Computer Matching by Government Agencies: The Failure of Cost/Benefit Analysis as a Control Mechanism

Roger Clarke

Principal, Xamax Consultancy Pty Ltd , Canberra

Visiting Fellow,Department of Computer Science, Australian National University

Version of November 1994

This paper was published in Information Infrastructure & Policy 4,1 (March 1995) 29-65

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

Abstract

During the twentieth century, public administration has increasingly involved intensive use of data about individuals. The explosion of network traffic is making vast additional volumes of transaction data available. Monitoring of individuals through their data is both possible and much cheaper than conventional physical and electronic surveillance. As a result, data surveillance is burgeoning. Computer matching is a key facilitative mechanism in the monitoring of populations.

External controls over dataveillance activities may not be necessary if intrinsic mechanisms are adequate. Cost/benefit analysis (CBA) is a well-established means of assessing the net value of a project. Used effectively, it should ensure that data processing applications are not commenced, and not continued, unless there are net benefits. This paper reports on research into the extent to which the use of CBA has prevented unjustified uses of computer matching by government agencies in the United States and Australia.

The conclusions reached are that:

- properly conducted CBA is capable of acting as a significant restraint on data surveillance schemes, provided that a number of conditions hold;
- those conditions have not existed in either the United States or Australia;
- intrinsic economic controls have therefore been largely ineffectual; and
- regulatory measures are necessary if CBA is to act as a control over unjustified use of data surveillance.

Contents

1. Introduction

There has long been an assumption that natural or intrinsic controls exist which will tend to constrain the excessive use of privacy-invasive applications of information technology. The tradition was established by the seminal works of Westin (Westin 1967, 1971; Westin & Baker 1974). It was recently revived by Laudon (1993), whose argument is based on economic grounds.

Economic controls over unreasonable uses of technology may take various forms. Their formal expression, however, is in cost/benefit analysis (CBA). If the Westin/Laudon thesis holds, then it is to be expected that CBA would have worked as a control over the use of data surveillance. The purpose of this paper is to test that contention, in relation to one particular and prevalent data surveillance technique, computer matching.

Computer matching is the comparison of machine-readable records containing personal data relating to many people, in order to detect cases of interest. The technique is called 'computer matching' in the United States, and 'data matching' in Australia and Canada. Computer matching became economically feasible in the early 1970s, as a result of developments in information technology (IT). The technique has been progressively developed since then, and is now widely used, particularly in government administration and particularly in the three countries mentioned above.

The paper commences by reviewing data surveillance in general, and providing an overview of the particular technique of computer matching. It then summarises CBA, as a prelude to an examination of the use of CBA in relation to computer matching in two countries. This is followed by an evaluation of CBA's effectiveness, which concludes that it has been ineffectual, and will remain so unless and until an appropriate set of extrinsic controls is imposed. A normative model is outlined, which it is claimed satisfies the requirements.

2. Data Surveillance

Surveillance is the systematic investigation or monitoring of the actions or communications of one or more people in order to collect information about them, their activities or their associates. Surveillance has long been undertaken through physical observation, and various technologies have been harnessed to support it, including telescopes, cameras, directional microphones, and more recently electronic tools of various kinds such as telephone 'bugs'.

Conventional forms of surveillance are labour-intensive, and therefore time-consuming and expensive. This has ensured that there has been an economic disincentive against its widespread use, and societies in which surveillance was very common have been unusual and regarded in modern democratic countries as almost pathological. Recent instances included the Soviet Union, East Germany under the Stasi, Romania and Communist China, particularly during the 'Red Guard' phase.

Data surveillance, usefully abbreviated to 'dataveillance' is the systematic use of personal data systems in the investigation or monitoring of the actions or communications of one or more people in order to collect information about them, their activities or their associates. Where the surveillance relates to an identified person, for a specific reason, it is 'personal dataveillance'. Where a (usually large) group of people is monitored, in order to identify interesting individuals, it is 'mass dataveillance'.

A variety of different dataveillance techniques are in common usage, including:

- personal dataveillance techniques:
    - the screening or authentication of exceptional transactions against internal norms;
    - front-end verification of exceptional transactions, against data relevant to the matter at hand, from internal databases and third party data;
    - front-end audit of exceptional individuals, against data related to<u>other</u> matters, from internal databases and third party data; and
    - cross-system enforcement;
- mass dataveillance techniques:
    - the screening or authentication of <u>all</u>transactions against internal norms;
    - front-end verification of <u>all</u> transactions, against data relevant to the matter at hand, from internal databases and third party data;
    - front-end audit of <u>all</u> individuals, against data related to<u>other</u> matters, from internal databases and third party data; and
    - analysis of data on internal databases and/or from third parties, including:
        - single-factor file-analysis; and
        - multi-factor file-analysis, or profiling;
- facilitative techniques:
    - the integration of data stored in dispersed databases;
    - data concentration, through organisational merger or the establishment and operation of data-interchange networks and hub systems; and
    - computer matching, the comparison of large quantities of personal data maintained by two or more data systems.

In comparison with conventional forms of surveillance, dataveillance is capable of being automated, and hence cheaper and more reliable. It therefore overcomes the economic disincentive, and thereby removes the most basic and effective of the protections against excessive use of surveillance. As a result, its use has burgeoned during the last 15 years. It has been initially applied in nations with sophisticated information technology capabilities, but is likely to be increasingly attractive also in emerging nations, many of which have less developed civil liberties traditions and mechanisms.

A comprehensive treatment of dataveillance is in Clarke [1988]. For other reviews, see Rule [1974], Laudon [1974], Smith

[1974-, 1976-], Westin & Baker [1974], Kling [1978], Rule et al [1980], OTA [1986b], Laudon [1986], Flaherty [1989] and Bennett [1992].

3. Computer Matching

(a) Introduction

Computer matching is a particular dataveillance technique. It involves the comparison of sets of machine-readable personal data records relating to many people, in order to detect cases of interest. Its use offers potential benefits, particularly financial savings. It is also error-prone, and its power results in threats to established patterns and values. The imperatives of efficiency and equity demand that computer matching be used, and the information privacy interest demands that it be used only where justified, and be subjected to effective controls.

A comprehensive description of the technique is in Clarke (1994b). It is predicated on the possession by one or more source organisations of one or more databases containing personal data records. A 'matching algorithm' is applied to them in order to find 'raw hits' or 'matches', in which the person to whom one record relates is inferred to be the same person to whom one or more other records relate. An 'inferencing procedure' is applied to the raw hits in order to draw conclusions about the person to whom the data purports to relate, or to his or her behaviour, actions or proclivities.

(b) Nature and Origins

Large numbers of organisations in federal, state and local government, and in the private sector, hold large amounts of data about individuals. In many circumstances, organisations, to protect their interests, seek confirmatory data.

Many such cross-checking activities are triggered by a transaction between the person and the organisation; for example an application for a job, a pension or a loan. This form of cross-checking activity is commonly referred to as 'front-end verification', and is a form of 'personal dataveillance'.

Cross-checking may also be undertaken in the absence of such a trigger. The motivation may be a generalised belief or suspicion that some people with whom the organisation deals may be transgressing standards, or that data concerning some people with whom the organisation deals may be incorrect and that the organisation's interests may be thereby harmed. This approach represents 'mass dataveillance' of a whole population, for a reason related to some suspected but as yet unidentified portion of that population.

Computer matching facilitates mass dataveillance. Its primary differences from front-end techniques are that it is undertaken some time after the event, and that it is applied to large numbers of people and records. The technique is used to detect people who may be of interest to the organisation conducting the match, or to its clients. For the most part, matching has been undertaken by government agencies, for the purpose of identifying people who may have (intentionally or unintentionally) received excessive benefits, or failed to pay appropriate taxes.

There are other, closely related facilitative mechanisms for mass dataveillance. One is 'data-linkage', by which is meant the storage in an individual's record on one file of that person's identifier in one or more other files, to enable prompt and reliable inter-relationship of data in the future. A second, 'data concentration', involves the merger of databases, or creation of new databases, to support a number of functions. This has been referred to as the 'national databank' issue, particularly in the United States, where a major discussion took place in the mid-1960s. A third mechanism is the use of a 'common, multi-purpose identifier'. This has given rise to debates over national identification schemes such as the Social Security Number (SSN) in the United States, the Social Insurance Number (SIN) in Canada, and in Australia the withdrawn Australia Card proposal and the Tax File Number (TFN).

The first computer matching program is generally identified in the literature as 'Project Match', conducted in 1977 by the then Department of Health, Education & Welfare (HEW), now the Department of Health & Human Services (HHS). Project Match compared the records of recipients of Aid to Families with Dependent Children (AFDC) with the payroll records of about 3 million federal employees. It was claimed to be a great success. It identified 33,000 raw hits, later filtered to 7,100, resulting in 638 internally investigated cases, of which 55 resulted in prosecutions [OMB 1986b p.18; see also Fischel & Siegel 1980, Weiss 1983, OTA 1986b p.42, Early 1986]. It appears, however, that these prosecutions resulted in only about 35 convictions, all for minor offences, with no custodial sentences and less than $10,000 in fines. This paradox of a project being hailed as a great success when the measurable financial costs are high and the measurable financial benefits very low, has been a feature of matching programs from the very beginning. With a few significant exceptions, it continues to be so.

By 1982, it was estimated that U.S. state and federal agencies routinely carried out about 200 programs [Cohen 1982]. The Reagan Administration instituted a drive for efficiency in government, and the President's Council on Integrity and Efficiency in Government (PCIE) "has been the strongest proponent of computer matching as a management tool" [Flaherty 1989, p.344]. The Congress' Office of Technology Assessment estimated a tripling in use between 1980 and 1984 [OTA, 1986, p.37]. Laudon estimated 500 programs in 1986 [1986, p.383].

In 1984-86, a GAO study found that the "current climate or environment surrounding computer matching" was a primary determinant of growth in its use. In particular, the report noted "(a) a rising concern about erroneous payments, (b) technological developments that make computer matching easier or more feasible ..., (c) reports of successful matches with large cost savings or cost avoidances, and (d) endorsement and recommendations by key oversight organisations" [GAO 1986c, p.2]. These are more consistent with a 'fashion' model of decision-making than a 'rational management' model. Unsurprisingly, growth continued.

The U.S. Federal Government has not only applied the technique within and among its own agencies, but it has also imposed requirements on State government agencies. A succession of statutes, culminating in the 1984 Budget Deficit Reduction Act, required state administrations to match data from a variety of their own data systems and those of other agencies of the same State, agencies of other States, Federal Government agencies, and the private sector. These requirements are supported by fiscal sanctions. As a result, some hundreds of interstate government matching programs are also in place, as are programs involving interchange between different, particularly adjacent, local government areas. See Kusserow [1983, 1984b] and GAO [1990c, p.25].

For more detailed historical accounts of computer matching in the United States, see Smith [1974 et seq], Kirchner [1981], Cohen [1982], Azrael [1984], Laudon [1986, pp.328-335], Greenberg & Wolf [1985, 1986], OTA [1986b] and Flaherty [1989, pp.344-358].

Computer matching is used for many different purposes, including the detection of errors and illegal behaviour, person location, confirmation of continuing eligibility for a benefit, and data quality audit. It can also contribute to other objectives such as supporting actions with financial benefits (e.g. the cancellation of unwarranted benefits and the collection of debts) and the construction and maintenance of databases.

The majority of contexts of use have to do with social control and efficient government administration, in particular:

- the comparison of data held by taxation authorities with that of financial institutions;
- the comparison of data held by social welfare agencies with that held by other agencies. In the United States in 1990, the Social Security Administration conducted 42 such programs, and its partners included many other welfare agencies, pension funds, immigration authorities, the parent locator service, the public debt administration agency, health and general insurers, prisons and death registries;
- the comparison of data held by housing authorities with income information from various sources;
- in law enforcement; and
- in other contexts, such as national service during the Vietnam War era, when Social Security records (which contain the dates of birth of the overwhelming majority of young people) were matched with files from the Department of Defense to select out those already in the military, and with Internal Revenue Service files to obtain mailing addresses [Laudon 1986, pp.331].

Computer matching is also used by the private sector for its own purposes; for example in the construction of consumer profiles through the merger of mailing lists. The most significant aspect of private sector participation appears, however, to be as providers of data to government agencies. This paper's focus is accordingly restricted to the use of computer matching by government agencies.

(c) Technical Description

No definition of the term 'computer matching' is to be found in computing dictionaries. Nor does the term appear in such landmark documents as Westin [1967, 1971], HEW [1973], Westin & Baker [1974], the U.S. Privacy Act 1974, FACFI [1976], PPSC [1977], NSWPC [1977], Lindop [1978] and OECD [1980]. Indeed in most of these references it is difficult to even trace the emergence of the concept. The technique has only become practicable as a result of the developments in information technology since the mid-1970s, and the term appears to have come into currency only after publication of descriptions of Project Match, conducted in 1977. The series of important documents through which the subsequent line of development can be traced includes OMB [1979a, b, 1982a, b], OMB/PCIE [1983], HHS [1983a, b, c, 1984], ALRC [1983], Laudon [1986], OTA [1986b, pp.37-63], SMOS [1987], the U.S. Computer Matching and Privacy Protection Act 1988, PCC [1989], O'Connor [1990] and PCA [1990].

The technique may be described in many different ways. This paper adopts as its working definition:

> *computer matching is any computer-supported process in which personal data records relating to many people are compared in order to identify cases of interest.*

A 'computer matching run' is an event in which one file is compared against one or more files. A 'computer matching program' is a set of one or more computer matching runs which are very similar in nature (in terms of the data which are accessed, the matching and inferencing criteria applied, etc.), and are undertaken to assist a single organisation or set of organisations in addressing a single set of objectives.

The actual organisational structure and processes vary significantly between agencies and programs. This section presents a model designed to accommodate all of the important variations, and in particular GAO [1986a p.9], HUD [1988], GAO [1988 p. 57], PCA [1990 pp.4-5] and DSS [1991 p.15].

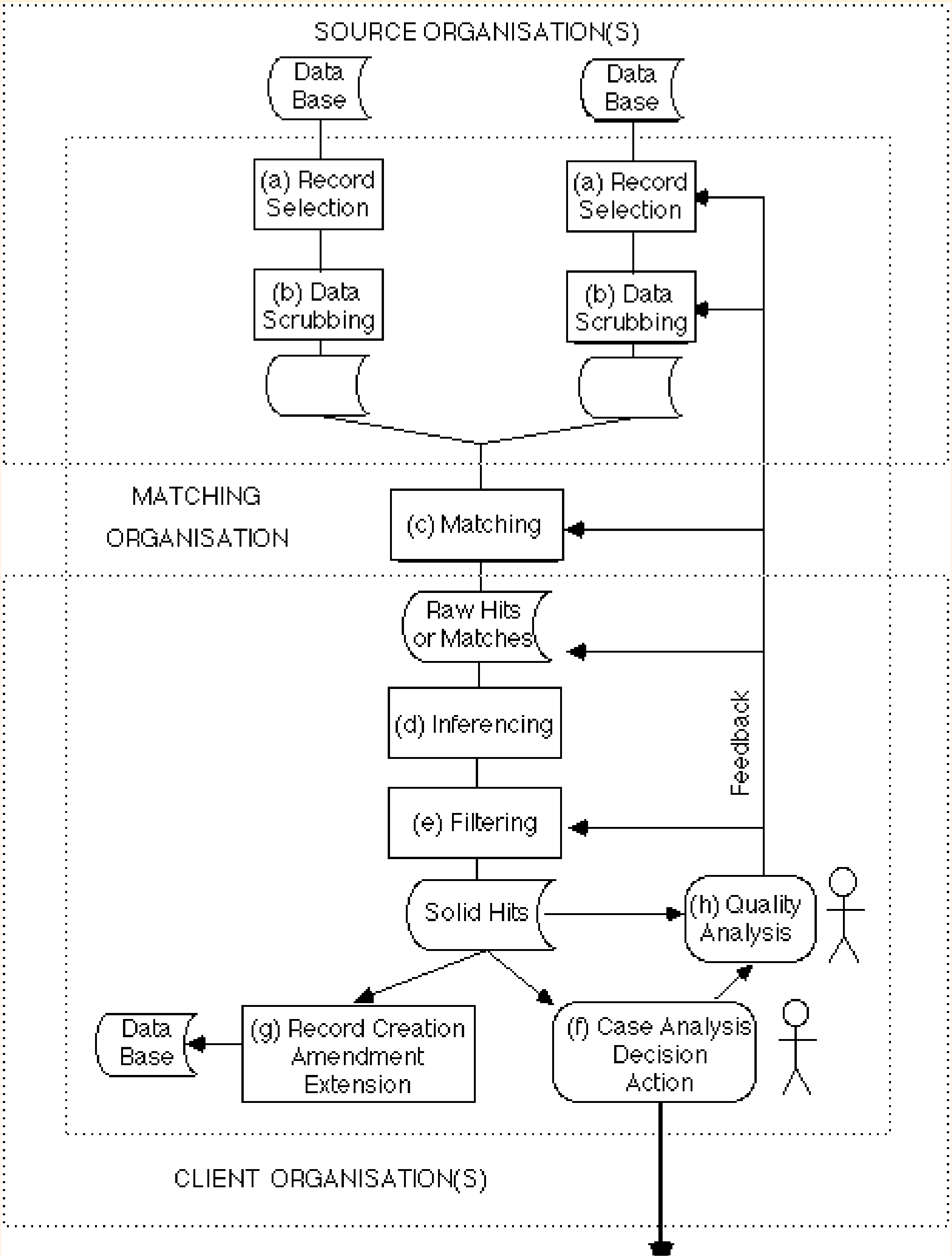
The participant organisations in computer matching can be classified into three groups:

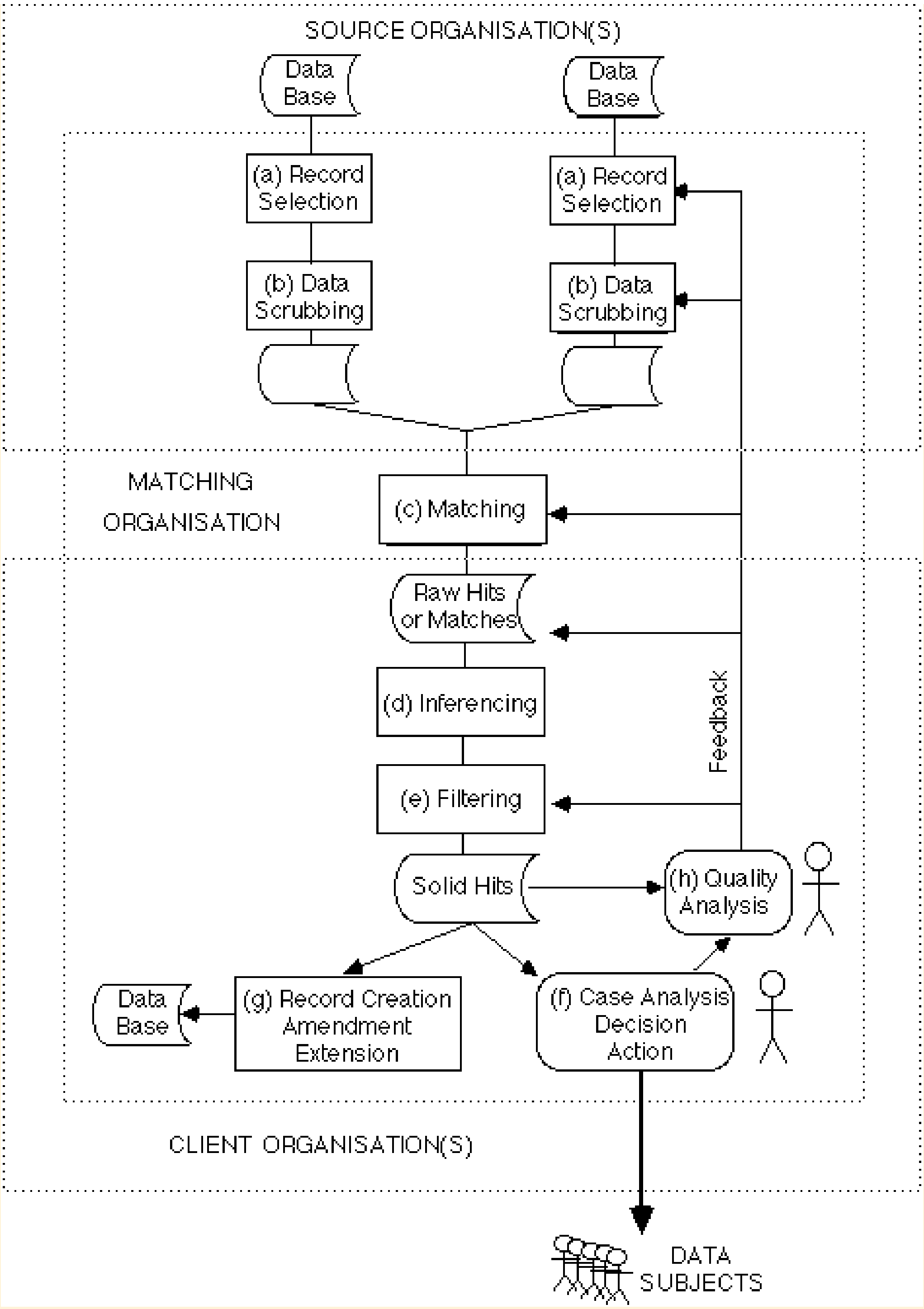
- a matching organisation, which undertakes the matching step and may also undertake other steps in the process;
- one or more source organisations, which provide input data and may also perform other steps; and
- one or more client organisations, which receive data, in order to make decisions and take actions and/or to maintain databases.

In any particular computer matching program, participants may include government agencies, corporations, or (less usually) unincorporated associations or individuals; and a single government agency may be active in more than one of these roles. The scope of this paper extends to computer matching in which any of the participant organisations is a government agency.

Computer matching is predicated on the possession by source organisations of at least one, usually two, but perhaps more databases containing personal data records. The steps in computer matching are depicted in Figure 1.

Figure 1: Conceptual Model of the Computer Matching Process





A detailed discussion of these steps is to be found in Clarke (1994b). In outline, they are:

1. selection from each database all or a sub-set of the records;
2. optionally, 'data scrubbing' operations, to change the organisation, format and/or content of one of more of the data sources into a form suitable for the matching step;
3. 'matching', whereby a 'matching algorithm' is applied to the file(s) of personal data records in order to find 'raw hits' or 'matches'. Generally these are matched pairs of records which are deemed to refer to the same data subject. In some cases, the algorithm may, however, involve a search for records on one file for which a match on the other canot be found;
4. 'inferencing', which involves an 'inferencing procedure' being applied to the outcome of the matching process (i.e. to the contents of the matched pairs of records, or to the existence or non-existence of matches), to draw conclusions about the person to whom the data purports to relate, or to his or her behaviour, actions or proclivities;
5. 'filtering' of the 'raw hits' to produce 'solid hits', and thus ensure efficient use of investigative resources and avoid insufficiently justified privacy-invasiveness;
6. analysis of the resulting information, and decision-making and action arising from it;

6. analysis of the resulting information, and decision-making and action arising from it;
7. optionally, creation of new records and/or amendments or extensions to existing records; and
8. optionally, quality analysis activities may generate feedback from any and all stages back to earlier steps.

This description has related to what might be termed 'identifier-based matching'. An alternative approach is technically feasible and increasingly economically feasible, in which the match is based not on one or more explicit identifiers but on the content of any and all data-items deemed relevant by an investigator, such as address, physical characteristics such as height and weight, or association with a place and time, other individuals or groups, or corporations. This more sophisticated approach, which might be referred to as 'content-based matching' is not further discussed in this paper.

(d) Conclusions

Computer matching is a powerful dataveillance technique, capable of offering benefits to the efficiency and effectiveness of government business greater than its financial costs. It is also a highly error-prone and privacy-invasive activity. Unless a suitable balance is found, and controls are imposed which are perceived by the public to be appropriate and fair, its use is liable to result in inappropriate decisions, and harm to people's lives. In a tightly controlled society, this is inequitable. In a looser, more democratic society, it risks a backlash by the public against the organisations which perform it, and perhaps against the technology which supports it.

Because of the current fashion of highly information-intensive procedures, the current boom in identification-based computer matching activity may be expected to continue for some years. Further refinements may be confidently expected in the data scrubbing, matching, inferencing and filtering steps, including the application of such techniques as direct access, multiple-file matching, associative memory, expert systems, neural networks and fuzzy logic.

Computer matching is a critical test of the resolve of information technologists to accept responsibility for the impact of their body of knowledge on people. It is not, in itself, an evil; but it is capable of being used evilly, or so insensitively that it will do significant harm to individuals, to groups, and to society as a whole. It shares that feature with many other techniques which are not yet empirically researchable, such as profiling, public networking, voice recognition, virtual reality in entertainment and education, the substitution of digital simulation for physical experimentation and intelligent robotics. It is vital that research be undertaken on such topics, and that that research be reported on in journals which reach the wide spread of academics and professionals, and not just discussed among a small clique of 'socially aware' fringe-dwellers (Clarke 1988).

4. Controls Over Computer Matching

The use of computer matching involves serious invasions of privacy. For a detailed analysis, see Clarke (1988, 1994c). Controls are essential, to ensure that the technique is applied only where the inherent invasiveness is justified, and that it is undertaken in such a manner as to minimise the harm arising from that invasiveness.

Intrinsic controls exist which tend to restrain agencies from applying computer matching, or to cause them to do so carefully. These include:

- the exercise of countervailing political power by the class of data subjects affected by the process, by their representatives, by the mass media, or by the general public. Given the imbalance of power between organisations and individuals, it is not realistic to expect this factor to be of any great significance except in particular circumstances;
- the displeasure of some organisation, such as a competitor or regulatory agency;
- self-restraint practised by the agency itself, influenced by professional norms, or by an appreciation of the delicacy of public confidence in its institutions and the resultant need to respect constitutional rights and moral concerns; and
- general blundering.

The intrinsic factor which might be expected to exercise the most significant degree of control over computer matching is economics: surely government agencies will not apply the technique in circumstances in which it is not worthwhile. The primary means whereby the economic factor will influence decision-making about computer matching programs is cost/benefit analysis, which is the primary focus of this paper.

Extrinsic controls also exist. The U.S. Congress acted on a number of occasions between 1970 and 1983 to alter the balance of power between citizens and the State. Laudon [1986, pp.372-4] lists fourteen major pieces of legislation in such areas as credit reporting, criminal justice information systems, education, taxation, banking, electronic funds transfer and debt collection. All were created within the framework set by Westin [Westin 1967, Westin 1971 and Westin & Baker 1974]. Westin had found no problems with extensive surveillance systems as such, only with the fairness of the procedures involved.

The centrepiece of this legislation, the Privacy Act of 1974, nominally precluded government agencies from transferring personal data among themselves without explicit legislative authority, or the consent of the data subject. In practice, the Act was quickly subverted. Publication of uses in the Federal Register proved to be an exercise in bureaucracy rather than control. The 'routine use' loophole in the Act was used to legitimise 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 as a whole). See PPSC [1977], Marx & Reichman [1984, p.449], OTA [1986b at pp.16-21] and Berman & Goldman [1989].

The original Bill had proposed to establish a permanent Privacy Commission, to ensure compliance by agencies with the law. Under threat of Presidential veto, this was replaced by a short term study commission, and oversight of the Act made the responsibility of the Office of Management and Budget (OMB). Unsurprisingly, given its primary responsibilities, OMB has consistently interpreted the Act in a manner unsympathetic to information privacy concerns [see, for example, Kirchner 1981, Laudon 1986, p.374-5 and Flaherty 1989, pp.346-9].

'Guidelines to [Federal] Agencies on Conducting Automated Matching Programs' were first created in 1979 [OMB, 1979a, b]. A much shorter, revised version was later issued [OMB, 1982a, b], which weakened some aspects of the guidelines, especially in relation to cost/benefit justification and demonstration that alternative means were insufficient. This appears to have resulted from submissions by agencies to the effect that such analyses were expensive and/or that programs were difficult to justify on the basis of quantifiable benefits, together with pressure from the President's Commission for Integrity and Efficiency. Subsequent regulatory documents include OMB/PCIE [1983] and OMB [1983].

At no stage does the Privacy Act or the OMB Guidelines appear to have provided any significant form of restraint on computer matching [Flaherty 1989, pp.349-350], and there is no mechanism whereby the Guidelines are enforced [Flaherty 1989, p.357, quoting a report by a Committee of the House of Representatives]. Very few of the matching programs which have been, or are being conducted, appear to have ever been the subject of explicit congressional approval or review. Laudon concluded that "as it stands now, the reigning principle is that virtually any federal system may be developed in the absence of a clearcut injury to individuals (the 'bodies floating in the river' criterion)". This colourful expression refers to a reported quip from the Secretary of the agency responsible for the original Project Match in 1977 [Laudon 1986, pp.382].

In 1990, the Computer Matching and Privacy Protection Act 1988 became law. It contains manifold exceptions and exemptions, has been subjected to restrictive interpretation by OMB, has been largely ignored by agencies of State governments inconvenienced but it, and has not been enforced. It is further discussed in a later section of this paper.

Extrinsic factors have not resulted in the exercise of effective control over computer matching in the United States. The regulatory regime in Australia is quite different, but similarly ineffectual (Clarke 1994a). It is critically important, therefore, that the most important of the intrinsic controls, the economics of the technique, have a significant impact. This paper accordingly focusses on the question as to whether or not cost/benefit analysis of computer matching schemes acts as an effective control over the technique's application.

5. Cost/Benefit Analysis

(a) Cost/Benefit Analysis in Principle

Cost/Benefit Analysis is a technique, well-grounded in microeconomic and management accounting theory, for assessing the net social value of a program or measure [Dasgupta & Pearce 1972, Mishan 1977, Sassone & Schaffer 1978, Thompson 1980, Gramlich 1981, Sugden & Williams 1985]. It has also been described in government publications [DoF 1991, 1992], and in forms designed expressly for use in relation to information technology applications [APSB 1981, DoF 1993].

CBA is undertaken from the perspective of the economy as a whole, not from that of any particular individual, organisation or group, and hence considers all gains and losses arising, regardless of to whom they accrue. Its aim is to ensure efficiency in the allocation of resources to society's aims. It is distinguished from financial evaluation, which is conducted from the viewpoint of an individual firm or agency. Although originally developed to assist in the evaluation of project proposals, it is also appropriately applied to those larger groupings of activities conventionally referred to as 'programs'. CBA is equally applicable to decisions about whether to commence a program, or to continue one that it already operational.

The technique involves the identification of all of the costs and benefits arising in relation to a program, and to the extent practicable and economic, their measurement. In decreasing order of desirability, the resulting data may comprise financial measures, financial estimates based on quantified measures, financial estimates based on quantified estimates, quantified measures of non-financial factors, quantified estimates, and qualitative descriptions. It is seldom possible for significant projects or programs to be evaluated using financial factors alone. It is therefore important that the judgement of executives and public representatives be applied to the full set of data relevant to the decision: "inevitably, some costs and benefits resist the assignment of dollar values. These ... are separately presented to the decision-maker ... with as much descriptive information as possible ... for assessment in conjunction with the quantified estimate of the net social benefits of the activity" [DoF 1991, pp.1,9].

An important element of cost/benefit analysis is the treatment of 'opportunity costs', "the benefits foregone from not having done something else with the resources" [GAO 1986b, p.27]; for example, in order to use the results of a matching program to prevent over-payments, an agency may need to divert resources from some other activity, such as the audit of a randomly selected sample of cases, or the follow-up of 'tip-offs'. Where labour is diverted from a profitable activity to a (perhaps temporarily) unprofitable one, it must be valued according to the more profitable alternative use. Hence, in addition to the wages and salaries paid, the net benefits foregone must be treated as a cost of the program, since, in the absence of the program, they would have been realised: "Costs are valued according to the willingness of others to pay for the resources involved and therefore reflect the best alternative foregone" [DoF 1991, p.16].

Further difficulties arise in respect of 'joint' costs and benefits, by which is meant costs and benefits which are attributable to multiple programs. In evaluating a particular program, care must be taken to identify and count that, and only that, share of costs and benefits which is appropriate. This requires the avoidance of implicit double-counting of costs or benefits, either in each of several related programs, or indeed within a single program.

In any environment in which early investment is necessary in order to achieve a significantly later return, the evaluation of costs and benefits needs to take account of the time-value of money. The common technique for doing so is to discount future monetary returns back to the base-year in which investment commenced, using current interest rates net of inflation. Other, less effective metrics are sometimes used, including the payback period, the discounted payback period, internal rate of return and the benefit/cost ratio [DoF 1991, pp.112-8]. It is necessary to select an appropriate rate to reflect the cost of capital and the risk inherent in the project. It is also important to properly handle the effects of inflation, either by adjusting all figures to dollars of the same time-period and using a 'real' discount rate, or by expressing all figures in contemporary dollars and using a discount rate adjusted for inflation (e.g. 8% real plus an allowance of 6.5 percentage points for a 6% inflation rate) [DoF 1991, pp.42-58].

It is also important that analysts take into account the risk and uncertainty involved in programs, including financial, commercial, industrial, technological and legislative risks [DoF 1993, pp.17-8]. CBA implicitly assumes a risk-neutral decision-maker, who uses the expected values of the outcomes, and is indifferent to the possibilities of significantly better or worse outcomes. There are many circumstances in which this is not appropriate, particularly where the 'downside' is extreme, or the distribution of costs and benefits falls unevenly on the population. The generally recommended approach to addressing risks is sensitivity analysis, supported by the technique of risk analysis in more complex cases [DoF 1991, pp.59-74].

Drawbacks of CBA include its misleading precision, and the scope for analyses to be manipulated to serve vested interests. These weaknesses can be addressed by clearly describing qualitative factors rather than using unjustified estimates to quantify them, and by clearly documenting the calculations whereby estimates are arrived at, and the assumptions underlying the estimating process.

Finally, cost/benefit analysis must be itself the subject of cost/benefit analysis. By this is meant that the ultimate objective is not a justification which is unassailable, but a reasoned judgement as to whether the activity is worthwhile to its clients, i.e. society as a whole. This places constraints on the resources and time which should be invested in refining the elements of a cost/benefit analysis. It does not, however, provide an excuse for omitting cost/benefit analysis, since without a sufficient basis for judgement, an agency will make inappropriate commitments of resources and achieve an inappropriate balance between control, service and other objectives. Generally, "the larger the project, the greater the resources at stake, and so the more that can usually be justified in terms of expenditure on information and analysis" [DoF 1991, p.12]. Means of improving the economics of CBA include investing in cost-benefit models appropriate to particular kinds of program.

(b) Cost/Benefit Analysis and Computer Matching

General background to the application of CBA to computer matching has been available for some time [GAO 1986b]. Guidelines have been published by the privacy 'watchdogs' of both Canada and Australia [PCC 1989, PCA 1992], and in the United States by the Office of Management and Budget [OMB 1979a, 1979b, 1982a, 1982b, 1983, 1989]. One U.S. agency is known to have promulgated comprehensive internal guidelines [SSA 1990a]. The practice is subject to statutory regulation in New Zealand.

The technique needs to be applied either on a continuing basis or at key points within a program. In particular, there is a need for pre-match, concurrent and post-match analyses [GAO 1986b, p.82-84]. In order to provide guidance for the evaluation of future proposals, both within and beyond the particular agency, it is important that concurrent and post-match analyses be related back to the original pre-match justification.

It is critical that all costs and benefits be included. In particular, the preliminary costs of initiating, developing, evaluating and deciding upon a proposed program are relevant. So too are all costs associated with creating software, amending software, negotiating with other agencies, auditing data, and gathering and scrubbing data. At the other end of the process, it is important that all costs of investigation, verification, communication, relevant case-management costs and all recovery costs including prosecution, appeals-handling and garnisheeing, be netted against the actual recoveries [GAO 1986b, p.26]. Canadian federal agencies have been provided with guidance regarding cost/benefit analysis, including a list of some twenty instances of direct, data processing, telecommunications, travel, training and consultant/ contractor costs, together with five instances of quantifiable savings. The costs of a matching program relative to its benefits "should be analyzed in terms not of the total cost in dollars but in terms of institutional resources, e.g. staff, equipment and materials, needed to perform [the] matching program in dollars but in terms of institutional resources, e.g. staff, equipment and materials, needed to perform [the] matching program and the amount of effort involved in developing and implementing it" [PCC 1989, p.8].

Cost/benefit analysis must be undertaken from an appropriate perspective and with an appropriate scope: "The goal of most computer matches ... should be to maximize benefits for or minimize costs to the general public" [GAO 1986b, p.23]. Accordingly, the scope of cost/benefit analysis should include the costs and benefits that arise in respect of not only the matching agency, but also all other organisations, both agencies and outside government, which are involved as sources, clients and recipients or otherwise incur costs or gain benefits. In addition, the costs of and benefits to individuals of all kinds must be included. The sole exception is that costs imposed on miscreants and criminals as a direct or indirect result of the match should not be treated as costs of the program.

It is also important that estimates reflect worldly imperfections, such as low quality data, the scope for serious criminals to counter or subvert the scheme, the evidentiary requirements to secure convictions and the collectability of, in particular, overpayments of welfare benefits and underpayments of taxes.

6. Research Method

The objective of the research reported on in this paper was to assess whether CBA has acted as a control mechanism over the use of computer matching. The research question involves a search for causality within a complex context. Hence a large number of factors is involved, and considerable depth of understanding is required. In these circumstances, experimental methods, quasi-experimental methods and questionnaire-based surveys are inappropriate.

In addition, there has been a long-standing reluctance by government agencies to make information publicly available concerning their computer matching activities. There is clearly some merit in the argument which they bring forward to the effect that the effectiveness of schemes may be compromised by public knowledge about their methods. This is insufficient to justify the shroud of secrecy that prevailed, and it seems reasonable to assume that there have been many schemes which were undertaken with dubious legality, justification and/or integrity procedures, and which agencies preferred not to submit to public scrutiny.

The situation has changed in recent years, however. In 1988, the United States Computer Matching and Privacy Protection Act effectively authorised all matching programs in return for some limited controls. By 1991-92, officers of U.S. agencies were robustly self-confident in their dealings with requests for information. Similarly, the Australian Privacy Act of 1988 had the effect of legitimising virtually all computer matching activities in that country. As late as 1987, Freedom of Information requests to Australian government agencies had resulted in precisely no meaningful information about their practices. By the early 1990s, however, some sunlight was beginning to find its way into the recesses of governmental activities, and empirical research, while still fraught with difficulties, at least became feasible.

The method adopted for this research combined several techniques. Considerable recourse was made to existing literature. There have been very few publications in the information systems and computer science literature, but several other disciplines, particularly law, sociology and political science provided some sources, as did government reports, particularly by the U.S. Office of Technology Assessment (OTA), and the proceeedings of Congressional and Australian Parliamentary hearings.

Field work was undertaken in two countries which are leaders in the use of computer matching, the United States [Clarke 1991b, 1992b, 1992c], and Australia [Clarke 1993b, 1994a]. This involved interviews with executives and operational staff in a variety of U.S. and Australian agencies, and the careful examination and analysis of written materials provided by those organisations, or otherwise publicly available. Wherever possible, multiple perspectives on the same object of study were used, such as internal documents, documents published by the agencies concerned, and reports by regulatory agencies (in the United States the General Accounting Office - GAO, and in Australia the Australian National Audit Office - ANAO and the Privacy Commissioner) and Parliamentary Committees. In addition to broad cross-sectional study, a single application was studied in depth, and longitudinally. The following sections report the outcomes.

7. Review of CBA and Computer Matching in the United States

The first major matching program was undertaken in 1976-77. The U.S. Government's original 'Guidelines to [Federal] Agencies on Conducting Automated Matching Programs' [OMB 1979] required that cost/benefit analyses be undertaken prior to a program being commenced. However, there are significant difficulties in undertaking a cost-benefit analysis of such a program. Many benefits are vague and unquantifiable, many expenses are hidden or already 'sunk', programs can be difficult to justify on the basis of quantifiable benefits, and justification exercises are expensive. Following lobbying by agencies, and pressure from the President's Commission for Integrity and Efficiency, a much shorter set of Guidelines was issued [OMB 1982], which in effect rescinded the requirements for cost/benefit justification and demonstration that alternative means were insufficient [OMB 1982. See also OMB/PCIE 1983, OMB 1985 and GAO 1986b, p.11].

There have been claims of dramatic success for matching schemes, but these have generally been made by the agencies which conducted them. There is not a great deal of data available, and little has been measured, or even subject to review by, an independent organisation. Anecdotal evidence, however, throws serious doubt on the efficacy of many matching programs.

In the original Project Match in 1977, the then social welfare agency (HEW) ran its welfare files against its own payroll files. The 33,000 raw hits which were revealed, and then reduced to 7,100, 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, five were dismissed, four pleaded guilty to misdemeanours (theft under $50), and only six were convicted of felonies. No prison sentences resulted, and the fines totalled under $2,000 [Early 1986, OTA 1986b, p.42].

"In a typical match of the welfare rolls of 34 jurisdictions involving over 5 million records, 35,000 cases appeared in which persons seemed to be receiving public assistance in more than one state ... that is seven-tenths of 1 percent ... Due to a number of errors in the underlying data records ..., less than one-fifth of these cases ... represented truly fraudulent recipients ... The match was declared a success" [Laudon 1986, p.332]. Less negative reports are to be found in Greenberg & Wolf [1984, 1985, 1986].

California, Florida and Massachusetts have undertaken serious study of computer matching, but most States appear to have simply accepted the requirement as an activity imposed upon them by the Federal Government's fiscal power, and to have given little consideration to the programs' cost-benefit profiles or the information privacy of their data subjects.

A 1982 match between Massachusetts welfare files and data of 7 million account holders with 117 banks found 6,500 hits, and 1,600 persons were issued with termination notices. Of the 1,600 cases, 15% were found to have involved erroneous SSNs, and a variety of other errors and complexities were identified. Either 420, or less than 50, actual terminations resulted (depending on which report is to be believed) [New York Times, 13 December 1982, Reichman & Marx 1985, Laudon 1986 p.331, OTA 1986b, p.43, Flaherty 1989, p.354]. Again, "In a pilot matching program which compared all HHS employee records with welfare rolls of surrounding counties of Washington D.C., ... seventy-five percent of [the raw hits] were dismissed after further investigation, leaving 158 'solid hits'" [Laudon 1986, pp.332-3]. Marx and Reichman report a New York State program in which half of the hits were spurious due to timing problems alone [1984, p.435].

Criticism of exaggerated estimates of benefits is not confined to the press and disaffected academics. An unpublished consultant's report to the Office of Technology Assessment concluded that the States had found only low returns from computer matching, and that the primary reason for continuation was that the programs are mandated by the Federal Government [OTA, 1985]. Again in 1987, "a majority of states [67%] expressed the opinion that start-up and operating costs of a system to obtain and use federal tax data would likely exceed the benefits in terms of program dollars saved", citing case follow-up as the primary cost [GAO 1987, pp.3, 12, 83]. In another instance, "[The Internal Revenue Service] ... questioned [GAO's] recommended disclosure of tax data to [Veterans' Affairs] because the potential savings [GAO] cite [are] over-estimated and the cost of investigating income discrepancies and safeguarding tax information are not included in [GAO's] cost estimates. Further, IRS stated that [GAO's] cost-benefit analysis was based on generalized revenue potentials and failed to account for the possible effects of tax data disclosure on voluntary compliance with the tax laws and, ultimately, on tax revenues" [GAO 1988, p.45,93-94].

Laudon [1986, p.134-9] bemoaned the lack of research undertaken in this area, and even the lack of methods whereby such studies could be undertaken. GAO [1986b] concurred that "the methods that have been used for assessing the costs and benefits of matching projects were not well developed, well described or standardized" [1986b, p.12]. It proposed criteria for reviewing cost-benefit analyses [p.90]. OTA concluded that few U.S. Government programs are subjected to prior cost/benefit assessment, and that "as yet, no firm evidence is available to determine the costs and benefits of computer matching and to document claims made by OMB, the inspectors general and others that computer matching is cost-effective" [OTA 1986, pp.37, 50-53].

The General Accounting Office has undertaken many investigations into matching and privacy in specific contexts [GAO 1983a, 1983b, 1984a, 1984b, 1985a, 1985b, 1986a, 1987, 1988, 1989a, 1989b, 1990a, 1990b, 1990c, 1991a, 1991b, 1991c, 1993]. Reports emanating from the Human Resources Division (HRD; for example GAO 1988, 1990b, 1991a, 1991c) have been strongly supportive of computer matching programs, to the extent of committing many of the exaggerations and errors common in other, less careful agencies. On the other hand, reports emanating from the Program Evaluation and Methodology Division (PEMD; for example GAO 1986b, 1986c, 1993) have adopted a more circumspect attitude, and sought a balance between administrative efficiency and other interests. The hawks, who reflect the rationalist economic orientation dominant in governments throughout the world during the late 1980s, have generally seen their proposals come to fruition, whereas the doves' proposals for controls have not.

Following a review of cost/benefit analysis in relation to 17 programs chosen expressly because their entry in a matching inventory indicated some reporting of costs or benefits, GAO concluded that "The benefit of recovering overpayments and debt was often presented in terms of the maximum potential amount that might be collected but without an acknowledgement debt was often presented in terms of the maximum potential amount that might be collected but without an acknowledgement of or adjustment for money that might actually be recovered. With one or two exceptions, estimates of the overpayment-avoidance benefit were presented with little or no description of the computation or its rationale or underlying assumptions" [GAO 1986b, pp.13,79]. "In general, ... there was little documentation available on the development of a match from its initial conceptualization or on the decision to perform the match to actual implementation" [GAO 1986c, p.2]. At State level, cost estimates were "not well supported" and "unreliable indicators" [GAO 1991b, pp.3,27].

At no stage between 1974 and 1988 do the Privacy Act of 1974 or the OMB Guidelines appear to have provided any significant form of restraint on computer matching. There has been no mechanism whereby the Guidelines were enforced [Flaherty 1989, pp.349-350, 357]. This is certainly not to say that all of the programs undertaken were subject to shoddy justification, nor that all programs would have necessarily shown a negative net value. It is contended, however, that assessment free from the heavily politicised and manipulated environment which existed would have resulted in many programs being quickly abandoned, or not even commenced.

Bills to regulate computer matching were introduced into the Senate in 1986 and 1987 by Senator Cohen (Republican, ME). After many amendments, the Computer Matching and Privacy Protection Act 1988 (CMPPA) took effect in January 1990. It establishes a number of fair information practice provisions to apply to matching programs. However, for various reasons, the Act is not applicable to many categories of matching program. It affects only 'computerized comparisons' of 'automated systems of records' held by 'Federal or other Government agencies', and there are many classes of program which are explicitly exempted. See Table 1.

Table 1: Matching Programs Not Subject to the 1988 Act

- matches performed to produce **aggregate statistical data** ;
- matches performed to support **research**, and which are not used for administrative decision-making about individuals;
- matches performed by an agency which performs as its principal function any activity pertaining to the **enforcement of criminal laws** , in relation to previously identified individuals;
- a variety of matches of **tax information** ;
- matches using records predominantly relating to **Federal personnel** , which are performed for 'routine administrative purposes' provided that the purpose is not to take adverse action against Federal personnel;
- matches using **records from systems maintained by a single agency** , provided that the purpose is not to take adverse action against Federal personnel; and
- matches performed in relation to foreign **counterintelligence and security clearance** purposes.

The Data Integrity Board which each agency participating in a matching program must establish, is required to review all programs annually and assess their costs and benefits. In general they are not to approve programs until they have been subjected to cost-benefit analysis, and have been demonstrated to be likely to be cost-effective. However, this requirement may be waived by the (internal) Data Integrity Board, in accordance with guidelines to be prescribed by the Director of OMB. Moreover, disapprovals may be overruled by the Director of OMB if he determines the program to be legal, cost-effective and in the public interest.

In all of the exempt classes, CMPPA created no new privacy protections: exemption frees agencies from any form of control, including even disclosure of the program's existence. Some of the exemptions are of broad scope, particularly that relating to tax information. The exclusion of matches within a single agency overlooks the existence of multi-function agencies, which are thereby permitted unfettered license to match data which is under the control of their various arms. In addition, the Act can be circumvented by moving the administration of a program into the same agency as the data system against which the data is to be compared. A subsequent study confirmed that "a significant proportion of governmentwide matching activity is excluded from the Act" [GAO 1990c, p.28]; for example, over half of the matching programs undertaken by the Social Security Agency are exempt [Clarke 1992a].

Even in respect of non-exempt matching programs, the only responsibilities the Act imposes are on Federal Government agencies, and they are only imposed in relation to matching programs in which one of the parties is a Federal Government agency and the other(s) is(are) Federal agencies or non-Federal (i.e. State or local government) agencies. The obligations therefore do not apply in the case of matching programs in which one Federal agency is dealing with one or more non-government organisations, and hence the law may be capable of subversion through the inclusion of an exempt organisation in a parallel matching program. Hence the requirements would not apply to, for example, a program involving the acquisition of any data by a Federal agency from a credit bureau, a financial institution or a network operator which provided EFT/POS services.

The Annual Reports required under the Act have been very slow in appearing, and very limited in scope. The primary conclusion of the first report (for the first part-year to March 1990) was that agencies had underestimated the impact of the verification and notice provisions. In due course a Congressman was found to sponsor an amendment Bill, and H.R. 5450 of 1990 significantly weakened the verification and notice provisions. The second Annual Report, for the remainder of 1990, 1990 significantly weakened the verification and notice provisions. The second Annual Report, for the remainder of 1990, took over two years to appear [OMB 1992], and, despite expectations, provided no usable information about the use of CBA. By the end of 1992, over 4 years after the CMPPA was passed, OMB had still not promulgated cost/benefit analysis guidelines, and regarded it as being of continuing low priority. GAO's proposed model [GAO 1986b] did not appeal to agencies.

In October 1993, an assessment by GAO concluded that there were serious deficiencies in the implementation of the CMPPA. In 40% of cases, agencies failed to even perform a meaningful CBA, and even where one was performed the quality was very low. The agencies' internal Data Integrity Boards accepted the analyses despite their severe methodological flaws, and served merely as a means of legitimising the programs [GAO 1993].

It is concluded that, in the United States:

- cost/benefit analysis has been applied to very few programs;
- the quality of such cost/benefit analysis as has been undertaken has generally been of low or exceedingly low quality;
- many programs which have been and are being undertaken might well have been judged by managers and representatives of the public to be of net negative worth if adequate information had been publicly available; and
- Congress' expectations that following the 1988 Act appropriate-quality cost/benefit analysis would be applied to at least some matching programs has been frustrated.

8. Review of CBA and Computer Matching in Australia

Computer matching has been intensively used in Australia since the late 1970s. There was discussion of regulation of computer matching in several parts of the Law Reform Commission's Report on Privacy [ALRC 1983], but no concrete conclusions were reached. It appears that, until the early 1990s, no Australian Government agency with regulatory responsibility even attempted to establish guidelines regarding the conduct of computer matching.

In this climate, very little information became publicly available concerning the justification of computer matching programs. In 1987-88, Freedom of Information requests to, and discussions with, a variety of Commonwealth Government agencies elicited the following information:

- despite the Freedom of Information Act's instructions to government agencies to be more open in their dealings with the public, many retained a strong preference for undertaking their work without publicity;
- all requests under the Freedom of Information Act for information concerning computer matching policies, programmes and guidelines were refused, variously because:
    - the relevant documents were exempt because they were 'documents affecting enforcement of law'; and/or
    - no document could be identified which contained the desired information.

Meanwhile a report by a committee representing government agencies, on a Review of Systems for Dealing with Fraud on the Commonwealth [Fraud 1987] firmly recommended "that the matching of data is an effective means of preventing and detecting fraud, and that its wider use should be considered and publicised" [p.5]. It was very supportive of the Australia Card proposal as a means of improving the reliability of computer matching [pp.100-102. See also Clarke 1987]. There have been other reports within the Commonwealth public sector which have supported or urged the increased use of matching, e.g. ANAO [1990], but these have seldom contained discussion of the protections necessary as a concomitant to the powers.

The Privacy Act of 1988 created a permanent Privacy Commissioner, and required him to, among many other things, 'research into and monitor developments in ... data-matching' [s.27(1)(c)]. Related provisions provided him with the function of examining a proposal for data-matching on request by a Minister or agency (k), and the power to do all things necessary or convenient to be done for or in connection with the performance of these functions (s.27(2)). The Privacy Commissioner also has computer matching responsibilities in relation to the Tax File Number (s.28 and Schedule 2). He is explicitly required to have regard to 'the protection of important human rights and social interests that compete with privacy' (s.29): privacy concerns about computer matching are not to be considered in isolation, and the benefits must be taken into account in assessing computer matching practices, and in preparing guidelines.

In 1990, the Privacy Commissioner discussed matching programs with five agencies, and concluded only that "whether cost/benefit analyses have formed part of the original decision to commence a matching program has been difficult to ascertain in many circumstances ... A definitive cost for distinct programs is often difficult to obtain" [PCA 1990, p.16]. Similar criticism was levelled by two regulatory agencies at the financial justification offered for a proposal by the Health Insurance Commission to significantly extend the data-intensity of its pharmaceutical benefits operations [ANAO 1991]. If this were typical of the level of competence with which computer matching programs were routinely justified, it would be no surprise that agencies were unenthusiastic about submitting their justifications to public scrutiny.

In late 1990, the Privacy Commissioner published an Exposure Draft of a set of Guidelines relating to data-matching, with the intention of communicating the his intended interpretation of his responsibilities and powers under the Act, and the standards intention of communicating the his intended interpretation of his responsibilities and powers under the Act, and the standards which he was considering applying in their execution. Agencies sought and gained lengthy extensions to the submission date; their submissions are not publicly available.

The Privacy Commissioner issued revised draft guidelines in September 1991 [PCA 1991b]. These contained major weakenings in the positions adopted in the original draft. A set of Final Proposed Guidelines was completed in December 1991 (although the document's existence and the procedure adopted were not publicly known until mid-1992). These corrected some of the weaknesses which the previous draft had created [PCA 1992]. They were sent to agencies with a request that they voluntarily adopt the Guidelines for a period of 18 months until August 1993. Very few agencies accepted the invitation to adopt the Guidelines voluntarily, and the Commissioner's hope that the delay in implementation would result in valuable experience being gained was forlorn.

The Guideline's requirements regarding prior cost-benefit analysis are diffuse (PCA 1992, pp.9, 16 and 20), and it is even arguable that such analysis is unnecessary ("If the program is being justified according to cost/benefit ...", p.16). Moreover, low standards of analysis apply. In particular:

- the document makes no reference to authoritative references on cost/benefit analysis;
- it refers to 'resource expenditure' and 'quantifiable savings', and hence appears to omit those factors which are quantifiable only in other-than-financial terms and those which are qualitative in nature;
- it fails to make clear that costs of and benefits to all agencies, other organisations and individuals should be included;
- it fails to mention opportunity costs;
- it fails to require that the basis of derivation be shown, that underlying assumptions be stated, and that 'guesstimates' be distinguished from empirically derived data; and
- it is unclear in its requirement relating to reference to experience, and especially to the post-evaluation assessments of prior programs (pp.13, 14). It appears that it is not mandatory to include a cost-benefit analysis in the evaluation, that there are no requirements that the post-match analysis facilitate comparison with the pre-match analysis, and that there is no clear requirement to compare outcomes with expectations. Long-running programs only need to be evaluated every three years, and (presumably by accident) the requirement for evaluation of a short-running program is "not later than three years after the commencement of operation".

There are also many exemptions from the Guidelines. Agencies and programs which are exempt from the Freedom of Information Act are also exempt from the Privacy Act and hence from the Computer Matching Guidelines. Most programs "the objective of which is to assist in undertaking law enforcement and fraud control activities ..." are excluded (PCA 1992, p. 5). This is so broadly phrased that a case could be made for virtually all uses of matching in relation to taxation and support payments to be exempt from the Guidelines. Only "automated data-matching programs" are covered, and only those "involving the scrutiny of the records of more than 5,000 individuals" (p.5). The size limitation appeared only in the 'final proposed version', and without justification, but is presumably because the cost of controls would appear to be proportionally relatively high in such cases. The exemption process represents a serious shortcoming in the protective regime, because rather than some aspects of the controls being suspended or qualified, the entire set of controls is discarded without any form of replacement. There is not even any suggestion that exempt programs be subjected to self-regulation with the Guidelines as a primary source.

The Privacy Commissioner conducted a further survey of computer matching programs during the fourth quarter of 1993, but the report has not yet been published. A draft review of the operation of the voluntary guidelines [PCA 1993] did not include any comments whatsoever on the cost-benefit analysis aspects, despite submissions by privacy interest groups to the effect that they were inadequate. The majority of the refinements mooted in the document appeared to be designed to satisfy the requests of agencies, by reducing the controls and increasing the exemptions.
It is apparent that CBA has at no stage operated in any effective manner to control computer matching in Australia. The creation of the Privacy Commissioner appears not to have made any difference in this regard to date, and the prospects of it making any difference appear slim.

The discussion in this section has intentionally omitted reference to a major computer matching program which was authorised by a special Act of Parliament in 1990, and is subject to a statutory control regime. The next section reports on a longitudinal case study of the pre-match justification and official post-match analyses of that program.

9. A Case Study of CBA in Australia

(a) Background

Between 1985 and 1987, the Australian Government developed a proposal for a national identification scheme based on an 'Australia Card'. In the face of overwhelming public opposition, it eventually withdrew the Bill (Clarke 1987). As an alternative, it first tightened the requirements relating to the Tax File Number, then, in violation of promises made in Parliament and widely publicised, quickly extended the use of the TFN to the entire welfare sector (Clarke 1991a, 1992a).

Parliament and widely publicised, quickly extended the use of the TFN to the entire welfare sector (Clarke 1991a, 1992a).

In the Budget of August 1990, the Minister for Social Security announced that the TFN was to be used also as the basis for what was referred to as the 'Parallel Data Matching Program' (PDMP). The program involves data from five major 'source agencies' (the Australian Taxation Office - ATO, the then Department of Community Services and Health - DCS&H, and the Departments of Social Security - DSS, Employment, Education & Training - DEET and Veterans' Affairs - DVA), and two other 'assistance agencies' (the Australian Electoral Commission - AEC, and the Health Insurance Commission - HIC, which runs the country's universal Medicare scheme).

Data is provided by source agencies to the Data Matching Agency within DSS. DSS checks it for validity. It extracts the TFNs and provides them to the ATO. The ATO returns to DSS the identity and income details associated with each TFN. DSS compares the data from the five source agencies and two assistance agencies, with the aim of finding identification discrepancies and anomalies in the patterns of benefit payments and declared income. DSS provides information about discrepancies and anomalies back to the source agencies.

This 'cycle' is repeated between 5 and 9 times each year. Each cycle involves over 10 million attempted matches, producing about 200,000 apparent discrepancies and anomalies, filtered down to about 15,000 for review. About 10,000 of these survive a preliminary manual examination, and result in enquiries being directed by one or more of the source agencies to the client. Action results in about 1,500 cases, such as variation or cancellation of a benefit, and/or action to recover over-payments of a benefit or to collect under-payment of tax (indicative figures only).

(b) Pre-Match CBA

Initially, the justification for the scheme was not released by the Government. Following personal representations to the Minister for Social Security, however, a copy of the documents which supported the Budget Submission were released [DSS, 1990b]. It appears to have been the first justification of an Australian computer matching scheme ever to become public.

The claimed benefits were $65 million in the first 6 months, and about $300 million p.a. thereafter. The document was short, and the only detailed analysis related to one specific benefit: family allowance. A number of aspects of the estimating basis are noteworthy [pp.4-7]:

- the proportion of family allowance recipients who receive an entitlement they should not, or who are overpaid, was estimated by random sampling techniques to be 2.3%;
- during the 1989-90 financial year, DSS undertook 232,000 reviews of family allowance recipients, based on targeting of "cases where it is believed there is a higher risk of incorrect payment". This resulted in 30,400 cancellations;
- these targeted reviews will continue; and
- the estimated gross savings in relation to family allowance payments, as a result of the new matching programs, was based on an assumption that 2% (38,000 of 1.9 million allowances) would be cancelled.

Four distinct problems with the justification are apparent. The first is that the estimates involved the double-counting of 30,400 cancellations which were to be discovered and cancelled because of detection by both control techniques. This was the only one of the eight benefits schemes for which this amount of detail was disclosed. The assumptions did, however, appear to be common to all schemes: "The Decision assumes a [total] Social Security income-support population of just over 4 million people with a randomly distributed incorrect payment rate of between 1 and 2 per cent ..." [p.5]. If the degree of inaccuracy was common across all schemes, the potential gains from the program as a whole would have been not $65 million in the first 6 months and $300 million p.a., but $13 million and $60 million respectively.

The second problem is that the estimates were based on the implicit assumption that 87% (2.0%/2.3%) of all detected overpayments would result in cancellation, an extremely high proportion not encountered in the limited United States evidence available, especially since delay factors exist, such as the right to dispute the Department's proposed ruling.

The third problem is the implicit assumption that the savings are repeatable, i.e. that every year 43,700 new errors and abuses would arise, of which 38,000 could be found and eradicated. Given that this population is probably one of the more stable benefit recipient groups, it seems highly unlikely that a large enough proportion of new applicants would abuse the scheme. If, on the other hand, family allowance cheats have been found by the Department to be persistent and early recidivists, then personal dataveillance techniques would be both more efficient and more readily justifiable than mass dataveillance techniques like computer matching.

The fourth problem is that the figures related only to gross savings: no allowance was made for one-time, recurrent or opportunity costs in the Department or elsewhere. At no point was any information provided as to what those costs might have been, nor even whether estimates had been prepared; nor were they extractable from the high level of aggregation in the Budget papers.

On the basis of the Department's own figures, the potential net financial benefits of the proposal dwindled away to a very small number. Yet this was sufficient for the proposal to gain the support of the Cabinet, which in effect means the support small number. Yet this was sufficient for the proposal to gain the support of the Cabinet, which in effect means the support of the various agencies which were to be involved in the scheme, have a checking role, or exercise considerable power in the pre-Budget rounds of discussions. During the Parliamentary stages, the arguments of privacy interest groups were ignored, and the Bill passed both Houses. The pre-match CBA exercise failed to operate as a constraint on a highly privacy-invasive scheme whose justification was seriously flawed.

(c) Post-Match Cost/Benefit Analysis - 1991-92

Part of the price which the Government and the agency had paid for the over-ambitious Australia Card scheme was the creation of a Privacy Commissioner who politicians and senior public service executives felt constrained to involve in discussions. As a result, the enabling legislation included a set of controls over the operation of the scheme, including the requirement that an Annual Report be prepared and tabled in Parliament, and a sunset clause [PC 1991a, Kelly 1992]. This has resulted in the PDMP being probably the most publicly documented computer matching scheme ever. The remainder of this section addresses two matters: what does post-match analysis tell about the quality of the original (1990) CBA; and what was the quality of the cost/benefit analyses undertaken in 1992 and 1993 to justify continuation of the scheme? It is based on the Department's Annual Reports on the Program [DSS 1991, 1992, 1993], together with a report by the Australian National Audit Office [ANAO 1993].

The scheme fared far less well than the Department had predicted. In 1991-92, the gross savings from cancellations and downward variations combined were $17 million rather than the $290 million from cancellations alone which had been confidently predicted 18 months earlier. A subsequent audit report provides further details: whereas $300 million in savings had been anticipated from 70,000 cancellations, $17 million had been realised from only 3,682 cancellations. And instead of the expected additional 40,000 cases of downward variations in payments, only 5,613 had resulted (ANAO 1993, p.xi, p.17). A variety of reasons were nominated for this huge shortfall (DSS 1992, pp.43-45, 51, 73-75; see also Clarke 1993b, p.4-5).

Some of the considerations were reasonable justification for deferral of the expected benefits, but others threw serious doubt on the quality of the original assessment. The Department was clearly far from confident that the setbacks were all temporary: as reality set in, the estimates were reduced to about one-third of the original claim (see Table 2) [DSS 1991, p.105-6; 1992, p.viii, 106; 1993, p.118 - note that due to the incomplete manner in which the agency reported, these figures relate only to the gross benefits arising from cancellations and downward variations].

Table 2: Gross Benefits As Estimated by DSS - 1990, 1992

	90-91	91-92	92-93	93-94	94-95	95-96
Projections - October 1990	65.0	>290.0	>300.0	>300.0	>300.0	>300.0
Projections - October 1992	-	-	107.3	114.1	81.2	99.1
DSS's 'Actual' Outcomes 0.0	17.0	-	-	-	-	

This example of wishful thinking attracted very little opprobrium. A subsequent Audit Report limited its criticism to "savings from the program have not been as significant as initially predicted" (ANAO 1993, p.viii). Parliament took no action on the matter. Perhaps such inaccuracy is an accepted part of the game of getting programs approved by the Cabinet and the Parliament.

The over-enthusiasm of the Department for the program is of historical interest. Of ongoing concern, however, was the Department's failure to apply conventional cost/benefit analysis principles to the exercise. Indeed, there was evidence of failure to even understand the concepts involved. In the 1992 Report, for example, net present value techniques were not applied, hardware and maintenance costs were overlooked, no costs were imputed for the efforts of other agencies and clients (which in the case of a program of such wide scope is essential), and the bases on which savings were projected into the future were not stated. The most glaring error was the complete omission of the staff costs involved in 137,000 manual examinations of files, 18,000 actual reviews, 10,000 actions against clients, 1,300 queries by clients, 150 formal appeals, 1,500 debt recovery actions (of which 700 involved negotiations with the debtor), and 100 briefings of the Director of Public Prosecutions. This omission was despite statements that "the real cost has been in the time and effort of staff administering the program" (DSS 1992, p.13) and "the reporting requirements are stringent and a lot of time and effort is needed to comply with them" (p.12).

The Privacy Commissioner expressed similar concerns, albeit more gently (private communication, 29 July 1993). An external audit of the Parallel Data Matching Program also criticised the quality of cost/benefit analysis undertaken, and pointed out that the Act "requires the tabling of a comprehensive report in both Houses of Parliament ... Sufficiently comprehensive cost/benefit information had not been included in either Report ..." (ANAO 1993, p.4). It appears that the quality was so low that the DSS may have been in breach of the Act.

The DSS's 1993 Progress Report acknowledged that "The Privacy Commissioner and some other commentators suggested [the

Department's previous reports] lacked sufficient accounting rigour and a more comprehensive cost-benefit analysis covering opportunity costs and costs absorbed by the Department should be conducted. This view was also shared by the Australian National Audit Office. The Department therefore decided to undertake such an analysis" (DSS 1993, p.79).

(d) Post-Match Cost-Benefit Analysis - 1993

In its third Annual Report, the Department stated that in 1992/93, the PDMP had achieved benefits of $70 million against costs of $13 million, for a net benefit of $56 million. In subsequent years, the Department projected close to double the benefits, with unchanged costs, and imputed additional benefits from voluntary compliance ranging from $40 to $90 million per annum. The net benefits from 1993/94 onwards were accordingly shown as $110-115 million per annum (DSS 1993, pp.79-91, 93-97, 115-129).

The information provided in the Department's Report of October 1993 was an improvement on the parlous attempts of previous years. The improvements were:

- recognition that opportunity costs are relevant;
- recognition that operational costs are relevant; and
- recognition that overpayment recovery is not automatic, by reducing the recovery figure from the previous (implicit) 100% to 70%. The leakages from full recovery include debts waived (debts of less than $200 appear to be commonly waived); death of the debtor (which debts are believed not to be pursued); emigration; inability to locate the debtor; discounts allowed as part of agreements reached with debtors; and defaults on agreements to pay.

Unfortunately the cost/benefit analysis contained in the DSS's 1993 Progress Report still contained significant errors. These were:

- the attribution to the data matching program of savings which resulted from the requirement for beneficiaries to provide their TFN. This included:
    - the cancellation of benefits to individuals who failed to provide a TFN;
    - the voluntary surrender of benefits when the TFN was requested; and
    - the denial of new benefits when a TFN was not supplied.
- These savings had almost nothing to do with the data matching program, and would have arisen whether or not the data matching program had ever existed (the only justification for including them in the Reports is that the action was authorised by the same package of statutes as authorised the matching program);
- the imputation of gross levels of deterrent effect. There are continuing attempts by the Department to claim inflated gains from unresearched deterrent effects which, if they exist, result from a whole raft of factors;
- the failure to provide the basis upon which opportunity costs were calculated, despite clear requirements that they be disclosed;
- the use of unreasonably low opportunity cost figures. A major scheme was abandoned in order to divert the resources to processing hits from the matching program. During the first 2-1/2 years, the salaries diverted to the program totalled $25 million, yet zero net opportunity costs were shown. The figures provided by the Department imply that, if the program had not been commenced, these people's salaries would have been paid in return for absolutely no benefits;
- there is no evidence that the costs of achieving overpayment recovery have been included. These would be expected to be significant, even for fairly straightforward cases, and very substantial where the case was contested, or the benefit-recipient elusive;
- the delays involved in actually achieving restitution of overpayments from benefit-recipients may not have been taken into account. The Progress Report fails to disclose the basis upon which the tables of figures were prepared. A fairly conservative presumption might be that repayments would (on average) commence in the middle of the year in which they are discovered, and that they would (on average) be spread evenly over 3 years. Yet the agency assumes simply that 70% of all debts outstanding at the end of one year will be collected during the next;
- the misapplication of the discounting rate and formulae. In applying the discount factor to allow for the effect of delays in receiving benefits, the Department made three separate errors.

It is also noteworthy that, on the basis of their own analyses, the other agencies involved in the scheme have not reached breakeven point, and in at least some cases appear not to anticipate any benefits.

(e) Conclusions

Table 3 shows the revised benefits projections provided by the agency, which progressively plummeted to 10-13% of the original estimates. The criticisms made in the previous section give rise to adjustments, and the final line of the table shows the adjusted figures.

Table 3: Gross Benefits As Estimated by DSS and the Author - 1990, 1992, 1993

	90-91	91-92	92-93	93-94	94-95	95-96

	90-91	91-92	92-93	93-94	94-95	95-96
Projections - October 1990	65.0	>290.0	>300.0	>300.0	>300.0	>300.0
Projections - October 1992	-	-	107.3	114.1	81.2	99.1
Projections - October 1993	-	-	-	35.8	39.8	40.2
DSS's 'Actual' Outcomes 0.0	15.0	30.8	-	-	-	
Author's Estimates	0.0	11.2	26.9	27.2	29.7	30.0

Table 4 shows the net benefits claimed by DSS. In addition, adjusted figures are shown after making an allowance for exaggeration of benefits and more reasonable levels of opportunity cost. The latter are based on the amount suggested by ANAO (1993, p.6). Finally, the cumulative net value in 1990/91 dollars is shown, after allowing for the time-value of money and using the recommended discount rate of 8% p.a [DoF 1991, p.57].

Table 4: Net Benefits As Estimated by DSS and the Author - 1993

	90-91	91-92	92-93	93-94	94-95	95-96
DSS's Estimates						
Benefits	-	15.0	30.8	35.8	39.8	40.2
Costs	8.0	13.3	13.7	14.7	13.8	13.3
Net Benefits	-8.0	1.7	17.1	21.1	26.0	26.9
DSS's Estimates Adjusted						
Benefits	-	11.2	26.9	27.2	29.7	30.0
Costs	8.0	25.7	25.3	23.8	22.1	22.1
Net Benefits	-8.0	-14.5	1.6	3.4	7.6	7.9
Discount Factor	1.0	0.926	0.857	0.794	0.735	0.681
Net Value	-8.0	-13.4	1.4	2.7	5.6	5.4
Cumulative Net Value	-8.0	-21.4	-20.0	-17.3	-11.7	-6.3

This case study shows that the original benefits estimates, which inveigled the Government and the Parliament into approving the program, were enormously exaggerated. The agency has, moreover, continued to abuse the technique of cost/benefit analysis in order to provide the appearance of a positive outcome, with the desired result that the Parliament has on two occasions authorised continuation of the program.

10. Evaluation of CBA As a Control Mechanism

Appropriately applied, cost/benefit analysis would be expected to prevent schemes from being implemented, unless they were at least financially viable. Moreover, they would be expected to cause schemes to be abandoned which were implemented for political or strategic reasons (i.e. despite inadequate economic justification), and which failed to fulfil their anticipated potential. The research reported on in this paper, comprising analysis of secondary sources, field work, and an in-depth, longitudinal case study, provide evidence of schemes being implemented with scant attention to economic factors, let alone to broader social considerations.

Although cost/benefit analysis is a well-documented method, agencies in the United States and Australia have a generally poor record in its application to computer matching. Some of the abuses of the cost/benefit analysis technique which have been and continue to be committed include:

- implicitly assuming that all losses from error and fraud are fully avoidable if computer matching is implemented, without allowing for the many reasons why savings may not eventuate, including:
    - poor quality data giving rise to a spurious 'hit';
    - poor quality data that would not secure a conviction, and perhaps not even enable the agency to defend against an appeal against an administrative determination;
    - cases which would give rise to only small recoveries, and which are therefore not worth pursuing;
    - cases where the agency exercises a prerogative not to pursue the amount, for humanitarian or other policy reasons;
    - 'skips', who can no longer be located; and
    - people who are no longer alive;
- failure to apportion benefits among the various programs which give rise to them, and hence double-counting of benefits;
- omission of the costs of investigation, prosecution and collection;
- omission of the costs of other agencies (including government solicitors), companies and individuals; and
- omission of opportunity costs.

There is no doubt that computer matching can be instrumental in the detection of a proportion of the errors, abuse and fraud in a variety of programs. There remain, however, considerable doubts about whether, in respect of any particular program, net financial benefits will arise, after allowing for all quantifiable and qualitative costs. Unless credible cost/benefit analyses are undertaken, both retrospectively and in advance, the potential economic safeguard against excessive use of computer matching cannot be realised.

Even if they were effective, economic controls 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 IT did not result in excessive privacy invasion [Westin 1967, Westin 1971 and Westin & Baker 1974]. This view was seriously undermined by Rule's work, which has demonstrated that, rather than supporting individual freedoms, administrative efficiency conflicts with it. Organisations 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 [Rule 1974, Rule et al 1980].

11. A Normative Regulatory Regime

The negative impacts of computer matching are sufficiently serious to demand controls. Natural or intrinsic controls are inadequate. In particular, economic constraints have proven ineffectual. A regulatory regime is necessary, to ensure that computer matching is undertaken only in circumstances, and in ways, which are beneficial to society.

The framework proposed by this author is normative because, in the absence of 'good' models, it is impractical for it to be derived from experience. The general principles proposed here are developed in part from previous proposals (and in particular Laudon [1986, pp.384-90] and Flaherty [1989, pp.359-407]), and in part from the author's own work. The framework comprises a few general principles, plus a set of detailed requirements of a legally enforceable code of practice. It is expressed in much greater detail in Clarke (1994c).

The establishment of extrinsic controls over computer matching is very unlikely to be even embarked upon until comprehensive information privacy laws are in place. Unfortunately, the primary consideration in the formulation of privacy laws throughout the world has been that the efficiency of business and government should not be hindered. What has been provided is an 'official response' which has legitimated dataveillance measures in return for some limited procedural protections commonly referred to as 'fair information practices' [Rule 1974, Rule et al 1980]. The first requirement of a control regime for computer matching is comprehensive and universally applicable data protection legislation which achieves a suitable balance between the various economic and social interests, rather than subordinates information privacy to administrative efficiency.

Privacy protection regimes based on cases being brought by private citizens against the might of large, information-rich and worldly-wise agencies have not worked, and are highly unlikely to do so in the future. To achieve appropriate balance between information privacy and administrative efficiency, it is necessary for an organisation to exist which has sufficient powers and resources to effectively represent the information privacy interest [ALRC 1983; OTA 1986b, pp.57-59, 113-122; Flaherty 1989, esp. pp.359-407]. It would be valuable to complement such a body with an effective option for individuals to prosecute and sue organisations which fail to comply with legal requirements. This can only come about if the requirements of organisations are made explicit, and this in turn is only likely to come about if detailed codes, prepared by a privacy protection agency on the basis of research and experience, are given statutory authority.

Computer matching is a specific technique, and cannot be appropriately enabled and controlled by generalised legislation [OTA 1986b, pp.57-59]. Either Parliaments must expend the effort to become competent in dealing with the issues, or they must create a specialist agency and invest it with the responsibility of bringing specific statutory instruments before it, dealing with specific programs or specific agencies. As Laudon concluded, "a second generation of privacy legislation is required" [1986, p.400]. The challenge is to regulate computer matching in such a way that it clears the path for worthwhile applications of the technique, while preventing unjustifiable information privacy intrusions.

When it suits their interests, agencies adopt the attitude that the agencies of government are merely branches within a single administrative apparatus, and hence all data transfers are internal to government. Monolithic government is inimical not only to the privacy interest, but to civil liberties generally. It is necessary for Parliaments to make clear that agencies are independent organisations for the purposes of data transfer, and that all data transfers are therefore subject to the rules regarding collection and dissemination. In addition, there is a danger that privacy protections may be subverted by the concentration of functions and their associated data into large, multi-function agencies. Hence boundaries must be drawn not only between agencies but also between functions and systems.

Care must be taken to ensure that exemptions do not rob privacy protection legislation of its effectiveness. The general principles of information privacy must be applied to all agencies and all systems, and the regulatory regime for computer matching to all programs. The widely practised arrangement of exempting whole classes must therefore be rolled back. It is entirely reasonable, on the other hand, for the specific nature of controls to reflect the particular features of an organisation entirely reasonable, on the other hand, for the specific nature of controls to reflect the particular features of an organisation, system or program.

As Laudon noted, "a pattern has emerged among executive agencies in which the identification of a social problem [such as tax evasion, welfare fraud, illegal immigration, or child maintenance default] provides a blanket rationale for a new system without serious consideration of the need for a system" [1986, p.385]. This 'blanket rationale' must be swept away. Computer matching programs must be subjected to conventional cost/benefit analysis, including estimation of the full range of actual and opportunity costs and financial benefits, quantification of as many as possible of the non-financial costs and benefits, and description of the non-quantifiable factors. Guides exist as to the nature of costs and benefits which should be considered [Kusserow 1984b, pp.33-40, PCC 1989, PCA 1990, SSA 1990, GAO 1990a, Clarke 1994c]. The justification of each program needs to be reviewed by an organisation whose interests are at least independent of those of the proponent organisation, and perhaps even adversarial.

Technological developments have rendered some of the early information privacy protections ineffective: "new applications of personal information have undermined the goal of the Privacy Act that individuals be able to control information about themselves" [OTA 1986b, p.4. See also Thom & Thorne 1983 and Greenleaf & Clarke 1984]. If a proper balance between information privacy and other interests is to be maintained, processes must be instituted whereby technological change is monitored, and appropriate modifications to law, policy and practice brought about. This needs to be specified as a function of the privacy protection agency, and the agency funded accordingly.

In addition to these general principles, detailed requirements must be imposed on organisations involved in computer matching. Those proposed in Clarke (1994c) comprise:

- pre-conditions to the commencement of a program, including legal authority, consideration of alternative measures, analysis and documentation of data quality, socio-economic justification through application of prescribed cost/benefit analysis techniques, public knowledge about the program, and the preparation and examination by the privacy protection agency of a set of terms of reference for the program dealing with a prescribed list of considerations;
- requirements regarding the conduct of the program, including controls over 'data scrubbing' (the massaging of data format and content), the matching process, and the inferencing and filtering steps;
- requirements regarding the use of the results, respect for due process in the taking of action arising from the results, the retention and destruction of the data arising from the program, post-program assessment of data quality, and re-evaluation of the program's costs and benefits including comparison against the cost/benefit analysis used to justify it in the first place; and
- general requirements, including overview by the privacy protection agency, public visibility of the program, consistency of program performance with the terms of reference, security and confidentiality, staff training, documentation, subject access to personal data, and auditability.

12. Conclusions

The research reported on in this paper unearthed very limited evidence to support the contention that cost-benefit analysis acts as a control over unreasonable or excessive use of computer matching by government agencies in the United States or Australia. Indeed, it appears to rarely, if ever, cause a proposed program to be not proceeded with; and seldom causes an existing program to be abandoned.

CBA is expensive, and not well understood. Government programs are commonly launched as strategic measures for which CBA is at best a justificatory exercise and at worst an unnecessary hindrance. Powerful agencies have abused CBA, and manipulated regulatory agencies and Parliaments to relieve themselves of the onerous duty of performing CBA.

The public's interest in unjustifiably privacy-invasive computer matching programs not being undertaken appears not to be capable of protection by intrinsic economic controls. In this arena at least, the tension between the State's interest in social control and individual citizens' interests in freedom from unreasonable interference is being consistently resolved in favour of the State.

To some people, that may be an entirely satisfactory outcome, but it flies in the face of democratic ideals. If the situation is to change, citizens' representatives need to impose their will far more convincingly. This in turn implies the need for far more active lobbying by public interest groups for an effective regulatory regime. Critical among the features of such a regime is a requirement that CBA be undertaken prior to and subsequently to the conduct of programs.

These findings have implications far beyond computer matching. Use of existing dataveillance techniques is increasing, and new waves of privacy-invasive technologies and techniques are arriving. Tracking of trade transactions, of funds flows, of vehicular movement and of message traffic may be nominally of entities other than people, but in practice dramatically enhance the surveillance of individuals. It can be confidently anticipated that proponents of these techniques will claim that unreasonable uses of these techniques are subject to economic control mechanisms, and that abuses will therefore not occur.

On the basis of the research reported on in this paper, serious doubt is cast over such claims.

ACKNOWLEDGEMENTS

My thanks to Bill Orme, then Executive Member of the N.S.W. Privacy Committee, who asked me to write a paper on this topic in 1976, when the term was only just coming into use and basic privacy protections seemed a higher priority. Sorry it's late, Bill. The assistance is gratefully acknowledged of fellow dataveillance researchers in Australia, the United States and Canada. So also is that of the Australian Privacy Commissioner, Kevin O'Connor, and his deputy Nigel Waters.

Staff of several U.S. government agencies were very forthcoming, and, since 1989, the Australian Department of Social Security and National Audit Office have also provided information valuable to the research. Their openness has enabled, for the first time, some amount of light to be thrown upon the practice and economics of computer matching by government agencies. Finally, reviewers and editors have contributed significantly to the tightening of the paper's focus. Responsibility for the interpretative comment in the paper is mine alone.

References

ALRC (1983) 'Privacy' Aust. L. Reform Comm., Sydney, Report No. 22 (1983)

ANAO (1990) Report No. 24, Australian National Audit Office, Canberra, 1990

ANAO (1991) 'Pharmaceutical Benefits Scheme: Review of Estimated Savings from Proposed System for Eligibility Checking' Joint Review by The Auditor-General and the Department of Finance, Canberra, Dec 1991

ANAO (1993) 'Audit Report No. 7 of 1993-94: Efficiency Audit: Department of Social Security - Data Matching' Australian National Audit Office, Canberra (October 1993)

APSB (1981) 'A Guide to Cost-Effectiveness Analysis of ADP Systems' Australian Public Service Board, Aust. Go vt Publ. Serv., Canberra, 1981

Azrael M.L. (1984) 'Lost Privacy in the Computer Age' The Law Forum, University of Baltimore School of Law (Spring 1984) 18-26

Bennett C. (1992) 'Regulating Privacy: Data Protection and Public Policy in Europe and the United States' Cornell University Press, New York, 1992

Berman J. & Goldman J. (1989) 'A Federal Right of Information Privacy: The Need for Reform' Benton Foundation Project on Communications & Information Policy Options, 1776 K Street NW, Washington DC 20006, 1989

Clarke R.A. (1987) 'Just Another Piece of Plastic for Your Wallet: The Australia Card' Prometheus 5,1 June 1987. Republished in Computers & Society 18,1 (January 1988), with an Addendum in Computers & Society 18,3 (July 1988)

Clarke R.A. (1988) 'Information Technology and Dataveillance' Commun. ACM 31,5 (May 1988). Re-published in C. Dunlop and R. Kling (Eds.) 'Controversies in Computing', Academic Press, 1991

Clarke R.A. (1991a) 'The Tax File Number Scheme: A Case Study of Political Assurances and Function Creep' Policy 7,4 (Summer 1991)

Clarke R.A. (1991b) 'Computer Matching Case Report: Rental Assistance for Low-Income Families' Working Paper available from the author (December 1991)

Clarke R.A. (1992a) 'The Resistible Rise of the Australian National Personal Data System' Software L. J. 5,1 (January 1992)

Clarke R.A. (1992b) 'Computer Matching in the Social Security Administration' Working Paper available from the author (January 1992a)

Clarke R.A. (1992c) 'Computer Matching and the Office of Management and Budget' Working Paper available from the author (January 1992b)

Clarke R.A. (1993a) 'Dataveillance by Governments: The Technique of Computer Matching' Working Paper, Dept. of Commerce, Australian National Uni. (July 1993)

Clarke R.A. (1993b) 'Matches Played Under Rafferty's Rules: The Parallel Data Matching Program Is Not Only Privacy-Invasive But Economically Unjustifiable As Well' Working Paper, Dept. of Commerce, Australian National Uni. (November 1993)

Clarke R.A. (1994a) 'Matches Played Under Rafferty's Rules: The Parallel Data Matching Program Is Not Only Privacy-Invasive But Economically Unjustifiable As Well' Policy 10,1 (Autumn 1994)

Clarke R.A. (1994b) 'Dataveillance by Governments: The Technique of Computer Matching' Forthcoming, Information Technology & People

Clarke R.A. (1994c) 'A Normative Regulatory Framework for Computer Matching' Forthcoming, J. Computer and Info. L.

Cohen (1982) 'Oversight of Computer Matching to Detect Fraud and Mismanagement in Government Programs' Hearings Before the Sub-Committee on Oversight of Government Management, Senate Committee on Governmental Affairs, U.S. Govt Printing Office, Washington DC (Dec 15-16, 1982)

Dasgupta A.K. & Pearce D. (1972) 'Cost Benefit Analysis: Theory and Practice' Macmillan, London, 1972

DoF (1991) 'Handbook of Cost-Benefit Analysis' Dept of Finance, Canberra, September 1991

DoF (1992) 'Introduction to Cost-Benefit Analysis for Program Managers' Dept of Finance, Canberra, March 1992

DoF (1993) 'Value For Your IT Dollar: Guidelines for Cost-Benefit Analysis of Information Technology Proposals' Dept of Finance, Canberra, July 1993

DSS (1991) 'Data Matching Program (Assistance and Tax): Report on Progress' Dept of Social Security and Data Matching Agency, Dept of Social Security, Canberra (October 1991)

DSS (1992) 'Data Matching Program (Assistance and Tax): Report on Progress - October 1992' Dept of Social Security, Canberra (October 1992)

DSS (1993) 'Data Matching Program: Report on Progress - October 1993' Dept of Social Security, Canberra (October 1993)

Early P. (1986) 'Big Brother Makes a Date' San Francisco Examiner, 12 Oct 1986

FACFI (1976) 'The Criminal Use of False Identification' U.S. Federal Advisory Committee on False Identification, Washington DC, 1976

Fischel M. & Siegel L. (1980) 'Computer-Aided Techniques Against Public Assistance Fraud: A Case Study of the Aid to Families with Dependent Children (AFDC) Program' Prepared by the U.S. Law Enforcement Administration by MITRE Corp. 1980

Flaherty D.H. (1989) 'Protecting Privacy in Surveillance Societies' Uni. of North Carolina Press, Chapel Hill, 1989

Fraud (1987) 'Review of Systems for Dealing with Fraud on the Commonwealth' Aust. Govt. Publ. Service (March 1987)

GAO (1983a) 'Action Needed to Reduce, Account For, and Collect Overpayments to Federal Retirees' Comptroller General of the United States, General Accounting Office, Washington DC, 1983

GAO (1983b) 'Computer Matches Identify Potential Unemployment Benefit Overpayments' Comptroller General of the United States, General Accounting Office, Washington DC, 1983

GAO (1984a) 'GAO Observations on the Use of Tax Return Information for Verification in Entitlement Programs' Comptroller General of the United States, General Accounting Office, Washington DC, 1984

GAO (1984b) 'Better Wage-Matching Systems and Procedures Would Enhance Food Stamp Program Integrity' General Accounting Office, Washington DC, 1984

GAO (1985a) 'Eligibility Verification and Privacy in Federal Benefit Programs: A Delicate Balance' General Accounting Office, GAO/HRD-85-22, 1985

GAO (1985b) 'A Central Wage File for Use by Federal Agencies: Benefits and Concerns' General Accounting Office, GAO/HRD-85-31, May 1985

GAO (1986a) 'Social Security: Pensions Data Useful for Detecting Supplemental Security Payment Errors' General Accounting Office, GAO/HRD-86-32, Mar 1986, 14 pp.

GAO (1986b) 'Computer Matching: Assessing Its Costs and Benefits' General Accounting Office, GAO/PEMD-87-2, Nov 1986, 102 pp.

GAO (1986c) 'Computer Matching: Factors Influencing the Agency Decision-Making Process' General Accounting Office, GAO/PEMD-87-3BR, Nov 1986, 30 pp.

GAO (1987) 'Welfare Eligibility: Deficit Reduction Act Income Verification Issues' General Accounting Office, GAO/HRD-87-79FS, May 1987, 93 pp.

GAO (1988) 'Veterans' Pensions: Verifying Income with Tax Data Can Identify Significant Payment Problems' General Accounting Office, GAO/HRD-88-24, Mar 1988, 100 pp.

Accounting Office, GAO/HRD-88-24, Mar 1988, 100 pp.

GAO (1989a) 'Interstate Child Support: Case Data Limitations, Enforcement Problems, Views on Improvements Needed' General Accounting Office, GAO/HRD-89-25, Jan 1989, 82 pp.

GAO (1989b) 'Child Support: State Progress in developing Automated Enforcement Systems' General Accounting Office, GAO/HRD-89-10FS, Feb 1989, 25 pp.

GAO (1990a) 'Computer Matching: Need for Guidelines on Data Collection and Analysis' General Accounting Office, GAO/HRD-90-30, Apr 1990, 16 pp.

GAO (1990b) 'Veterans' Benefits: VA Needs Death Information From Social Security to Avoid Erroneous Payments' General Accounting Office, GAO/HRD-90-110, Jul 1990, 14 pp.

GAO (1990c) 'Computers and Privacy: How the Government Obtains, Verifies, Uses and Protects Personal Data' General Accounting Office, GAO/IMTEC-90-70BR, Aug 1990, 68 pp.

GAO (1991a) 'Federal Benefit Payments: Agencies Need Death Information From Social Security to Avoid Erroneous Payments' General Accounting Office, GAO/HRD-91-3, Feb 1991, 23 pp.

GAO (1991b) 'Computer Matching Act: Many States Did Not Comply With 3-Day Notice or Data-Verification Provisions' General Accounting Office, GAO/HRD-91-39, Feb 1991, 29 pp.

GAO (1991c) 'Welfare Benefits: States Need Social Security's Death Data to Avoid Payment Error or Fraud' General Accounting Office, GAO/HRD-91-73, Apr 1991, 13 pp.

GAO (1993) 'Computer Matching: Quality of Decisions and Supporting Analyses Little Affected by 1988 Act' GAO/PEMD-94-2, 18 October 1993

Gramlich E.M. (1981) 'Benefit-Cost Analysis of Government Programs' Prentice-Hall 1981

Greenberg D.H. & Wolf D.A. (1984) 'An Evaluation of Food Stamp and AFDC/Wage Matching Techniques: Final Report' Prepared for the U.S. Dept. of Agriculture, Food and Nutrition Service by SRI Int'l (May 1984)

Greenberg D.H. & Wolf D.A. (1985) 'Is Wage Matching Worth All the Trouble?' Public Welfare (Winter 1985) 13-20

Greenberg D.H. & Wolf D.A. (with Pfiester J.) (1986) 'Using Computers to Combat Welfare Fraud: The Operation and Effectiveness of Wage Matching' Greenwood Press Inc., Oct 1986

Greenleaf G.W. & Clarke R.A. (1984) 'Database Retrieval Technology and Subject Access Principles' Aust. Comp. J. 16,1 (Feb, 1984)

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

HHS (1983a) 'Computer Matching in State Administered Benefit Programs: A Manager's Guide to Decision-Making' Dept. of Health and Human Services, 1983a

_____ (1983b) 'Inventory of State Computer Matching Technology' Dept. of Health and Human Services, 1983b

_____ (1983c) 'Nationwide Impact of Verifying Resources by Matching Recipients of Public Assistance with Bank Records' Dept. of Health and Human Services, 1983c

_____ (1984) 'Computer Matching in State Administered Benefit Programs' U.S. Dept. of Health and Human Services, June 1984

HUD (1988) 'OIG Communiqué', U.S. Department of Housing and Urban Development', Washington DC (September 1988)

Kelly P. (1992) 'Australian Federal Privacy Laws and the Role of the Privacy Commissioner in Monitoring Data Matching' in Clarke R. & Cameron J. (Eds.) 'Managing Information Technology's Organisational Impact II' Elsevier/North Holland, Amsterdam, 1992

Kirchner J. (1981) 'Privacy: A history of computer matching in federal government programs' Computerworld (December 14, 1981)

Kling R. (1978) 'Automated Welfare Client Tracking and Welfare Service Integration: The Political Economy of Computing' Commun. ACM 21,6 (June 1978) 484-93

Kusserow R.P. (1983) 'Inventory of State Computer Matching Technology' Dept. of Health & Human Services, Office of Inspector-General (March 1983)

_____ (1984a) 'The Government Needs Computer Matching to Root out Waste and Fraud' Comm ACM 27,6 (June 1984) 542-545

_____ (1984b) 'Computer Matching in State-Administered Benefit Programs' Dept. of Health & Human Services, Office of Inspector-General (June 1984)

Laudon K.C. (1974) 'Computers and Bureaucratic Reform' Wiley, New York, 1974

Laudon K.C. (1986) 'Dossier Society: Value Choices in the Design of National Information Systems' Columbia U.P., 1986

Laudon K.C. (1993) 'Markets and Privacy' Proc. Int'l Conf. Inf. Sys., Orlando FL, Ass. for Computing Machinery, New York, 1993, pp.65-75

Lindop (1978) 'Report of the Committee on Data Protection' Cmnd 7341, HMSO, London (December 1978)

Marx G.T. & Reichman N. (1984) 'Routinising the Discovery of Secrets' Am. Behav. Scientist 27,4 (Mar/Apr 1984) 423-452

Mishan E.J. (1977) 'Cost-Benefit Analysis' 2nd Edition, George Allen & Unwin, London, 1977

NSWPC (1977) 'Guidelines for the Operation of Personal Data Systems' New South Wales Privacy Committee, Sydney, 1977

O'Connor K. (1990) Paper for the Communications & Media Law Association, available from the Privacy Commissioner, Human Rights & Equal Opportunities Commission, G.P.O. Box 5218, Sydney, 26 April 1990

OECD (1980) 'Guidelines for the Protection of Privacy and Transborder Flows of Personal Data' Organisation for Economic Cooperation and Development, Paris, 1980

OMB (1979a) 'Guidelines to Agencies on Conducting Automated Matching Programs' Office of Management and Budget, Washington DC, March 1979

OMB (1979b) 'Privacy Act of 1974: Supplemental Guidance for Matching Programs' 44 Fed. Reg. 23, 138. 1979

OMB (1982a) 'Computer Matching Guidelines', Office of Management and Budget, Washington DC, May 1982

OMB (1982b) 'Privacy Act of 1974: Revised Supplemental Guidance for Conducting Matching Programs' 47 Fed. Reg. 21, 656, 1982

OMB (1983) 'Agency Computer Match Checklist' Memorandum M-84-6, Office of Management and Budget, Washington DC, December 29, 1983

OMB (1989) 'Final Guidelines Interpreting the Provisions of Public Law 100-503, the Computer Matching and Privacy Protection Act of 1988' Fed. Reg. 54, 116 Mon. June 19, 1989, 13 pp.

OMB (1992) 'Annual Report on the Federal Agencies' Implementation of the Computer Matching and Privacy Protection Act of 1988 for the Calendar Year 1990' Office of Management and Budget, Washington DC, October 1992

OMB/PCIE (1983) 'Model Control System for Conducting Computer Matching Projects Involving Individual Privacy Data' Office of Management and Budget & President's Commission for Integrity & Efficiency 1983

OTA (1985a) 'Fair Information Practices in Seven States' Unpublished Internal Consultant's Report by R.E. Smith, Office of Technology Assessment, Jan 1985

OTA (1985b) 'Federal Government Information Technology: Electronic Surveillance and Civil Liberties' OTA-CIT-293, U.S. Govt Printing Office, Washington DC, Oct 1985

OTA (1986a) 'Federal Government Information Technology: Management, Security and Congressional Oversight' OTA-CIT-297, U.S. Govt Printing Office, Washington DC, Feb 1986

OTA (1986b) 'Federal Government Information Technology: Electronic Record Systems and Individual Privacy' OTA-CIT-296, U.S. Govt Printing Office, Washington DC, Jun 1986

OTA (1987) 'Defending Secrets, Sharing Data: New Locks and Keys for Electronic Information' OTA-CIT, U.S. Govt Printing Office, Washington DC, Oct 1987

OTA (1988) 'Electronic Delivery of Public Assistance Benefits: Technology Options and Policy Issues' OTA-BP-CIT-47, U.S. Govt Printing Office, Washington DC, Apr 1988

PCA (1990) 'Data Matching in Commonwealth Administration: Discussion Paper and Draft Guidelines' Privacy Commissioner, Human Rights & Equal Opportunities Commission, G.P.O. Box 5218, Sydney, Australia (October 1990) (56 pp.)

PCA (1991a) 'Interim Report on Operation of Data-Matching Program under the Data-Matching Program (Assistance and Tax) Act 1990' Privacy Commissioner, Human Rights Australia, Sydney (September 1991)

PCA (1991b) 'Data Matching in Commonwealth Administration: Revised Draft Guidelines' Privacy Commissioner, Human Rights & Equal Opportunities Commission, September 1991

PCA (1992) 'Data-Matching in Commonwealth Administration: Report to the Attorney-General' (incorporating Proposed Guidelines) Privacy Commissioner, Human Rights Australia, Sydney (June 1992)

PCA (1993) 'Data-Matching: Operation of Voluntary Guidelines: Review' Discussion Paper, Privacy Commissioner, Human Rights Australia, Sydney (December 1993)

PCC (1989) 'Data Matching Review: A Resource Document for Notification of the Privacy Commissioner of Proposed Data Matches' Privacy Commissioner, Ottowa (July 1989)

PPSC (1977) 'Personal Privacy in an Information Society' Privacy Protection Study Commission, U.S. Govt. Printing Office, July 1977

Reichman N. & Marx G.T. (1985) 'Generating Organisational Disputes: The Impact of Computerization' Proc. Law & Society Ass. Conf., San Diego, June 6-9, 1985

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

Rule J.B., McAdam D., Stearns L. & Uglow D. (1980) 'The Politics of Privacy' New American Library, 1980

Sassone P.G. & Schaffer W.A. (1978) 'Cost-Benefit Analysis: A Handbook' Academic Press, 1978

Smith R.E. (1974-) 'Privacy Journal' monthly, since November 1974

Smith R.E. (1976-) 'Privacy Journal Compilation of State and Federal Privacy Laws', updated annually, since 1976

SMOS (1987) 'Review of Systems for Dealing with Fraud on the Commonwealth' Aust. Govt. Publ. Service (March 1987)

SSA (1990) 'Guide for Cost/Benefit Analysis of SSA Computer Matches' Office of the Chief Financial Officer, Office of Program and Integrity Reviews, Social Security Administration, March 1990

Sugden R. & Williams A. (1985) 'The Principles of Practical Cost-Benefit Analysis' Oxford U.P., 1985

Thom J. & Thorne P. (1983) 'Privacy Legislation and the Right of Access' Austral. Comp. J. 15,4 (November 1983) 145-150

Thompson M. (1980) 'Benefit-Cost Analysis for Program Evaluation' Sage, 1980

Weiss L.B. (1983) 'Government Steps Up Use of Computer Matching To Find Fraud in Programs' Congressional Qtly Wkly Report February 26, 1983

Westin A.F. (1967) 'Privacy and Freedom' Atheneum, 1967

_____ (ed.) (1971) 'Information Technology in a Democracy' Harvard U.P., 1971

Westin A.F. & Baker M. (1974) 'Databanks in a Free Society' Quadrangle, 1974

Navigation

Go to Roger's Home Page

Go to the contents-page for this segment.

Send an email to Roger

Created: 18 September 1996

Last Amended: 27 January 1998

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