pieces across the geography and years of our life. Retrieval is impractical and often impossible. A central data bank removes completely this safeguard" (Representative Frank Horton, 1966-67 hearings on a proposed national data center, quoted in [59, p. 56]). Concerns such as these resulted in the proposal not proceeding.

Centralized storage, however, is no longer a precondition of the dossier society that Horton feared. For dataveillance purposes a single centralized data bank is unnecessary, provided that three conditions are fulfilled:

1. a range of personal data systems must exist, each processing data for specific purposes;
2. some, preferably all, personal data systems must be connected via one or more telecommunications networks;
3. the data must be identified consistently.

A recent report of the Office of Technology Assessment (O.T.A.) of the U.S. Congress discussed the manner in which use of IT is quickly leading to a de facto national identification system [43, pp. 3, 68, 74]. There are also reports suggesting that both the NClC (National Crime Information Center) and NSA (National Security Agency) are providing foci for such a system (e.g., [54, pp. 185-186]).

Beyond assisting in the investigation of people's pasts. IT is also dramatically improving the monitoring of people's ongoing activities and present location. A person's most recent financial (or indeed any other kind of) transaction indicates where the person can currently be found. If that location is communicated to surveillance staff immediately, they can literally be on their way to the scene before the person leaves the checkout counter. A recent Australian report commenced with the sentence "EFTPOS is not a Greek island" [3]. The comment had a poignancy that its authors may not have appreciated. People go to Greek islands to escape from it all. EFT/POS constitutes a real- time locator service: You cannot escape from it at all. This was recognized at least as long ago as 1971, when it was suggested as an appropriate surveillance tool for the KGB (Armer, quoted in [59, p. 115]).

Physical and even communications surveillance are labor- intensive activities, which have so far proved difficult to automate. Dataveillance is essentially computer based, with the "watch and report" responsibility delegated to a reliable, ever- wakeful servant. It is increasingly cost- effective for organizations to place people under surveillance via their records and transactions, and traditional methods are being relegated to the role of complementary techniques. Furthermore, because dataveillance is cheaper than traditional methods, there is a tendency for more organizations to monitor more people: Both personal and mass surveillance are becoming routinized [35, 51].

THE CENTRAL ROLE OF IDENTIFICATION SCHEMES

Of the three requirements for a dispersed national data center identified earlier, the first two are already fulfilled. The third has not as yet been achieved, because of the difficulties of reliably identifying surveillance subjects, in associating stored data with individuals, and in associating new data with old.

The vast majority of personal data systems use schemes based on documentary evidence, possession of tokens, and personal knowledge. None of these can provide a satisfactory basis for a high- integrity system [9]. Many organizations that need to recognize identities in successive transactions assign their data subjects a more- or- less arbitrary unique identifying code. Some organizations prefer to (or have to) identify individuals by their names, usually supplemented by additional data such as date of birth. This approach involves a great deal of ambiguity, and name- matching routines have been developed to apply algorithms to such data in order to display synonyms "most- likely- first."

These various schemes may be of a reasonable level of integrity where data subjects have an interest in their accuracy, but are otherwise of, at best, only moderate integrity. Organized crime finds such low- integrity schemes as a social- security number a positive boon in its aims to legitimize false identities.

A high- integrity identification scheme is only possible if some physiological attribute is used that the person cannot alienate, and that the organization can capture, recognize and store with its records. During 1987 the New South Wales (N.S.W.) Police

that the organization can capture, recognize and store with its records. During 1987 the New South Wales (N.S.W.) Police Department implemented a fingerprint- record system based on Japanese technology. Although some U.S. state and local government law- enforcement agencies have installed such systems, the N.S.W. initiative is quite significant, since its bureau operates on behalf of all police departments throughout Australia, and it appears to have been the first national system to enter operation anywhere in the world. Apart from criminal records and limited applications in building security arrangements, fingerprint identification has not been socially acceptable. It can confidently be expected that there will be considerable efforts to make it so.

Historically, organizations have developed their identification schemes independently of one another, and large, multifunction organizations have run multiple schemes. However, organizations are increasingly using a single code for multiple purposes. For example, since at least the early 1970s, financial institutions have been moving toward "client- oriented" data management, whereby all of a client's data carry the same identifying code. The Australian Department of Social Security has committed itself to a common identification scheme for recipients of all classes of benefits by 1990.

Some identifiers were designed to be used for multiple purposes, whether by a single organization or by several. For example, in European countries it is normal for the same number to be used for the national superannuation fund as for taxation. Despite successive reports recommending the contrary (e.g., [48, 68]), the United States is continuing its trend toward using the (originally single- purpose) social security number as a de facto national identification code. For example, in 1985 a database called ESVARS was established, explicitly to enable any organization (federal, state, or private sector) to verify social- security numbers [43, p. 73].

General- purpose schemes, for use by all organizations for all purposes, have been attempted in a number of countries during wartime, including the United Kingdom and Australia, when the "inducement of rationing" has made them workable [55, p. 314]. The few countries that have considered such a scheme in peacetime have rejected the idea. The United States did so in the mid- 1970s [68]. The Australian government proposed such a scheme in 1985, although the proposal was subsequently amended to a multipurpose scheme for three main agencies, and finally withdrawn in the third quarter of 1987 when serious public concern arose about its implications [9]. As the scope of use of an identification scheme moves from single- use via multiple- use toward general- purpose use, the ease with which dataveillance can be undertaken increases significantly.

TECHNIQUES OF DATAVEILLANCE

This section discusses the techniques used in personal dataveillance, mass dataveillance, and facilitative mechanisms. Figure 2 provides a summary.

Figure 2: Techniques of Dataveillance

- **Personal Dataveillance**, of previously identified individuals
  - integration of data hitherto stored in various locations within a single organisation
  - screening or authentication of transactions against internal norms
  - front-end verification of transactions that appear to be exceptional, against data relevant to the matter at hand, and sought from other internal databases or from third parties
  - front-end audit of individuals who appear to be exceptional, against data related to *other* matters, and sought from other internal databases or from third parties
  - cross-system enforcement against individuals, where a third party reports that the individual has committed a transgression in his or her relationship with the third party
- **Mass Dataveillance**, of groups of people
  - screening or authentication of all transactions, whether or not they appear to be exceptional, against internal norms
  - front-end verification of all transactions, whether or not they appear to be exceptional, against data relevant to the matter at hand, and sought from other internal databases or from third parties
  - front-end audit of individuals, whether or not they appear to be exceptional, against data related to *other* matters, and sought from other internal databases or from third parties
  - single-factor file analysis of all data held or able to be acquired, whether or not they appear to be exceptional, variously involving transaction data compared against a norm, permanent data or other transaction data
  - profiling, or multi-factor file analysis of all data held or able to be acquired, whether or not they appear to be exceptional, variously involving singular profiling of data held at a point in time, or aggregative profiling of transaction trails over time
- **Facilitative Mechanisms**
  - computer or data matching is a process in which personal data records relating to many people are compared in order to identify cases of interest
  - data concentration is the combination of personal data through organisational merger or by the operation of data-interchange networks and hub systems

---

Personal Dataveillance Techniques

Organizations maintain records about individuals they are concerned with (their *data subjects).* In most cases data subjects are clients of the organization, because they have a known and fairly explicit relationship with them. This relationship may be direct (as with financial institutions and the Internal Revenue Service (lRS)) or indirect (as is sometimes the case with superannuation funds). With some record- keeping organizations, there may be no overt relationship (e.g., counterintelligence agencies, private investigators, and credit bureaus).

Subjects of personal dataveillance have attracted attention for some reason. The reason may be benign, for example, because they have applied for employment or a service. An investigation will usually have the intention of disqualifying the person from the employment or service they seek, but sometimes the organization may be considering whether the person may qualify for extra assistance, say, because of aboriginal ancestry. Another class of reasons for investigation is suspicion that the person has committed a crime or misdemeanor. A transaction may have taken place that appears inconsistent with existing records, or potentially incriminating information may have been received from outside the organization.

Dataveillance depends on data that identify people. Despite the increase in information intensity in recent years, there remain many economic relationships in which the parties do not necessarily identify themselves. These include barter and cash transactions ranging from hunting and fishing licenses, through gambling, and bus, train, and ferry tickets to quite large consumer items, including expensive cars and boats. Authorities throughout the world, concerned about the ease with which organized crime "washes" its illegally gained cash, have set, or are considering setting, maximum limits on unidentified cash transactions.

Given that identified records exist, a variety of dataveillance techniques are available. The most primitive technique, *record integration,* brings together all of the data an organization holds about each person. For many organizations this is not the trivial exercise it appears. Data may be dispersed in many ways, such as geographically across different offices and files, or under different codes or names (e.g., where the person has changed name; operates under multiple identities, including married and maiden names, and business and company names; operates joint accounts, sometimes with another party's name first; uses various combinations of given names; or has a name that is subject to spelling variants). In addition, during the early years of administrative and commercial applications of computing, it has been cost- effective and even necessary to store transaction data separately from permanent data, and transaction data for each period of time in distinct files. Financial institutions have undergone the transition to client- oriented data storage, the insurance industry is going through it, and airlines have commenced the changeover.

An approach adopted by most organizations is to monitor new transactions. Each transaction an organization receives (e.g., an application for employment, a loan, or a government benefit) is processed according to standard rules to determine whether the transaction is valid and acceptable. Additional rules may be applied, expressly designed to detect both inaccuracies and attempts to cheat the decision criteria. Exceptional cases are generally submitted to a more senior authority for more careful, nonroutine consideration.

Where the processing rules depend only on data already available to the organization, these practices are generally referred to as screening or authentication. Some commercial and administrative activity is impractical without such basic data processing, and in many cases the law requires it, as, for example, in the processing of applications for government benefits. Moreover, the stewardship responsibilities to which any organization is subject generally require that controls be built into transaction processing. subject to cost- effectiveness constraints.

Front- end verification of transactions represents a further development beyond screening [25; 43, pp. 67-83]. It involves the collection of data from other personal data systems in order to facilitate the processing of a transaction. The source of the data may be elsewhere in the same organization, for example, a driver's- licensing authority might consult its traffic- offenses database when renewing licenses. More commonly, however, front- end verification involves communication between two or more distinct organizations, either on an ad hoc basis or under a standing data- interchange arrangement.

Front- end verification is a personal- dataveillance technique when the transaction has been identified as exceptional, and the purpose of collecting the additional data is to establish whether there is any inconsistency between the various sources of data. An inconsistency may disqualify the transaction or be evidence of some wrongdoing such as providing misleading information. The data- interchange arrangements necessary to support front- end verification are not well documented in the literature. Rule provides one good reason: "No topic evoked less candour ... or gave rise to more vivid displeasure when I insisted on pursuing it" [55, p. 308].

Front- end verification tests transactions. A broader form of personal dataveillance is what might be termed *front- end audit.* This uses the occasion of the detection of an exceptional transaction as an opportunity to further investigate other matters relating to the individual. For example, when a driver is stopped for a traffic offense, it is becoming standard practice for the police officer to initiate on- line inquiries. These typically concern the vehicle (whether it is currently registered and whether it has been reported stolen), the vehicle's registered owner, and the driver (whether the driver is being sought for questioning or has an

reported stolen), the vehicle's registered owner, and the driver (whether the driver is being sought for questioning or has an outstanding arrest warrant). The first transaction generally arises because there are reasonable grounds for believing that an offense has been committed. The justification for the subsequent transactions is less clear.

Intersystem and interorganizational arrangements can be pursued a step further by means of *cross- system enforcement* . This technique makes an individual's relationship with one organization dependent on his or her performance in relation to another. For example, there have been proposals in some U.S states whereby renewal of a driver's license or entry to a turnpike would be precluded until the person has paid all outstanding parking fines. Steps have been taken in this direction (e.g., to preclude the sale of books of turnpike tickets), but to date their effectiveness appears to be doubtful.

In those cases the systems to be used for cross- enforcement are different, but to some extent related. There have been suggestions in New York City that even the issuing of a marriage license might be made dependent on payment of outstanding parking fines. In this case the link between the two systems is rather tenuous it is merely that the same organization has responsibility for both functions. Such a mechanism was included in the Australian government's proposed national identification scheme: The individual's rights under Medicare to free treatment or a refund of medical expenses would have been suspended until that person's obligation to have a national identity' card was fulfilled [9].

Mass Dataveillance Techniques

Personal dataveillance is concerned with identified individuals about whom some kind of concern or suspicion has arisen. On the other hand, mass dataveillance is concerned with groups of people and involves a generalized suspicion that some (as yet unidentified) members of the group may be of interest. Its purposes are to identify individuals who may be worth subjecting to personal surveillance, and to constrain the group's behavior.

Screening or authentication of transactions, discussed earlier, is arguably a form of mass surveillance to the extent that it is routinely or automatically applied to every transaction, whether or not it appears to be exceptional. Similarly, when data are routinely sought from other internal databases or third parties in order to undertake front- end verification of all transactions, mass dataveillance is being undertaken. This is a recent development in which lT's role has been criticized: "In the past, such verification was done manually on a random basis or when the accuracy of information provided was suspect. Today, . . . computerized databases and on- line networking make it possible to carry out such verification routinely" [43, p. 67]. Similarly, front end audit is a mass- surveillance technique if the investigation of multiple aspects of a person's performance arises without any explicit cause, rather than as a result of some exceptional transaction.

> *Personal dataveillance is concerned with identified individuals about whom some kind of concern or suspicion has arisen.*

The application of such techniques to existing records, rather than to new transactions, is referred to here as *file analysis: "*The files are most useful where they enable the system quickly and unerringly to single out the minority of their clients who warrant some measure of social control" (Rule, quoted in [67]). File analysis can be effective in searching out what Marx and Reichman refer to as "low- visibility offenses" [35]. In a recent instance in the United Kingdom, government investigators applied file-analysis techniques to detect and prosecute multiple applications for shares in "privatized" government enterprises such as Telecom and British Petroleum.

Screening, front- end verification, front- end audit, and file analysis may all be undertaken with varying degrees of sophistication. Transaction data may be compared against a formal standard or other norm, for example, highlighting those tax returns that include deductions above a certain value or show more than, say, eight dependents. The norms against which the data are compared may be either legal or other a priori norms that have been set down in advance by some authority, possibly for good reasons, possibly quite arbitrarily. Alternatively, they may be a posteriori norms that were inferred from analysis of the collection of records.

Alternatively, transaction data may be compared against permanent data, for example, highlighting tax returns where the spouse's name does not match that on file. Or transaction data may be compared against other transaction data, for example, highlighting people whose successive tax returns `show varying numbers of dependents.

The previous examples are each based on a single factor. Judgments of any complexity must be based on multiple factors, rather than just one. *Profiling,* as it is commonly known, may be done on the basis of either a priori arbitrary or pragmatic norms, or on a posteriori norms based on empirical evidence. Rule noted in 1974 that the IRS used an a posteriori technique for predicting the "audit potential" of different returns. It did this by inferring unknown characteristics from known characteristics by applying discriminant analysis to a small random sample of returns [55, p. 282]. Marx and Reichman's description of this technique is "correlating a number of distinct data items in order to assess how close a person comes to a predetermined characterization or model of infraction" [35, p. 429]. These authors further distinguish "singular profiling" from "aggregative profiling," which involves analyzing transaction trails over a period of time.

Sophisticated profiling techniques are claimed to hold great promise because they can detect hidden cases amid large populations

Sophisticated profiling techniques are claimed to hold great promise because they can detect hidden cases amid large populations. Benefits could be readily foreseen from profiles of young people with proclivities toward particular artistic and sporting skills; propensity' for diseases, disorders, delinquency, or drug addiction; or suicidal or homicidal tendencies. A recent O.T.A. report noted that most U.S. federal agencies have applied the technique to develop a wide variety of profiles including drug dealers, taxpayers who underreport their income, likely violent offenders, arsonists, rapists, child molesters, and sexually exploited children [43, pp. 87-95].

Facilitative Mechanisms

Mass- dataveillance techniques may be successfully applied within a single personal- data system, but their power can be enhanced if they are applied to data from several. These systems might all be operated by the organization concerned or by a number of distinct organizations. In such cases a preliminary step may he undertaken:

> *Computer matching is the expropriation of data maintained by two or more personal-data systems, in order to merge previously separate data about large numbers of individuals.*

Matching has become technically and economically feasible only during the last decade, as a result of developments in IT. The first large program reported was Project Match, undertaken by the U.S. Department of Health, Education and Welfare (HEW), now known as Health and Human Services (HHS). By 1982 it was estimated that about 500 programs were carried out routinely in U.S. state and federal agencies [69], and O.T.A. estimated a tripling in use between 1980 and 1984 [43, p. 37]. Moreover, a succession of federal laws, culminating in the 1984 Budget Deficit Reduction Act, imposed matching on state administrations as a condition of receiving federal social- welfare funding. (For references descriptive of and supportive of matching, see [25], [39]-[41], [65]. and [66]. Cautionary and critical comments are to be found in, [22], [23], [26], [35], [43], [50], and [61]).

Matching makes more data available about each person and also enables comparison between apparently similar data items as they are known to different organizations. Rather than relating to a single specified person for a specific reason, matching achieves indiscriminate data cross-referencing about a large number people for no better reason than a generalized suspicion: "Computer matches are inherently mass or class investigations, as they are conducted on a category of people rather than on specific individuals .. . in practice, welfare recipients and Federal employees are most often the targets" [43, p. 40].

Matching may be based on some common identifier that occurs in both files, in which case the error rate (measured by the proportion of undetected matches and spurious matches) will tend to be fairly low. There are few opportunities for such matching, however, and instead it is usually necessary to correlate several items of information. Intuitively, name, birth date, and sex seem appropriate, but it appears that greater success has been achieved by using some component of address as a primary matching criterion.

Individuals may be judged to be interesting because of:

- the existence of a match where none should exist,
- the failure to find a match where one was expected,
- inequality between apparently common data items (e.g., different numbers of dependents), or
- logical inconsistency among the data on the two files (e.g., the drawing of social- welfare a during a period of employment).

Curiously, the current U.S. government matching guidelines [40] define matching only in terms of the first of these criteria.

An additional facilitative mechanism for both personal and mass dataveillance is referred to here as *data concentration.* The conventional approach is to merge existing organizations in search of economies of scale in administration. If the capabilities of large- scale data processing equipment were to continue to increase, the merger of social- welfare and internal- revenue agencies could be anticipated enabling welfare to be administered as "reverse taxation."

Dataveillance is, by its very nature, intrusive and threatening.

Organizational merger is an old- fashioned "centralized" solution. The modern, dispersed approach to data concentration is to establish systems that can function as the hub of a data- interchange network. For example. the U.S. government has developed new systems to facilitate routine front- end verification. These initiatives, in the name of waste reduction, involve both federal government sources (including IRS and criminal records) and private- sector credit bureaus, and their use not only extends across many federal government agencies, but is also imposed on state welfare administration agencies [14; 42, pp. 68-74]. The Australian government's proposal for a national identification scheme involved just such a coordinating database [9].

DATAVEILLANCE'S BENEFITS AND DANGERS

In this section the advantages dataveillance techniques offer are briefly discussed. Greater space is then devoted to the threats

dataveillance represents. Figure 3 summarizes these dangers.

Figure 3: Real and Potential Dangers of Dataveillance

- **Dangers of Personal Dataveillance**
  - wrong identification
  - low quality data
  - acontextual use of data
  - low quality decisions
  - lack of subject knowledge of data flows
  - lack of subject consent to, data flows
  - blacklisting
  - denial of redemption
- **Dangers of Mass Dataveillance**
  - **To the Individual**
    - arbitrariness
    - acontextual data merger
    - complexity and incomprehensibility of data
    - witch hunts
    - ex-ante discrimination and guilt prediction
    - selective advertising
    - inversion of the onus of proof
    - covert operations
    - unknown accusations and accusers
    - denial of due process
  - **To Society**
    - prevailing climate of suspicion
    - adversarial relationships
    - focus of law enforcement on easily detectable and provable offences
    - inequitable application of the law
    - decreased respect for the law and law enforcers
    - reduction in the meaningfulness of individual actions
    - reduction in self-reliance and self-determination
    - stultification of originality
    - increased tendency to opt out of the official level of society
    - weakening of society's moral fibre and cohesion
    - destabilisation of the strategic balance of power
    - repressive potential for a totalitarian government

Benefits

Significant benefits can result from dataveillance. The physical security of people and property may be protected, and financial benefits may accrue from the detection and prevention of various forms of error, abuse, and fraud. Benefits can be foreseen both in government activity (e.g., tax and social welfare) and in the private sector (e.g., finance and insurance). (For the limited literature on the benefits of matching, see [16], [25], [43, pp. 50-52], [65], and [66]. Literature on the benefits of other dataveillance techniques is very difficult to find).

Some proponents claim that the deterrent effect of public knowledge that such techniques are applied is significant, perhaps even more significant than direct gains from their actual use. There may be symbolic or moral value in dataveillance, irrespective of its technical effectiveness.

Few people would contest the morality of an organization applying the more basic techniques, for example, record integration and screening. Some would go so far as to regard organizations that did not apply modern IT in such ways as failing to fulfil their responsibilities to taxpayers and shareholders. Nevertheless, dataveillance is, by its very nature, intrusive and threatening. It therefore seems reasonable that organizations should have to justify its use, rather than merely assuming its appropriateness.

Dangers of Personal Dataveillance

Because so few contemporary identification schemes use a physiological identifier, they are, at best, of moderate integrity. Rather than individuals themselves, what is monitored is data that purport to relate to them. As a result there is a significant likelihood of wrong identification.

The vast majority of data systems operators are quite casual about the quality of most of their data; for example, the O.T.A. reported that few federal government agencies have conducted audits of data quality [43, p. 26]. For many organizations it is cost- effective to ensure high levels of accuracy only of particular items (such as invoice amounts), with broad internal controls designed to ensure a reasonable chance of detecting errors in less vital data. Some errors are intentional on the part of the data subject, but many are accidental, and some are a result of design deficiencies such as inadequate coding schemes. Similar problems arise with other elements of data quality such as the timeliness and completeness of data. Even in systems where a high level of integrity is important, empirical studies have raised serious doubts [30; 43, pp. 52-53]. Data quality is generally not high, and while externally imposed controls remain very limited, it seems likely that the low standards will persist.

People and matters relating to them are complicated, and organizations generally have difficulty dealing with atypical, idiosyncratic cases or extenuating circumstances [35, p. 436]. A full understanding of the circumstances generally requires additional data that would have seemed too trivial and/or expensive to collect, but also depends on common sense, and abstract ideas like received wisdom, public opinion, and morality [54]. When the data are used in their original context, data quality may be sufficient to support effective and fair decision making, but when data are used outside their original context, the probability of misinterpreting them increases greatly. This is the reason why information privacy principles place such importance on relating data to the purpose for which they are collected or used [44], and why sociologists express concern about the "acontextual" nature of many administrative decision processes [35].

Much front- end verification is undertaken without the subject's knowledge. Even where an organization publicizes that it seeks data from third parties, the implications of the notice are often unclear to the data subject. International conventions stipulate that data should not be used for purposes other than the original purpose of collection, except with the authority of law or the consent of the data subject (e.g., [44]). Where consent is sought, the wording is often such that the person has no appreciation of the import of the consent that is being given, or the bargaining position is so disproportionately weighted in favor of the organization that the data subject has no real option but to comply. Effective subject knowledge and consent mechanisms are necessary, both as a means of improving data quality, and to avoid unnecessary distrust between individuals and organizations.

Front- end audit and cross- system enforcement give rise to additional concerns. Their moral justification is not obvious, and they create the danger of individuals being effectively blacklisted across a variety of organizations. Credit- bureau operations are extending in some countries into insurance, employment. and tenancy. Acute unfairness can arise, for example. when organizations blacklist a person over a matter that is still in dispute. It is particularly problematic where the person is unaware that the (possibly erroneous, incomplete. or out- of- date) data have been disseminated. Finally, even where individuals have brought the problems upon themselves, blacklisting tends to deny them the ability to redeem themselves for past misdemeanors.

Dangers of Mass Dataveillance to the Individual

Mass dataveillance embodies far greater threats. In respect of each individual, mass surveillance is clearly an arbitrary action, because no prior suspicion existed. The analogy may be drawn with the powers of police officers to interfere with the individual's quiet enjoyment. If a police officer has grounds for suspecting that a person has committed, or even is likely to commit, an offense, then the police officer generally has the power to intercept and perhaps detain that person. Otherwise, with rare and, in a democratic state, well justified exceptions, such as national security emergencies and, in many jurisdictions, random breath testing, even a police officer does not have the power to arbitrarily interfere with a person.

With mass dataveillance, the fundamental problems of wrong identification, unclear, inconsistent, and context- dependent meaning of data, and low data quality are more intense than with personal dataveillance. Data arising from computer matching are especially problematic. Where there is no common identifier, the proportion of spurious matches (type (1) errors) and undetected matches (type (2) errors) can be very high. The causes include low quality of the data upon which computer matching depends (variants, misspellings, and other inaccuracies, and incompleteness), inappropriate matching criteria, widely different (or subtly but significantly different) meanings of apparently equivalent data items, or records with differing dates of applicability. Marx and Reichman report a New York State program in which half of the matches were spurious due to timing problems alone [35, p. 435]. In addition, the meaning of the record as a whole must be properly understood. Although it might seem improper for a person to be both in employment and in receipt of a social- welfare benefit, many pensions and allowances are, in law, either independent of, or only partially dependent on, income from other sources.

Data on the error rates of matching programs are difficult to find: They are mostly conducted away from the glare of public, or indeed any other kind of, supervision. In an incident in Australia in 1986, the federal agency responsible for the Medicare scheme calmly, and without apparent legal authority, expropriated and merged data from several federal government agencies, relating to all inhabitants of the small island state of Tasmania. The agency reported the 70 percent hit rate across the databases as a good result, confirming its belief that a national identification scheme could be based on such a procedure. They ignored the implication that across the national population the records of nearly five million persons would remain unmatched, and failed to apply any tests to establish what proportion of the 70 percent were spurious matches and what proportion of the 30 percent nonmatches were failures of the algorithm used. Australians embrace a popular mythology that everyone in Tasmania is related to everyone else. For this reason alone, the agency might have been expected to recognize the need for such testing.

to everyone else. For this reason alone, the agency might have been expected to recognize the need for such testing.

The complexities of each system (particularly a country's major data systems such as taxation and social welfare) are such that few specialists are able to comprehend any one of them fully. It is arguably beyond the bounds of human capability to appreciate the incompatibilities between data from different systems and to deal with the merged data with appropriate care. Computer matching, therefore, should never be undertaken without the greatest caution and scepticism.

Profiling makes a judgment "about a particular individual based on the past behavior of other individuals who appear statistically similar" [43, p. 88]. Statistical techniques such as multivariate correlation and discriminant analysis have limited domains of applicability that are often poorly understood or ignored. Even if the statistical procedures are properly applied, a profile needs to be justified by systemic reasoning. In the hands of the inadequately trained, insufficiently professional, or excessively enthusiastic or pressured, profiling has all the hallmarks of a modern witch- hunting tool.

Profiling is not restricted to retrospective investigation. It purports to offer the possibility of detecting undesirable classes of people before they commit an offense. O.T.A. documents a "predelinquency" profile developed for the U.S. Law Enforcement Assistance Administration [43, p. 90]. Even if the technique is perceived to be successful, its use seems to run counter to some fundamental tenets of contemporary society. It is unclear on what moral and, indeed, legal grounds profiling may be used to reach administrative determinations about individuals or discriminate between individuals. Such vague constraints may not be sufficient to stultify an attractive growth industry. With computer displays and printouts lending their (largely spurious) authority to such accusations, how will the tolerance needed in complex social environments be maintained?

Not only in government, but also in the private sector, dangers arise from both the effectiveness and ineffectiveness of profiling. The combination of consumer profiles with cheap desktop publishing is dramatically altering the cost- effectiveness of customized "mail shots." Applied to cable television, the technique will enable the operator to selectively transmit "commercials" to those subscribers who seem most likely to be susceptible to the client's product (or perhaps just the advertisement). Whereas Vance Packard could only prophesy the development of such technology [47]. the components can now be identified. and the economics described.

Conventional justice is expensive and slow. Some procedures are now being structured, particularly in such areas as taxation, such that a government agency makes a determination, and individuals who disagree must contest the decision [61]. This inversion of the onus of proof exacerbates the problems of misinterpretation resulting from data merger, and uncertainty arising from correlative profiling. It is further compounded by the imbalance of power between organization and individual. Marx and Reichman provide an example in which individuals were confronted by a complex of difficulties: A remote examination authority statistically analyzed answer sheets, and threatened students who had sat in the same room and given similar (incorrect) answers with cancellation of their results unless they provided additional information to prove they did not cheat [35, p. 432].

Some dataveillance is undertaken with dubious legal authority or in the absence of either authority or prohibition. To avoid being subjected to public abuse and perhaps being denied the right to undertake the activity, it is natural for organizations to prefer to undertake some operations covertly. There are also cases where the benefits of surveillance may be lost if it is not undertaken surreptitiously (e.g., because of the likelihood of the person temporarily suspending, rather than stopping, undesirable activities; or of "skips" on consumer credit transactions).

To protect the mechanism or the source, an individual may not be told that dataveillance has been undertaken, the source of the accusation, the information on which the accusation is based or even what the accusation is. Such situations are repugnant to the concept of due process long embodied in British law and in legal systems derived from it. Dataveillance tends to compromise the individual's capacity to defend him or herself or to prosecute his or her innocence. In its most extreme form, one Kafka could not anticipate, the accuser could be a poorly understood computer program or a profile embodied in one.

Social Dangers of Mass Dataveillance

At the social level, additional problems arise. With personal dataveillance, investigation and monitoring normally take place after reasonable grounds for suspicion have arisen. Mass surveillance dispenses with that constraint because the investigation is routinely performed and the suspicion arises from it. The organization therefore commences with a presumption of guilt on the part of at least some of the data subjects, although at the beginning of the exercise it is unknown which ones. The result is a prevailing climate of suspicion.

The organizational functionary who communicates with the data subject often only partially understands the rationale underlying the decision, prefers not to admit that lack of understanding, and is often more concerned with case resolution than with public relations. Hence, there is an increased tendency for organizations and data subjects to develop adversarial relationships. Moreover, since organizations generally have the information, the size and the longevity, the bargaining positions are usually unequal.

Some of the "atypical. idiosyncratic, and extenuating cases" that are uncovered by mass dataveillance are precisely the deviants who are being sought. But others are just genuinely different, and such people tend to have difficulties convincing bureaucrats

who are being sought. But others are just genuinely different, and such people tend to have difficulties convincing bureaucrats that their idiosyncrasies should be tolerated. Dataveillance encourages investigators to focus on minor offenses that can be dealt with efficiently, rather than more important crimes that are more difficult to solve. Law enforcers risk gaining a reputation for placing higher priority on pursuing amateur and occasional violators (particularly those whose records are readily accessible, like government employees and welfare recipients), rather than systematic, repetitive, and skilled professional criminals. The less equitably the law is perceived to be enforced, the greater the threat to the rule of law.

An administrative apparatus that has data available to it from a wide variety of sources tends to make decisions on the person's behalf. Hence, a further, more abstract, yet scarcely less real impact of dataveillance is reduction in the meaningfulness of individual actions, and hence in self- reliance and self- responsibility. Although this may be efficient and even fair, it involves a change in mankind's image of itself, and risks sullen acceptance by the masses and stultification of the independent spirit needed to meet the challenges of the future.

Some people already opt out of official society, preferring bureaucratic anonymity even at the costs of foregoing monetary and other benefits, and, consequently, attracting harassment by officialdom. There may already be a tendency toward two- tiered societies, in which the official documentary level of government facts and statistics bears only an approximate relationship to the real world of economic and social activity. If uncontrolled dataveillance were to cause the citizens of advanced Western nations to lose confidence in the fairness with which their societies are governed, it would be likely to exacerbate that trend.

An increase in the proportion of economic activity outside mainstream society would prompt, and be used to justify, a further increase in the use of mass surveillance. Assuming that world politics continues to be polarized into an East-West confrontation, it would be very easy to justify tighter social controls since any sign of serious weakening ,in the moral fiber and integrity of the West would be destabilizing. Since "mastery of both mass communications and mass surveillance is necessary for an elite to maintain control" [57, p. 176], IT will be a major weapon whereby ruling groups seek to exercise control over the population.

Finally, it is necessary to mention (but not over dramatize) the risk of dataveillance tools supporting repressive actions by a future invader, or by the "dirty- tricks department" of some democratically elected government gone, as Hitler's did, somewhat off the rails: "Orwell foresaw--and made unforgettable-- a world in which ruthless political interests mobilized intrusive technologies for totalitarian ends. What he did not consider was the possibility that the development of the intrusive technologies would occur *on its own* , *without the spur of totalitarian intent.* This, in fact, is what is now happening" [57, p. 179).

In general, mass dataveillance tends to subvert individualism and the meaningfulness of human decisions and actions, and asserts the primacy of the state.

SAFEGUARDS
Intrinsic Controls over Dataveillance

Some natural controls exist that tend to limit the amount of dataveillance undertaken. The most apparent of these is its expense. There have been claims of dramatic success for matching schemes, but these have generally been made by the agencies that conducted them, and independent audits are hard to come by. The U.S. government's original (1979) guidelines on matching required that cost/benefit analyses be undertaken prior to the program being commenced [39]. However, there are many difficulties in undertaking a cost/benefit analysis of such a program. Many benefits are vague and unquantifiable, and many expenses are hidden or already "sunk." As a result, the requirement was rescinded in 1982 and has not been reimposed [40]. Moreover, there is seldom any other legal or even professional requirement that a cost/benefit analysis be performed [61, p. 540]. In 1986 O.T.A. concluded that few U.S. government programs are subjected to prior cost/ benefit assessment [43, pp. 50-52].

Although reliable audits are difficult to find, anecdotal evidence throws doubt on the efficacy of matching. In the original Project Match, HEW ran its welfare files against its own payroll files. The 33,000 raw hits that were revealed required a year's investigation before they could be narrowed to 638 cases, but only 55 of these were ever prosecuted. Of a sample of 15 cases investigated by the National Council for Civil Liberties after HEW released the names of the people involved, 5 were dismissed, 4 pleaded guilty to misdemeanors (theft under $50), and only 6 were convicted of felonies.

No prison sentences resulted, and the fines totaled under $2,000 [14, 49]. A 1983 match between Massachusetts welfare and bank files found 6,500 hits in five million records, resulting in 420 terminations of benefits, but also much confusion and recrimination [50]. Recent U.S. government reports have also raised doubts about the economic worth of many matching programs. A more positive report on several local government systems is to be found in [16].

There is very little evidence concerning the economics of other dataveillance techniques. Effective cost/ benefit assessment, however, appears to be very rare (e.g., [43, pp. 80-81)). Unless credible cost/benefit analyses are undertaken, at least retrospectively, and preferably in advance, the potential economic safeguard against excessive use of dataveillance cannot be realized.

Economic controls, even if they were effective, may not be sufficient to protect individual freedoms. In the early years of

Economic controls, even if they were effective, may not be sufficient to protect individual freedoms. In the early years of personal- data systems, the dominant school of thought, associated with Westin, was that business and government economics would ensure that lT did not result in excessive privacy invasion [74-76]. This view has been seriously undermined by Rule's work, which has demonstrated that, rather than supporting individual freedoms, administrative efficiency conflicts with it. Organizations have perceived their interests to dictate the collection, maintenance, and dissemination of ever more data, ever more finely grained. This is in direct contradiction to the interests of individuals in protecting personal data [55- 59). Meanwhile, the onward march of IT continues to decrease the costs of dataveillance.

Another natural control is that surveillance activities can incur the active displeasure of the data subject or the general public. Given the imbalance of power between organizations and individuals, it is unrealistic to expect this factor to have any relevance outside occasional matters that attract media attention. Another, probably more significant control is that an organization's activities may incur the displeasure of some other organization, perhaps a watchdog agency, consumer group, or competitor.

In any case, these natural controls cannot be effective where the surveillance activities are undertaken in a covert manner. Intelligence agencies in particular are subject to few and generally ineffective controls. Also, many controls, such as the power to authorize telephone interception, may not be subject to superordinate control. Intrinsic controls over dataveillance are insufficient to ensure that the desirable balance is found.

Extrinsic Controls over Dataveillance

The establishment of extrinsic controls over dataveillance cannot even be embarked upon until comprehensive information privacy laws are in place. Proper protection of privacy- invasive data handling was stillborn in the United States in the early 1970s by the limited official response associated with Westin [74-76], the Privacy Act of 1974, and the PPSC report [48]. Westin found no problems with extensive surveillance systems as such, only with the procedures involved, and the PPSC's aim was to make surveillance as publicly acceptable as possible, consistent with its expansion and efficiency [59, pp. 75, 110].

The U.S. Privacy Act was very easily subverted. Publication of uses in the Federal Register has proved to be an exercise in bureaucracy rather than control. The "routine use" loophole in the act was used to legitimize virtually any use within each agency (by declaring the efficient operation of the agency to be a routine use) and then virtually any dissemination to any other federal agency (by declaring as a routine use the efficient operation of the federal government) (see [35, p. 449; 43. pp. 16-21; 48]).

Rule's thesis [55-59]--that privacy legislation arose out of a concern to ensure that the efficiency of business and government was not hindered--has been confirmed by developments in international organizations and on both sides of the Atlantic. The OECD's 1980 Guidelines for the Protection of Privacy were quite explicitly motivated by the economic need for freedom of transborder data flows [44]. In the United States, the President's Council for Integrity and Efficiency (PCIE) and Office of Management and Budget (0MB) have worked not to limit matching, but to legitimize it [25]. In the United Kingdom, the Data Protection Act of 1984 was enacted explicitly to ensure that U.K. companies were not disadvantaged with respect to their European competitors.

There have been almost no personal- data systems, or even uses of systems, that have been banned outright. Shattuck [61, p. 540] reported that, during the first five years, the OMB's cavalier interpretation of the Privacy Act had resulted in not a single matching program being disapproved. Few sets of Information Privacy Principles appear to even contemplate such an extreme action as disallowing some applications of IT because of their excessively privacy- invasive nature. Exceptions include those of the New South Wales Privacy Committee [38], which are not legally enforceable, and, with qualifications, Sweden. This contrasts starkly with the conclusions of observers: "At some point . . . the repressive *potential* of even the most humane systems must make them unacceptable" [59, p. 120]; and "We need to recognize that the potential for harm from certain surveillance systems may be so great that the risks outweigh their benefits" [33, p. 48].

Some countries, such as Australia, have no information privacy legislation, and only incidental protections exist, such as breach of confidence, telephonic interception, trespass, and official secrecy [18]. In jurisdictions where information privacy safeguards do exist, they are piecemeal, restricted in scope, and difficult to enforce. In particular, many countries restrict the protections to government data or computer- based systems, or make no provision for such conventional safeguards as detailed codes of practice, oversight by an adequately resourced and legally competent authority. or the right to sue for damages.

Moreover, technological developments have rendered some information privacy protections ineffective. For example, the O.T.A. concluded that "the Privacy Act offers little protection to individuals who are subjects of computer matching" [43, p. 38] (see also [17] and [63]).

Avenues of Change

Only once the principles of fair information practices have been engrained into our institutions and our ways of thought will it be possible to address the more complex, subtle, and pervasive threats inherent in contemporary IT.

In some countries the courts have absolved themselves of responsibility to change the law for policy reasons, unequivocally

In some countries the courts have absolved themselves of responsibility to change the law for policy reasons, unequivocally asserting not just Parliament's primacy in, but its exclusive responsibility for, law reform. In the United States, although the Bill of Rights does not mention a right to privacy, the courts have progressively established such a right based on elements of several of the amendments. The likely present view of the U.S Supreme Court, however, might be indicated by this quotation: "I think it quite likely that self- discipline on the part of the executive branch will provide an answer to virtually all of the legitimate complaints against excesses of information- gathering" (Rehnquist, 1971, then a spokesperson for the Justice Department, now Chief Justice, quoted in [59, p. 147]). Moreover, courts throughout the world have difficulty with cases involving recent developments in technology [10, 72]. Accordingly, they prefer to await statutory guidance from parliaments, with their generally better financed and less- fettered access to technological know- how.

However, parliaments also tend toward inaction on difficult technological matters, particularly when they are proclaimed to be the salvation of the domestic economy or are tangled up with moral issues, such as "dole cheating" and "welfare fraud." Consumer protection laws in many countries still have yet to be adapted to cater for the now well- developed EFTS. Although the early literature on EFTS omits mention of its social impact, testimony was given before U.S. Senate subcommittees at least as early as 1975 on the repressive potentials of computerised payment systems [59, p. 115] (see also [24] and [55]). The call for protection was still necessary in 1984 in the United States [77] and in 1986 in Australia [3]. Parliaments in some countries such as Australia look less like sober lawmaking institutions than gladiatorial arenas. There are serious difficulties in convincing such legislatures to constrain the development of new "wonder technologies."

The conclusion is inescapable that the populations of at least some of the advanced Western nations are severely threatened by unbridled, IT- driven dataveillance.

POLICY PROPOSALS
New and Improved Safeguards

Since its brief period in the sun in the early 1970s, privacy has become unfashionable among lawmakers, and the momentum that the fair information practices/data protection/information privacy movement once had, has been lost. The PPSC concluded that "the real danger is the gradual erosion of individual liberties through the automation, integration and interconnection of many small, separate record- keeping systems, each of which alone may seem innocuous. even benevolent, and wholly justifiable" [48, p. 533). Its recommendations were ignored and are now in serious need of resuscitation, not just in the United States, but also in other countries whose information privacy protection regime has not kept pace with developments in IT.

In some countries an effective foundation for dealing with information privacy problems was established during the 1970s. In others, such as the United States, the first attempt failed to establish an adequate basis. Still others have not taken the first step. That necessary foundation can be roughly equated with the OECD's 1980 guidelines [44].

Additional steps must now be taken. It is clear today that the dictates of administrative efficiency are at odds with individual freedoms, and the power of dataveillance techniques is far greater than a decade and more ago. It is now essential that governments consider each dataveillance technique and decide whether it should be permitted under any circumstances at all; if so, what those circumstances are and how each proposal should be assessed in order to judge its compliance with those criteria; what code of practice should apply to its use; and what control mechanisms will ensure that each of these safeguards operates effectively and efficiently.

Further, it must be recognized that IT continues to develop, and mechanisms are needed to ensure that legislators in particular, and the public in general. are kept up- to- date with the salient features of new applications.

The Responsibilities of I.T. Professionals

It would be inappropriate for the purveyors of any technology to be responsible for decisions regarding its application. The technologist has an unavoidable interest in the outcome, and cannot appreciate and take into consideration the interests of the many different social groups who may consider themselves to be affected.

However, this necessary neutrality must not be interpreted as an excuse for inaction. IT professionals and academics alike have a moral responsibility to appreciate the power of the technology in which they play a part. Academe should commit some amount of research effort to the testing of the contentions in this paper as well as originate and evaluate proposals for technical safeguards. Both groups must publicize the nature and implications of their work, both for classes of affected individuals and for society as a whole. This applies as much to the negative consequences as it does to the potential benefits.

Finally, where there are acknowledged shortfalls in the regulatory environment in which IT is being applied, the IT practitioner has a responsibility to lobby for effective and efficient safeguards. This article has argued that existing safeguards are entirely inadequate. This implies a responsibility to approach lawmakers about the urgent need for developments in information privacy law. Although the actions of individual practitioners can be significant, coordinated policy efforts by professional bodies, such as the British and Australian Computer Societies, and by common- interest groups, such as ACM and the IEEE Computer

as the British and Australian Computer Societies, and by common- interest groups, such as ACM and the IEEE Computer Society, are likely to have greater effect.

I.T. as an Antidote to Information Concentration

If society is to control its fate, it must recognize a new Law of Requisite Variety in Information Systems (see Figure 4). Dispersion of authority and power, and, hence, of information, has long been regarded as vital to the survival of individualism and democracy. This law goes further, by recognizing that dispersion of data is also economic. Contrary to conventional wisdom, it is not administratively wasteful to treat the organs of executive government as distinct agencies, but rather administratively sensible.

Figure 4: The Law of Requisite Variety in Information Systems

- Society demands many different services, and many different organisations exist to provide them. Each of these organisations designs its information systems to support the functions it performs.
- For each system, the data definitions, the level of integrity of identification, and the degree of data quality are chosen to ensure cost-effectiveness. The definitions, identification mechanism and data quality features of each system are therefore qualitatively different from those of every other system.
- Hence a single information system cannot economically serve the interests of all organisations. Clusters of organisations may be supported by a single information system, but only at the risk of compromising the effectiveness of each of them. Economies of scale will only be achieved when the functions and priorities of the organisations are closely aligned.

Society may be better served by an alternative to centralization and its concomitant notions of rigidity and risklessness. If looseness, diversity, tolerance, initiative, enterprise, experimentation, and risk management were adopted as the bases for social and economic organization, then society could develop the adaptiveness needed to cope with technology- induced change. In the words of one philosopher, "The problem is to combine that degree of security which is essential to the species, with forms of adventure and danger and contest which are compatible with the civilised way of life" [60, p. 21).

Some elements will be critical to a human- oriented IT. For example, the alternative approach to identification proposed by Chaum [8] proposes that the capabilities of "smart cards" be used not only for the benefit of organizations, but also of individuals. Each organization would know each individual by a different "digital pseudonym," which would be the joint property of both parties. Each individual could deny organizations the ability to link their data about him or her without consent. Both parties would have their interests protected. By such approaches, contemporary. decentralizing IT can support the evolution of human- oriented society, rather than hasten the demise of the age of individualism.

CONCLUSION

Dataveillance applications of IT have serious implications for individualism and society. The limited improvements in information practices that were achieved during the last decade have been outpaced by technological developments. Yet until and unless comprehensive information privacy protection is in place, effective controls over the new and emerging techniques of dataveillance will not be possible.

This article does not argue that personal and mass dataveillance are intrinsically evil and should be proscribed. However, their serious implications must be traded off against their benefits in each and every instance. Moreover, those benefits must not be assumed, but carefully assessed. We must appreciate the implications of the new technological capabilities, and create safeguards such that some applications are proscribed and the remainder controlled. We need to harness the new, decentralizing potential of IT as a means of achieving a looser, more tolerant, diverse, robust, and adaptive society.

Acknowledgments

This article arises from collaborative research undertaken with Graham Greenleaf, of the Faculty of Law at the University of New South Wales. Assistance is also gratefully acknowledged from research assistants Louise Macauley and Chris Keogh, and Jim Nolan, executive member of the New South Wales Privacy Committee. The constructive criticism of referees and the area editor was also very helpful.

REFERENCES

1. Ackroyd. C.. Margotis. K., Resenhead. 1 . and Shallice, T.*The* Technology of Political Control. Penguin Books, New York, 1977

2. Askin. F. Surveillance: The social science perspective. Columbia Hum. Rights Law Rev. 4, 1(Winter 1972). (see also the remainder of the issue)

3. Australian Science and Technology Council. Towards a Cashless Society. ASTEC. Canberra Australia. May1986

4. Bramford, I. The Puzzle Palace. Penguin Books, New York, 1983.

5. Brenion. M. *The Privacy invaders* . Coward- McCann, 1964.

6. Bornham, 0. *The Rise of the Computer State*. Random House/Weidenfeld and Nicolson, 1983.

7. Campbell, D., and Connor, S. *On the Record* . Michael Joseph, 1986.

8. Chaum, D. *Security without identification* Transaction systems to make Big Brother obsolete. Commun ACM 28. 10 (Oct. 1985). 1030-1044.

9. Clarke. R.A. Just another piece of plastic for your wallet: The Australia card scheme. Prometheus 5, 1 (June 1987). 29-45.

10. Clarke, R.A. Judicial understanding of information technology. Comput. J 31. 1 (Feb. 1988).

11. Cowen, Z. The Private Man. Australian Broadcasting Commission. 1969.

12. Crispin A. Who's Watching You. Penguin Books, New York, 1981

13. Donner, F.J. The Age of Surveillance. Knopf. New York, 1980.

14. Early. P. Big Brother makes a date. San Francisco Exam. (Oct. 12. 1986).

15. Ellul, J. The Technological Society. Knopt. New York, 1964.

16. Greenberg. D.H.. and Wolf, D.A. Is wage matching worth all the trouble? Public Welfare (Winter 1985). 13- 20.

17. Greenleaf, G.W., and Clarke, R.A. Database retrieval technology and subject access principles. Aust. Comput. J. 16, 1 (Feb. 1984). 27- 32

18. Greenleaf, G.W.. and Clarke, R.A. Aspects of the Australian Law Reform Commission's information privacy proposals. J. Law and Inf Sci. 2, 1(Aug. 1986). 83-110.

19. Gross, M.L. The Brain Watchers Signet. 1963.

20. Hoffman, L.J., Ed. Computers and Privacy in the Next Decade. Academic Press, Now York, 1980.

21. Huxley. A. Brave New World Penguin Books, New York, 1975 (originally published in 1932).

22. Kircher, J. A history of computer matching in federal government programs. Computerworld (Dec 14. 1981).

23. Kling, R. Automated welfare client- tracking and service integration: The political economy of computing. Commun ACM 21.6 (June 1978). 484-493.

24. Kling, R. Value conflicts and social choice in electronic funds transfer system developments. Commun. ACM 21, 8 (Aug. 1978). 642- 657.

25. Kusserow, R.P, The government needs computer matching to root out waste and fraud. Commun. ACM 27. 6 (June 1984), 542- 545.

26. Langan, K.J. Computer matching programs: A threat to privacy? Columbia J. Law Soc. Probl. 15, 2 (1979)

27. Laudon, K.C. Computers and Bureaucratic Reform. Wiley. New York. 1974.

28. Laudon, K.C. Complexity in large federal databanks. Soc./Trans. (May 1979).

29. Laudon, K.C. Problems of accountability in federal databanks. In Proceedings of the American Association for the Advancement of Science (May). American Association for the Advancement of Science. 1979

30. Laudon, K.C. Data quality and due process in large interorganizational record systems. Commun. ACM 29. 1 (Jan. 1986). 4-11.

31. Laudon, K.C. Dossier Society, Value Choices in the Design of National information Systems. Colombia University Press, New York, 1986

32. Long. E.v. The Intruders. Praeger. New York, 1967.

33. Marx, G.T. The new surveillance. Technol. Rev. (May- June 1985)

34. Marx, G.T. I'll be watching you: Reflections on the new surveillance. Dissent (Winter 1985).

35. Marx, G.T., and Reichman, N. Routinising the discovery of secrets. Am. Behav. Sci. 27. 4 (Mar.- Apr. 1984). 423-452.

36. Miller, AR. The Assault on Privacy. Mentor, 1972.

37. Netor, A. Dossier. Stein and Day. 1974.

38. New South Wales Privacy Committee. Guidelines for the Operation of Personal Data Systems. NSWPC. Sydney. Australia, 1977.

39. Office of Management and Budget. Guidelines to Agencies on Conducting Automated Matching Programs. 0MB. Mar. 1979.

40. Office of Management and Budget. Computer Matching Guidelines. 0MB. May 1982.

41. Office of Management and Budget President's Commission for Integrity and Efficiency. Model Control System for Conducting Computer Matching Projects Involving individual Privacy Data. OMB/PClE. 1983.

42. Office of Technology Assessment. Federal government information technology: Electronic surveillance and civil liberties. OTA- ClT- 293, U.S. Congress, Washington, D.C.. Oct. 1985.

43. Office of Technology Assessment. Federal government information technology: Electronic record systems and individual privacy. OTA. ClT- 296, U.S. Congress, Washington, D.C.. June 1986.

44. Organisation for Economic Cooperation and Development. Guidelines for the Protection of Privacy and Transborder Flows of Personal Data. OECD. Paris, France, 1980.

45 Orwell, G. 1984. Penguin Books, Now York, 1972 (originally published in 1948).

46. Oxford Dictionary. vol. X 1933, p. 248.

47. Packard, Y. The Naked Society. McKay, New York, 1964.

48. Privacy Protection Study Commission. Personal Privacy in an information Society. U.S. Government Printing Office. Washington. D.C., July 1977.

49. Raines, J.C. Attack on Privacy. Judson Press, 1974.

50. Reichman, N., and Marx, G.T. Generating organisational disputes: The impact of computerization. In Proceedings of the Law and Society Association Conference (San Diego, Calif., June 6- 9). Law and Society Association, 1985.

51. Rodota, S. Privacy and data surveillance: Growing public concern Inf. Stud. 10, DECD. Paris, France, 1976.

52. Rosenberg, J.M. The Death of Privacy. Random House, 1969.

53. Rosenberg. R.S. Computers and the information Society. Wiley. New York, 1986.

54. Roazak, T. The Cult of Information. Pantheon, 1986.

55. Rule, J.B. Private Lives and Public Surveillance: Social Control in the Computer Age. Schocken Books. 1974.

56. Rule, J.B. Value Choices in E.F.T.S. Office of Telecommunications Policy. Washington, D.C.. 1975.

57. Rule, J.B. 1984-The ingredients of totalitarianism. In 1984 Revisited-Totalitarianism in Our Century. Harper and Row, New York, 1983, pp 166-179.

58. Rule. J.B. Documentary identification and mass surveillance in the United States. Soc. Probl 31, 222 (1983)

59. Rule. J.B. McAdam, D. Stearns, L. and Uglow, D. The Politics of Privacy. New American Library. 1980

60. Russell, B. Authority and the Individual. George Allen and Unwin. 1949.

61. Shattuck, J. Computer matching is a serious threat to individual rights. Commun. ACM 27, 6 (June 1984). 538-541

62. Stone, M.G. Computer Privacy. Anbar, 1968

63. Thom, J., and Thorne, P. Privacy legislation and the right of access. Aust. Comput. J. 15, 4 (Nov 1983). 145- 150

64. Thompson, A.A. A Big Brother in Britain Today. Michael Joseph, 1970.

65. U.S. Dept. of Health and Human Services. Computer Matching in State Administered Benefit Programs: A Manager's Guide to Decision- Making HEW. Washington, D.C.. 1983.

66. U S. Dept. of Health and Human Services. Computer Matching iii State Administered Benefit Programs. HEW, Washington, D.C., June 1984

67. U.S. Dept. of Health, Education and Welfare. Secretary's Advisory Committee on Automated Personal Data Systems

67. U.S. Dept. of Health, Education and Welfare, Secretary's Advisory Committee on Automated Personal Data Systems. Records, Computers and the Rights of Citizens. MIT Press, Cambridge. Mass., 1973.

68. U.S. Federal Advisory Committee on False Identification. The Criminal Useof False Identification. FACFI. Washington, D.C., 1976.

69. U.S. Senate. Oversight of Computer Matching to Detect Fraud and Mismanagement in Government Programs. U.S. Senate, Washington, D.C.. 1982.

70. Warner, M., and Stone, M. The Data Bank Society: Organisations, Computers and Social Freedom. George Allen and Unwin, 1970

71. Webster's 3rd Edition. 1976, p. 2302.

72. Weeramantry, C.G. The Slumbering Sentinels: Lawand Human Rights in the Wake of Technology. Penguin Books. New York, 1983.

73. Wessell, M.R. Freedom's Edge: The Computer Threat to Society. Addison- Wesley, Reading, Mass., 1974.

74. Westin, A.F. Privacy and Freedom. Atheneum, New York, 1967

75. Westin, A.F., Ed. information Technology in a Democracy. Harvard University Press, Cambridge. Mass , 1971.

76. Westin, A.F., and Baker, M. Databanks in a Free Society. Quadrangle, New York, 1974.

77. Yestingsmeier, J. Electronic funds transfer systems: The continuing need for privacy legislation. Comput. Soc. 13, 4 (Winter 1984). 5- 9

78. Zamyatin, Y. We, Penguin Books. New York, 1983 (originally published in Russian, 1920).

CR Categories and Subiect Descriptors: [Computer Applications]: J.1 Administrative Data Processing; K.4.1 (Computers and Society): Public Policy Issues; K.4.2 [Computers and Society): Social Issues; K.5.2 (Legal Aspects of Computing]: Governmental Issues

General Terms: Human Factors, Legal Aspects, Management, Security Additional Key Words and Phrases: Data protection, data surveillance, dataveillance, front- end verification, mass surveillance. matching schemes, profiling surveillance

---

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or spocific permission.

Navigation

Go to Roger's Home Page.

Go to the contents-page for this segment

Send an email to Roger

Created: 7 October 1996

Last Amended: 26 January 1998

These community service pages are a joint offering of the Australian National University (which provides the infrastructure), and Roger Clarke (who provides the content).

The Australian National University
Visiting Fellow, Faculty of
Engineering and Information Technology,
Information Sciences Building Room 211

Xamax Consultancy Pty Ltd ACN: 002 360 456
78 Sidaway St
Chapman ACT 2611 AUSTRALIA
Tel: +61 2 6288 1472, 6288 6916