Chapter 5

Computer Profiling
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>summary</td>
<td>87</td>
</tr>
<tr>
<td>Background</td>
<td>87</td>
</tr>
<tr>
<td>Findings</td>
<td>89</td>
</tr>
<tr>
<td>Finding 1</td>
<td>89</td>
</tr>
<tr>
<td>Finding 2</td>
<td>91</td>
</tr>
<tr>
<td>Finding 3</td>
<td>93</td>
</tr>
<tr>
<td>Finding 4</td>
<td>94</td>
</tr>
</tbody>
</table>
Chapter 5

Computer Profiling

SUMMARY

While computer profiling is not currently a subject of major policy debate, the potential policy issues raised by the future growth of computer profiling are important. In computer profiling, a record system (or record systems) is searched for a specified combination of data elements, i.e., the profile. Profiling involves the use of inductive logic to determine indicators of characteristics and or behavior patterns that are related to the occurrence of certain behavior.

A profile is developed by a government agency to select characteristics of types of individuals, and to determine the probabilities of such individuals engaging in activities or behavior of interest to that agency. For example, the Drug Enforcement Agency (DEA) has developed profiles of the types of persons more likely to be engaging in illegal drug activity; the Internal Revenue Service (IRS) has developed profiles of categories of taxpayers more likely to be under-reporting taxable income; and the Federal Bureau of Investigation (FBI) has developed profiles of violent offenders. Profiles can be valuable tools for investigative, administrative, and intelligence purposes because they reduce the population that is of interest to an agency, and thus may increase the agency’s efficiency and effectiveness.

OTA found that:

- Federal agencies are currently using computer profiling and it is likely that its use will expand in the near future.
- Important privacy and constitutional implications are raised by computer profiling because people may be treated differently before they have done anything to warrant such treatment.
- The validity of computer profiles in accurately selecting the desired subset of individuals is subject to debate, and thus also raises questions about the relevancy of data used and the appropriateness of using computer profiles for certain decisions.
- At the present time, there are no policy guidelines for agency use of computer profiling.

BACKGROUND

Before computers were used to process and store information, systematic data on large numbers of individuals were not retained (or if retained were not readily accessible). Moreover, there was no easy means to analyze the data that did exist in order to construct profiles. Information technology in general—and computers in particular—have removed these constraints. Detailed, historical information on individuals can be compiled from various computerized databases. Computers can be used to analyze complex and disparate information and, based on that analysis, to design a profile. Additionally, computers can be used to search a record system on the basis of a profile. These technological changes make profiles both more powerful and more available. Most importantly, technology is now making possible many new profiling applications for which judgments of social acceptability have yet to be made.

Profiling involves the use of inductive logic to determine indicators of characteristics and or behavior patterns that are related to the occurrence of certain behavior. A judgment is
made about a particular individual based on the past behavior of other individuals who appear statistically similar, that is, who have similar demographic, socioeconomic, physical, or other characteristics. Generally, in the Federal Government, the behavior of interest is actual or potential violation of a law or administrative regulation.

In the past, and as is often still the case, people who appeared suspicious or acted strangely were often watched more carefully and their stories were verified from outside sources. Searches through Federal record systems were often conducted on the basis of a list of characteristics that experience had shown were problematic. Such profiles were often crude and could easily lead to the stereotyping of individuals. Today, profiling is much more sophisticated as a result of advances in behavioral psychology and statistics. As most behavior is complex, sophisticated modeling may be done to determine the interrelations among certain indicators. There are two general models of profiling. One is singular profiling, which models distinct characteristics or activities, e.g., sex, age, income, or number of dependents. When these characteristics appear together or in a certain pattern, that individual is flagged by the profile. The second model of profiling is aggregative profiling, which is based on the frequency with which selected factors appear across cases. This model is designed to find systematic and repetitive violators.

Profiles have been used for decisionmaking in a variety of areas, ranging from insurance and advertising to motor vehicle or real estate licensing to entrance to the medical and legal professions. Profiles used range from those that are benign and socially acceptable (e.g., granting driver's licenses to 16 year olds, who inmost States are judged to be physically and mentally mature enough to drive a car) to those that are discriminatory and socially unacceptable (e.g., denying rental housing to minorities or students or denying professional employment opportunities to women).

Profiles have been used by the government to help agencies uncover possible misrepresentation of eligibility to receive Federal funds or benefits, possible noncompliance with or violation of agency regulations, and possible violation of civil or criminal statutes. In the government, profiles can be created, to some extent, for the convenience of implementing public policies, as they replace subjective judgments with objective decisionmaking criteria. Profiles can be useful during any stage of an agency's interaction with individuals. For example, in eligibility benefit programs, profiling may be used at the application stage to determine if an applicant is likely to misrepresent his or her income, or at the redetermination stage to ascertain if it is likely that an individual's status has changed. In law enforcement, profiling may be used in discovering likely suspects (e.g., airplane hijackers) or in determining an appropriate sentence for someone convicted of a crime. Profiles can be valuable tools for investigative, administrative, and intelligence purposes because they reduce the population that is of interest to an agency, and thus may increase the agency's efficiency and effectiveness.

Because computer profiling may result in selected individuals being treated differently from those not selected, it has raised a number of policy questions involving civil, constitutional, and equal rights considerations. The primary conflict is between the rights of the individuals selected (e.g., equal protection and due process) and the purpose of the government in using computer profiles and their effectiveness in achieving that purpose. No matter how sophisticated the profile, the question of treating people differently before they have acted remains.

Computerized profiling also introduces some very important new policy issues. If the use of computer profiling in the Federal Government were to be expanded, the long-term societal effects on behavior patterns, and the possible effects on individuality and creativity, would warrant attention. Additionally, the va-
lidity of computer profiles in accurately selecting the desired subset of individuals is subject to debate, and thus also raises questions about the relevancy of data used and the appropriateness of using computer profiles for certain decisions.

**FINDINGS**

Finding 1

Federal agencies are currently using computer profiling and it is likely that its use will expand in the near future.

Federal agencies have developed profiles for a number of purposes, mainly for identifying individuals most likely to be involved in an illegal activity or most likely to misrepresent their financial or personal situation in applying for a Federal benefit. The OTA survey revealed that 16 Federal agencies presently use computer profiling. For example, the IRS uses computer-generated generic profiles to identify potential compliance deficiencies; the Department of Education uses profiles, based on criteria including taxes paid, marital status, and size of household, to select Pell Grant applicants for validation; the Bureau of Indian Affairs profiles the public social service support and facilities usages and needs of individual corporate groups of Indians for budgetary planning and allocation of resources; and the Federal Reserve Board uses surveys of retail consumers to obtain statistical data concerning financial status and behavior of households and businesses, access to and use of consumer credit, asset holdings, financial practices, effect of charge card transactions, and the like.

According to the OTA survey, some agencies are planning to add this capability to existing systems. For example, the redesign of the Treasury Enforcement Communications System, known as TECS II, will incorporate profiling. The U.S. Army Criminal Investigation Command is considering developing a system of profiling potential victims and criminal offenders for use in the conduct of crime prevention surveys and in the development of investigative leads. Some agencies have conducted pilot programs of profiling that are no longer in use, for example, the Office of the Inspector General in the Department of Energy developed, with DOE Defense Programs, a profile of the ‘Insider Criminal.”

The use of profiles for law enforcement purposes has been widely documented. Computers were not necessarily used in preparing these, but they are illustrative of the type of computer profiles already under development. The Drug Enforcement Agency (DEA) has developed a profile of airplane passengers likely to be smuggling drugs, and a profile to detect those transporting marijuana on trains. The Coast Guard has a profile of vessels likely to be smuggling drugs into the country. The Customs Bureau also has a “smuggler’s profile.” The Federal Aviation Administration used a hijacker profile as part of its screening program at domestic airports until it began routine searches of all carry-on items and magnetometer screening of all passengers.

The FBI has developed numerous profiles, including those of various violent criminals and serial murderers. This work is being expanded under the auspices of the FBI National Center for the Analysis of Violent Crimes. Also, based in large part on interviews with felons convicted of serial murders, the FBI has developed profiles of serial murderers, especially

---


See, for example, *United States v. Klein*, 592 F.2d 909 (5th Cir. 1979), and *United States v. Asbury*, 586 F.2d 973 (2d Cir. 1973).

serial sex murderers. The FBI is currently developing software for preparing computerized profiles of violent offenders, based on the concept already implemented for arson offenders in the computer-assisted Arson Information Management System (AIMS). In 1983, the Office of Juvenile Justice and Delinquency Prevention of the Department of Justice funded the University of Pennsylvania School of Nursing to identify the variables that fit profiles of rapists, child molesters, and sexually exploited children.

In the 1970s, the Law Enforcement Assistance Administration funded "pre-delinquency" programs to create computer models to identify those young people who were likely to become delinquent. The computer models or profiles included factors that were common among known delinquent youths, such as area of residence, family situation, school performance, ethnic group, and medical history. Young people who most closely matched the profile were to be given special treatment. In 1983, the Office of Juvenile Justice and Delinquency Prevention funded the Rand Corp. to develop strategies based on the "pre-delinquency" presumption.

Computer profiles can also be used as a way of avoiding errors in Federal Government eligibility and benefit programs and as a way of allocating scarce investigative resources. Based on a computer profile, caseworkers can determine during the application process which applicants may need more careful checking. Characteristics often associated with errors could include basic factors such as age, race, or education level; some combination of factors; or more indirect factors, such as length of family separation, residency, or living with a specified relative. In 1979, the Supplemental Security Income's Office of Family Assistance reported that the following characteristics were used in error-prone profiles: earned income, home ownership, age 26 to 40, recent separation, bank account, and overdue redetermination of benefits.

In eligibility benefit programs, computer profiles or screens can also be used to search databases of recipients prior to conducting a computer match. The records that were selected by the profile would be the only ones subject to computer matching. A smaller number of records would then be matched. If the computer profile was effective in selecting those records most likely to contain errors, then the percentage of verifiable hits would increase. In this way, computer profiles or screens may make computer matching more effective and efficient. Additionally, cuts in the Federal budget may increase the pressure to use computer profiling not only to detect and prevent fraud and errors, but also to allocate the time of caseworkers or investigators.

There has been no survey of the use of computer profiles in social service programs at the Federal level. The President's Council on Integrity and Efficiency (PCIE) has released three inventories of Federal computer applications to prevent/detect fraud, waste, and mismanagement. The applications include matches, profiles, edits, scans, screens, analyses, and extracts. If one adopts the PCIE categorization, there were no profiles used prior to 1982, 13 profiles used in the period 1982-83, and five profiles used in the period 1984-85. However, agencies have sometimes placed computer applications that appear to be profiles in a different category, e.g., Project Sonoma—Welfare Fraud Profile is listed as a match. Some computer screens appear to be based on a computer profile (e.g., a Department of Education screen designed to identify, by selected criteria, guaranteed student loans.

---


"Pre-Delinquent Funding: Déjà Vu," Privacy Journal, April 1984, p. 3.


maintained by State Guaranty Agencies that are in excess of the regulatory maximum of 10 years), while others do not (e.g., prescription payments made by Blue Cross and Blue Shield, screened to ascertain whether that company was computing and claiming Medicaid prescription drugs in accordance with Federal procedures).

Information on State use of computer profiles is also sketchy. The Carter Administration's Eligibility Simplification Project reported on the use of error-prone profiles, primarily at the State level. According to its study, West Virginia had used computer profiling, or a selective case action system, for Aid to Families With Dependent Children (AFDC), food stamp, and Medicaid cases, based on a quality-control sample generated monthly by the computer. The profile was based on a statistical method of evaluating previous error situations and was modified periodically. Reportedly, from 1973 to 1976, the case error rate and payment error rate declined by 20 percent. The Eligibility Simplification Project found similar results with the use of error-prone profiling in South Carolina and New Hampshire. The Eligibility Simplification Project found that other States appeared to be experimenting with the use of such profiles in determining social service eligibility. A survey of seven States conducted for OTA in 1984 revealed that computer profiling was not used by those States.

Finding 2

Important privacy and constitutional implications are raised by computer profiling because people may be treated differently before they have done anything to warrant such treatment. Computer profiles involve categorizing people based on selected criteria, and then selecting a subset of these people for special treatment. The equal protection guarantees of the fifth and 14th amendments were designed to ensure that individuals were treated in a manner similar to other individuals, and that the government not treat individuals differently simply because they were members of a group. Although the government can classify people for special treatment, it cannot do so based on impermissible criteria (e.g., race, religion, or national origin), nor can it use a classification to arbitrarily burden a group of individuals. In computer profiling, the criteria used might be those that are already viewed as discriminatory under existing law—e.g., race, religion, national origin, and sex. For example, in DEA's drug courier profile, being Hispanic has appeared as one of the criteria. With sophisticated profiling, it may also be possible to use a number of related indicators rather than a category whose use would be illegal.

The equal protection clauses may also require that the criteria on which the profile is based be related to the behavior in question; otherwise, the selected group may be arbitrarily burdened. Additionally, the government program would need to be rationally related to achieving a legitimate purpose such as detecting fraud, waste, and abuse or apprehending drug smugglers.

The use of computer profiling may also conflict with the due process clauses of the fifth and 14th amendments that protect an individual against arbitrary treatment and provide an individual with certain procedural guarantees. Some argue that computer profiles eliminate the discretion and arbitrariness of investigative authorities, caseworkers, and parole officers. Others respond that profiles merely replace a crude form of profiling (hunches, for example) with a more sophisticated one. In either case, the due process clauses require rules and procedures to limit discretion and protect individuals from arbitrary treatment. In some instances, use of computer profiling may not provide for adequate rules and procedures.

With respect to the use of profiles in eligibility programs, Senator William Cohen reported that:

We have profiles that have been developed by computer, and disability payments that have been discontinued with no human contact coming about until such time as those cases are appealed to an administrative law judge. Two-thirds of the cases appealed are being reversed.17

The extreme result of a computer profile would be that benefits are terminated, which would not occur without a hearing. The more common result would be that a selected individual is subject to a more thorough investigation than others because he or she fits a profile. To some extent, this individual is regarded with suspicion based on the profile. Individuals may not know that they are being treated differently, and even if they do, may not know why.

With respect to the use of computer profiles in law enforcement, the primary issue is whether fitting a profile constitutes probable cause or reasonable suspicion and is reason to search or detain an individual. In determining whether an investigative stop is lawful, the courts balance the need for the search against the intrusion to the person. To justify the intrusion, law enforcement agents must be able to identify specific and articulable facts that show the intrusion is reasonably warranted.24

There have been a number of court cases involving the use of the drug courier profile, and, hence, this will serve as an example of the legal issues that arise with use of profiles for law enforcement purposes. Although this profile is not currently generated by a computer nor are computers necessarily used to search relevant databases, the legal issues would be similar whether or not a computer was involved. Agents typically use the drug courier profile as a tool in conducting surveillance on a group of people, generally those boarding or departing a plane. If agents see a person whose behavior fits a number of criteria in the profile, then they follow the person. If agents believe it is justified, they stop the individual, identify themselves as law enforcement agents, and request to see identification. Based on the information revealed and the behavior of the person, the agents may then “request” that the suspect accompany them to an office in the airport. There the person is told that he or she is suspected of carrying drugs, advised of his or her rights, and asked for permission to search his or her luggage and person.25

In cases in which the sole or primary justification for an investigative stop has been the drug courier profile, the lower courts have not been consistent in their rulings. For example, in United States v. McCaleb, 552 F.2d 717 (6th Cir. 1977) and State v. Washington, 364 So. 2d 958 (La. 1979), the courts reversed the appellants’ convictions based on investigative stops triggered by meeting a drug courier profile because their activities were too consistent with innocent behavior. In United States v. Vasquez, 612 F.2d 1338, an investigative stop based in part on a profile was judged valid.

In 1979, the Supreme Court ruled on two instances involving the use of the drug courier profile. In the first case, United States v. Mendenhall, 446 U.S. 544, the Court ruled that the investigative stop of Mendenhall, which was based on her fitting characteristics of the drug courier profile, was constitutional. However, the majority did not agree on why it was constitutional, giving little guidance to the lower courts on the acceptability of the profile in establishing justification for an investigative stop. One month later, the Court handed down.

a second decision dealing with the drug courier profile, Reid v. Georgia, 448 U.S. 438. In this case, the Court held that the investigative stop of Reid, based on his matching characteristics of the drug courier profile, was not constitutional. The Court described the drug courier profile as "a somewhat informal compilation of characteristics believed to be typical of persons unlawfully carrying narcotics."

Based on these two cases, the legal status of the present drug courier profile is in question. Moreover, the Reid opinion may imply that the constitutionality of the profile could turn on its sophistication. If this is true, then the use of computer-generated profiles in law enforcement may be considered a more valid investigative tool than the more informal profiles.

Federal court decisions since Mendenhall and Reid have not clarified the status of the use of a drug courier profile in an investigative stop. "In 1981, in United States v. Cortez, 101 S. Ct. 690, the Supreme Court approved use of a profile by border patrol agents to detect the smuggling of illegal aliens from Mexico to the United States.

Finding 3

The validity of computer profiles in accurately selecting the desired subset of individuals is subject to debate, and thus also raises questions about the relevancy of data used and the appropriateness of using computer profiles for certain decisions.

Profiles vary in their complexity and in the formality of statistical techniques on which they are based. Because computers are such powerful tools in analyzing and manipulating vast quantities of data, it is likely that profiles will become even more complex and formal. Regardless of their complexity and formality, profiles by definition are prone to some degree of error, as they are merely probability statements.

In formal profiles, when a general population is characterized and a profile developed, the profile is only a statistical average of that general population. The similarities among the population will be accentuated, while the differences will be ignored. If the profile was based on a sufficiently large population, it will have some value in selecting those of interest, but there will also be some margin of error in the profile. The types of errors will be false positives (identifying those who fit the profile, but do not fit the category sought) and false negatives (passing by those who do not fit the profile, but do fit the category sought). In developing the profile, the statistician will incorporate the degree of error that the user is willing to tolerate.

The more informal, crude profiles are greatly influenced by the experience and concerns of those who develop them. For example, in the case of the drug courier profile, the criteria that make up the profile have varied over time and with the city in which DEA agents are working. Some subset of the following are generally considered as the profile: the use of small bills for ticket purchase, travel to and from major drug import centers, travel for short periods of time, absence of luggage or empty luggage, travel under an alias, unusual itinerary, unusual nervousness, use of public transportation, making a phone call after deplaning, leaving a fictitious callback telephone number with the airline, attempting to conceal that someone is waiting for them or that they are traveling with someone, purchase of a one-way ticket, Hispanic origin, youth, luggage without identification tags, ticket purchased at the last minute or late arrival, and deplaning last. There is no record establishing how and why these characteristics have come to be included in the profile. There may also be some criteria that DEA keeps confidential.

The OTA survey asked agencies to provide both information on the development and testing of profile programs and any evaluation reports. Of the 16 agencies that reported profil-
ing activities, none had this information available. There are no known studies of the degree of error in profiles used in eligibility verification programs.

A principal policy issue involves determining the accuracy of a computer profile and its effectiveness in achieving the desired outcome. The cost-effectiveness of computer profiles has never been systematically studied. There are a number of costs that may need to be considered: 1) developmental costs, including research, testing, validation, and evaluation; 2) computer costs, including hardware and software; and 3) administrative costs, including follow-up on individuals who fit the profile. The costs to individuals who may needlessly be subject to investigation may also need to be considered. Additionally, as with computer matching, there may be hidden or secondary costs that need to be examined.

There are also a number of benefits that need to be considered, primarily increasing the effectiveness and efficiency of an investigation because the relevant population has been narrowed, and preventing and deterring illegal behavior.

Some information is available on the effectiveness of profiling for law enforcement purposes. None contains specific cost-benefit categories or figures. A 1981 FBI evaluation of psychological profiling found that, of 192 cases examined, in 77 percent the profile helped focus the investigation, in 20 percent it helped locate possible suspects, and in 17 percent the profile actually identified the suspect. (Totals exceed 100 percent since more than one type of assistance may apply to a single case.) The vast majority of cases were murder or rape investigations. 8

There are some sketchy statistics on the effectiveness of the drug courier profile in selecting persons carrying drugs. In United States v. Van Lewis, 409 F. Supp. 535 (E.D. Mich. 1976), testimony from DEA revealed that agents at the Detroit airport had searched 141 persons in 96 encounters, found narcotics in 77 of these encounters, and arrested 122 persons. Forty-three of the searches in which narcotics were found were nonconsensual. In 15 of the 25 consent searches, no illegal narcotics were found. In testimony in United States v. Price, 599 F.2d 494 (2d Cir. 1979), a DEA agent stated that about 60 percent of those he stopped, based on the drug courier profile, were carrying narcotics. However, it appears that no national statistics are available on the effectiveness of the drug courier profile.

Finding 4

At the present time, there are no policy guidelines for agency use of computer profiling.

The use of computer profiling raises a number of important policy questions. In determining the appropriate use of computer profiling, a number of factors warrant consideration, including:

1. The nature of the decision for which the profile is used. In other words, under what circumstances is it appropriate to use computer profiling? In answering this question, two distinctions may prove helpful. The first is the government purpose in using profiling—e.g., detection of fraud, waste, and abuse; detection of violent criminals; and detection of discrimination. It may be appropriate to use computer profiling for all of these purposes and for any other purposes. Alternatively, the dangers of categorizing people and the speculative nature of profiles may outweigh their general use, but not their use for specific purposes.

The second distinction is whether only one individual, or one group or class of individuals, is subject to the computer profile. A profile may provide the key by which a database of many individuals is searched. One individual may also be selectively compared to a profile. Because an individual may be affected differently


under the two circumstances, different standards could be considered for its use.

2. The nature and source of the data used. To be consistent with equal protection law, one could argue that computer profiles should not include criteria traditionally considered discriminatory, e.g., race, religion, national origin, or sex. It may also be necessary to eliminate or restrict the use of attributes that may substitute for the overtly discriminatory criteria. Additionally, it may be necessary to restrict the use of results of sophisticated invasive or intrusive psychological or physiological tests, e.g., genetic testing, in profiles.

In setting standards for the use of data, it may also be helpful to consider the source of the data in determining its relevance for a profile. For example, it may not be appropriate for IRS profiles to include information not provided by the taxpayer or not directly relevant to financial matters.

3. The rights of individuals, with respect to both decisions based on profiles and being the subject of profiling, regardless of use. Should individuals be informed that their records are being searched on the basis of a profile or that they are being compared to a profile? If they do not want to be subject to profiling, what are their remedies? If an individual is accorded different treatment because of the way he or she compares to a profile, what rights does he or she have and how can they be implemented?

4. The accuracy of the profile. Given that profiles themselves are prone to errors, some testing may be necessary prior to the use of a profile. Independent validation and testing of any software program used for profiling may be necessary to determine bias and accuracy. If profiles are to be used, guidelines may need to be developed for validation and testing. It may be necessary that this testing be done by a group (or groups) other than the one that developed the profile. Although it maybe difficult to get an exact accounting of costs and benefits, some outlining of the significant costs and benefits that are expected could also be done.

With respect to the drug courier profile, William Conley has suggested that testing should be done in two steps. First, establishing the percentage of those previously arrested who displayed a particular characteristic. Second, determining what percentage of all airplane passengers exhibit the same characteristic.”

“**Ibid., p. 863.**