January 10, 2003

The Honorable John Ashcroft
Attorney General
United States Department of Justice
Main Justice Building, Room 5137
950 Pennsylvania Avenue, N.W.
Washington, D.C. 20530

Dear Attorney General Ashcroft:

We are writing to inquire about the current "data mining" operations, practices and policies at the Department of Justice. Improved access to and the sharing of information among intelligence and law enforcement agencies at the federal, state and local levels is crucial in promoting our national security interests. These national security interests are most effectively and efficiently served, however, when the information being collected and shared is relevant, reliable, timely and accurate. As one recent expert report observed, "Data mining, like any other government data analysis, should occur where there is a focused and demonstrable need to know, balanced against the dangers to civil liberties. It should be purposeful and responsible." *(Protecting America's Freedom in the Information Age, A Report of the Markle Foundation Task Force*, October, 2002, p. 27.)

Adequate oversight by the Congress, and especially by the appropriate committees of jurisdiction, is essential in helping to ensure that adequate standards are set and met, so that these activities can be both effective and respectful of the constitutional rights of the American people. Accordingly, we are interested in learning the extent to which the Department is relying on data mining to deal with the terrorism threat or other criminal activity, and how this technology is being used.

We raise this inquiry against the backdrop of public concern over the Total Information Awareness System (TIA) being developed under the supervision of Admiral Poindexter within the Defense Advanced Research Project Agency (DARPA). TIA is intended, according to Department of Defense officials, to generate tools for monitoring the daily personal transactions by Americans and others, including tracking the use of passports, driver's licenses, credit cards, airline tickets, and rental cars. The Administration's goal is to turn these tools over to law enforcement agencies. According to press reports, one such tool, a software program called "Genoa," has already been delivered by DARPA to the Department of Justice.

Advances in the technological capability to search, track or "mine" commercial and government databases and Americans' consumer transactions have provided powerful tools that have dramatically changed the ways that companies market their products and services. Collection and use by government law enforcement agencies of such commercial transactional data on law-abiding Americans poses unique issues and concerns, however. These concerns include the specter of excessive government surveillance that may intrude on important privacy interests and chill the exercise of First Amendment-protected speech and associational rights.

Moreover, as Federal law enforcement agencies obtain public source and proprietary data for mining, the sheer volume of information may make updating the data and checks for reliability and accuracy difficult, if not impossible. Reliance on data mining by law enforcement agencies may produce an increase in false leads and law enforcement mistakes. While the former is a waste of resources, the latter may result in mistaken arrests or surveillance. Such mistakes do occur, even without data-mining.[1] In short, while the only ill effect of business reliance on outdated or incorrect information may be misdirected marketing efforts, data mining mistakes made by a law enforcement agency may result in misdirection or misallocation of limited government resources and devastating consequences for mistakenly targeted Americans.

We are interested in determining the extent to which the Justice Department is relying on data-mining and how the Department is addressing these concerns with appropriate safeguards on the collection, use and dissemination of information obtained through data mining. Specifically, we ask for and would appreciate your responses to the following questions.

1. **Data-Mining Operations Underway Within the Department of Justice.**

   (A) Please identify any private sector or proprietary databases obtained or being used by the Department of Justice for data-mining or pattern-recognition activities as well as any databases from government agencies outside DOJ being used for such purposes.

   (B) Have any private sector or proprietary databases referred to in (A) above been aggregated by the Department with any data from government agency databases for data-mining or pattern-recognition activities?
   (C) Is the Department using any data-mining tools to obtain information for law enforcement purposes unrelated to the detection and prosecution of terrorism?

   (D) To the extent that the Department is using proprietary data provided by private intermediaries, (i) what procedures are you using to preserve the confidentiality policies

---

[1] A recently declassified FBI memorandum, dated April 14, 2000, makes this point with startling details about incidents of mistaken surveillance activity, including a Foreign Intelligence Surveillance Act (FISA) order being improperly implemented with unauthorized videotaping of a meeting; wiretapping a cellular telephone that had been dropped by the target and assigned to an innocent user, who "was therefore the target of unauthorized electronic surveillance for a substantial period of time;" unauthorized monitoring of an e-mail account; and "unauthorized searches, incorrect addresses, incorrect interpretation of a FISA order and overruns of ELSUR [electronic surveillance]."

of these intermediaries? (ii) Is the Department compensating the private intermediaries for assisting in the data mining? (iii) Has the Department taken any steps to shield the private intermediaries from liability for their cooperation with the government?

(E) What procedures, if any, does the Department follow to ensure the accuracy and reliability of information currently collected and stored in databases used for data-mining?

(F) By contrast to the use of private sector or proprietary databases, in the search for proper data mining tools, to what extent is the Department of Justice developing new tools and to what extent is it making use of existing tools developed in the private sector or used by other government agencies (such as search engines and data mining software)? What are the pros and cons of these differing approaches?

2. Foreign Terrorist Tracking Task Force. On October 29, 2001, the President directed the Department to establish the Foreign Terrorist Tracking Task Force (FTTTF) to "ensure that, to the maximum extent permitted by law, Federal agencies coordinate programs to . . . 1) deny entry into the United States of aliens associated with, suspected of being engaged in, or supporting terrorist activity; and 2) locate, detain, prosecute, or deport any such aliens already present in the United States." Your April 11, 2002, order establishing the FTTTF would do more than ensure that agencies "coordinate programs" and requires the FTTTF to have "electronic access to large sets of data, including the most sensitive material from law enforcement and intelligence sources." In response to my request for more detailed description of the mission and activities of the FTTTF, you stated in response to written questions that:

"The FTTTF has identified a number of specific projects which it can coordinate or run to fill gaps in existing government efforts relating to prevention of terrorist activities. For example, the FTTTF is pursuing projects to: 1) create a unified, cohesive lookout list; 2) identify foreign terrorists and their supporters who have entered or seek to enter the U.S. or its territories; and 3) detect such factors as violations of criminal or immigration law which would permit exclusion, detention or deportation of such individuals. In addition, the FTTTF is in the process of identifying other intelligence-related projects that it can support through its collaborative capability to co-locate data from multiple agency sources."

(A) Redundancy within government programs can be both expensive and ineffective. The "projects" of the FTTTF appear to overlap other initiatives underway within the Department. For example, the FBI has an Information Sharing Task Force and participates in 47 Joint Terrorism Task Forces (JTTF) to unify all levels and branches of law enforcement in preventing and investigating terrorist activity and helps coordinate the JTTF in Regional Terrorism Task Forces (RTTF). Director Mueller has also created a permanent Terrorism Watch List, a new Office of Intelligence, a new Integrated Intelligence Information Application (IIIA) database, and new hiring and recruiting initiatives. **Please explain how the Department's FTTTF "lookout list" differs in substance and use from the FBI's Terrorism Watch List and how the FTTTF's**

"other intelligence-related projects" will differ from the functions of the FBI's JTTF, and IIIA database, and new Office of Intelligence. Please also explain how the FTTTF's "lookout lists" differ from or interface with those used by Customs, INS, and State Department (and successor agencies) for border control purposes and by the Transportation Security Administration.

(B) The FBI's new Office of Intelligence is intended to provide strategic analysis and gather information from current and past cases and other agencies, to look for patterns and analyze risks, and to meet the needs of other organizations responsible for homeland security. The separate FTTTF supervised by the Deputy Attorney General is required, with a budget of over $20 million, to conduct its own intelligence analysis projects and create and maintain its own databases and lookout list. **Since Director Mueller routinely briefs the President with the CIA Director on terrorist threats, please explain why you decided to place the FTTTP in the Deputy Attorney General's office rather than within the FBI as part of its new Office of Intelligence?** [2]

(C) The FBI performs its intelligence gathering mission under the supervision of a Director appointed for a ten-year term in a structure designed, in part, to insulate the exercise of Bureau powers from political considerations, and pursuant to formal Attorney General guidelines and Congressional oversight. You revised those guidelines in April 2002, but the guidelines you issued are limited to the FBI. They state, for example, "The FBI is authorized to operate and participate in identification, tracking and information systems for the purpose of identifying and locating terrorists..." **Are the investigative restrictions applicable to FBI agents also applicable to employees conducting data mining and operating the FTTTF under the guidance of the Deputy Attorney General? Conversely, given your guidelines on tracking terrorists are limited to the FBI, what is the source of and what are the guidelines defining the authority of the FTTTF?**

(D) What information is necessary to trigger a data-mining inquiry on a particular individual or targeted activity to ensure that this technique is only being used for purposes relevant to detecting, preventing or punishing terrorism or other criminal activity?

3. <u>Admiral Poindexter's Total Information Awareness Project (TIA).</u> According to the Department of Defense, the Defense Advanced Research Project Agency (DARPA) has established the Total Information Awareness (TIA) Project to develop technologies for rapid language translation, commercial transaction data mining, and interagency analysis and decision-making tools.

(A) To what extent are you and the Department of Justice consulting or collaborating with Admiral Poindexter or the Department of Defense in designing and implementing TIA surveillance tools and related programs?

---

[2] This question was originally directed to Deputy Attorney General Thompson in May 2002, but no response has been provided.

(B) Have any TIA generated or developed technologies been delivered to the Department of Justice and, if so, (i) are any being used? (ii) describe the purposes for which they are being used; and (iii) are any of the tools for data mining and pattern recognition?

(C) TIA has programs called Genoa I and II. Has this program been delivered in whole or in part to the Department of Justice and, if so, (i) is it being used? (ii) Describe the purposes for which it is being used; and (iii) is this a tool for data mining or pattern recognition?

(D) TIA has a program called EELD (Evidence Extraction and Link Discovery). Has this program been delivered in whole or in part to the Department of Justice and, if so, (i) is it being used? (ii) Describe the purposes for which it is being used; and (iii) is this a tool for data mining or pattern recognition?

(E) TIA has a program called Genisys. Has this program been delivered in whole or in part to the Department of Justice and, if so, (i) is it being used? (ii) Describe the purposes for which it is being used; and (iii) is this a tool for data mining or pattern recognition?

(F) TIA has a program called TIDES (Translingual Information Detection, Extraction and Summarization. Has this program been delivered in whole or in part to the Department of Justice and, if so, (i) is it being used? (ii) Describe the purposes for which it is being used; and (iii) is this a tool for data mining or pattern recognition?

(G) Is the FTTTF coordinating its work in any way with the TIA?

(H) What safeguards, if any, do you believe should be included in any data mining tools developed by TIA to ensure the accuracy and reliability of the information collected and stored in databases? Have you recommended such safeguards to the Department of Defense?

1. **Compliance With The Privacy Act.**

(A) Does the Privacy Act impose any restriction on data-mining activities by the Department and, if so, what are those restrictions?

(B) Does the Department employ any outside contractors to perform data mining services and, if so, how does the Privacy Act apply, if at all, to the out-sourcing of data mining activities?

(C) The Privacy Act, 5 U.S.C. §552a(e)(4), requires agencies to "publish in the Federal Register upon establishment or revision a notice of the existence and character of the system of records." Have you promulgated any regulations regarding the FTTTF?

(D) The Privacy Act, 5 U.S.C. §552a(e)(4)(E), requires publication of the policies and practices of the agency regarding storage, retrievability, access, controls, retention and disposal of the records. Have you published such policies and practices regarding the FTTTF?

(E) Generally, the Privacy Act prohibits governmental agencies from disclosing records to another agency, unless it falls under the "routine use" exception. 5 U.S.C. §552a(b)(3). Does the Department rely on this "routine use" exception to obtain databases from other agencies for aggregation in the FTTTF and other databases within the Department?

(F) The Privacy Act, 5 U.S.C. §552a(e)(4)(D), requires Federal Register publication of "each routine use of the records contained in the system, including the categories of users and the purpose of such use." If the answer to (E) above is affirmative, has the Department published any Federal Register notice required by the Privacy Act? If so, please provide a copy of any such notice and, if not, please explain why.

(G) The Privacy Act imposes restrictions on "matching" programs conducted by the government or the private sector on behalf of the government, unless the matching is conducted "subsequent to the initiation of a specific criminal or civil law enforcement investigation" or "for foreign counterintelligence purposes." How does the Department ensure that the FTTTF and other Department databases comprised of aggregated data from other agencies are operated within these restrictions?

(H) Does the Department believe that any amendments to the Privacy Act would be helpful to facilitate data mining by the Department and, if so, does the Department intend to transmit to the Congress any amendments to the Privacy Act to clarify the legality of data-mining by Federal agencies?

2. **Coordination With the Department of Homeland Security.**

(A) The Homeland Security Act expressly authorizes the new department to request, access, receive, analyze and integrate information from government agencies and private sector entities, and to establish and utilize "a secure communications and information technology infrastructure, including data-mining and other advanced analytical tools, in order to assess, receive and analyze data and information. . . ." [P.L. 107-296, Sections 201(d)(1), (13), (14)]. Does the Department of Justice have any such express statutory authority to conduct data mining? If so, please describe that authority.

(B) Do you anticipate the Department of Justice's data mining operations being transferred to the new Department of Homeland Security? If not, please explain why.

(C) Do you believe it is valuable to have a coordinated data mining effort with one agency clearly held accountable for setting guidelines of data uniformity and reliability and, if so, which agency do you believe should take this primary position in order to avoid duplication of effort?

We appreciate your attention to this important matter.

Sincerely,

PATRICK LEAHY Chairman	RUSSELL D. FEINGOLD United States Senator	MARIA CANTWELL United States Senator