U.S. Department of Justice
Office of Information Policy
Suite 11050
1425 New York Avenue, NW
Washington, DC 20530-0001

Telephone: (202) 514-3642

September 25, 2014

Mr. Steven Aftergood
Federation of American Scientists
1725 DeSales Street, NW Suite 600
Washington, DC 20036
saftergood@fas.org

Re: OLP/11-01073 (F)
VRB:DRH:JBG

Dear Mr. Aftergood:

This is a final response to your Freedom of Information Act request dated and received in this Office on August 8, 2011, in which you requested the latest copy of a Department of Justice report to Congress on the subject of "data mining," pursuant to section 804 of the Implementing Recommendations of the 9/11 Commission Act of 2007 (P. Law 110-53). This response is made on behalf of the Office of Legal Policy.

By e-mail dated July 27, 2012, we advised you that we had located the 2011 data mining report, and were processing that report in response to your request. You confirmed that this would be satisfactory. Accordingly, I have determined this report, and corresponding letter transmitting the report to Congress, are appropriate for release without excision, and copies are enclosed. Please be advised that the enclosed report covers the time period January 1, 2008 to September 30, 2009.

For your information, Congress excluded three discrete categories of law enforcement and national security records from the requirements of the FOIA. *See* 5 U.S.C. § 552(c) (2006 & Supp. IV 2010). This response is limited to those records that are subject to the requirements of the FOIA. This is a standard notification that is given to all our requesters and should not be taken as an indication that excluded records do, or do not, exist.

If you are not satisfied with my response to this request, you may administratively appeal by writing to the Director, Office of Information Policy, United States Department of Justice, Suite 11050, 1425 New York Avenue, NW, Washington, DC 20530-0001, or you may submit an appeal through this Office's eFOIA portal at http://www.justice.gov/oip/efoia-portal.html. Your appeal must be received within sixty days from the date of this letter. If you submit your appeal by mail, both the letter and the envelope should be clearly marked "Freedom of Information Act Appeal."

Sincerely,

Vanessa R. Brinkmann
Senior Counsel

Enclosures

U.S. Department of Justice

Office of Legislative Affairs

Office of the Assistant Attorney General

Washington, D.C. 20530

December 15, 2011

The Honorable John A. Boehner
Speaker
U.S. House of Representatives
Washington, D.C. 20515

Dear Mr. Speaker:

We are pleased to transmit the report required by the Federal Agency Data Mining Reporting Act of 2007. This reporting requirement is contained within Section 804 of the Implementing Recommendations of the 9/11 Commission Act of 2007, Pub. L. No. 110-53 ("Section 804"). Section 804 requires the Attorney General to submit a report to Congress regarding "the organization and operations of every program engaged in 'data mining,'" as defined in the statute.

Following a thorough review by the components and agencies of the Department of Justice, we have identified four initiatives that meet the definition of "data mining" under Section 804 that were conducted during the period from January 1, 2008 to September 30, 2009. Specifically, the report discusses programs carried out during the same time period by the Federal Bureau of Investigation and United States Attorneys' Offices, as reported by the Executive Office of United States Attorneys. The report also provides information on certain advanced analytic activities conducted by the Department during the same period that do not meet the definition of "data mining" set forth in Section 804, but that may nonetheless be perceived as "data mining," as that term is commonly understood. The report is enclosed.

We apologize for the delay in transmitting this report. Please do not hesitate to contact this office if we may be of further assistance with this or any other matter.

Sincerely,

Ronald Weich
Assistant Attorney General

Enclosure

UNITED STATES DEPARTMENT OF JUSTICE

REPORT ON "DATA-MINING" ACTIVITIES FROM JANUARY 1, 2008-SEPTEMBER 30, 2009

SUBMITTED PURSUANT TO SECTION 804 OF THE

IMPLEMENTING RECOMMENDATIONS OF THE 9/11 COMMISSION ACT OF 2007

Section 804 of the Implementing Recommendations of the 9/11 Commission Act of 2007, Pub. L. No. 110-53 (Act or Section 804), requires the heads of all agencies in the Federal Government to submit, within 180 days of enactment of the Act and annually thereafter, a report regarding the organization and operations of every program engaged in "data mining," as defined in the statute. For each such initiative, the head of the agency must provide:

- A thorough description of the data mining activity, its goals and, where appropriate, the target dates for the deployment of the data mining activity.
- A thorough description of the data mining technology that is being used or will be used, including the basis for determining whether a particular pattern or anomaly is indicative of terrorist or criminal activity.
- A thorough description of the data sources that are being or will be used.
- An assessment of the efficacy or likely efficacy of the data mining activity in providing accurate information consistent with and valuable to the stated goals and plans for the use or development of the data mining activity.
- An assessment of the impact or likely impact of the implementation of the data mining activity on the privacy and civil liberties of individuals, including a thorough description of the actions that are being taken or will be taken with regard to the property, privacy, or other rights or privileges of any individual or individuals as a result of the implementation of the data mining activity.
- A list and analysis of the laws and regulations that govern the information being or to be collected, reviewed, gathered, analyzed, or used in conjunction with the data mining activity, to the extent applicable in the context of the data mining activity.
- A thorough discussion of the policies, procedures, and guidelines that are in place or that are to be developed and applied in the use of such data mining activity in order to:
    - protect the privacy and due process rights of individuals, such as redress procedures; and

- ensure that only accurate and complete information is collected, reviewed, gathered, analyzed, or used, and guard against any harmful consequences of potential inaccuracies.

This report first discusses "data mining" as a conceptual matter, as well as the privacy concerns that may be implicated by advanced analysis of information obtained and retained by the government. It then discusses qualifying programs conducted during the period from January 1, 2008 to September 30, 2009, in response to each of the seven requested pieces of information listed above and set forth in Section 804. Specifically, the report discusses programs carried out during the same period by the Federal Bureau of Investigation (FBI) and United States Attorneys' Offices, as reported by the Executive Office of United States Attorneys (EOUSA). Finally, this report provides information on certain advanced analytic activities conducted by the Department during the same period that do not meet the definition of "data mining" set forth in Section 804, but that may nonetheless be perceived as "data mining," as that term is commonly understood.

I. Background

Section 804 requires the head of every agency in the Federal Government to issue, within 180 days of enactment of the legislation and annually thereafter, a report regarding the organization and operations of every program engaged in "data mining." Section 804(b)(1) defines "data mining" as:

> a program involving pattern-based queries, searches, or other analyses of 1 or more electronic databases, where—
>
> > (A) a department or agency of the Federal Government, or a non-Federal entity acting on behalf of the Federal Government, is conducting the queries, searches, or other analyses to discover or locate a predictive pattern or anomaly indicative of terrorist or criminal activity on the part of any individual or individuals;
> >
> > (B) the queries, searches, or other analyses are not subject-based and do not use personal identifiers of a specific individual, or inputs associated with a specific individual or group of individuals, to retrieve information from the database or databases; and
> >
> > (C) the purpose of the queries, searches, or other analyses is not solely—
> >
> > > (i) the detection of fraud, waste, or abuse in a Government agency or program; or
> > >
> > > (ii) the security of a Government computer system.

Three elements of the Act's definition warrant further discussion. First, the Act's definition includes *all* electronic databases that are searched in a manner described in the statute. Second, the definition is prospective in nature, as it applies only to those searches that attempt to locate a "*predictive* pattern or anomaly indicative of terrorist or criminal activity" (emphasis added).[1] As a result, programs dedicated to solving past crimes or incidents do not fall within this definition. Third, Section 804 excludes subject-based queries by focusing on queries, searches, or other analyses undertaken or authorized by the Federal Government to locate a predictive pattern or anomaly indicative of terrorist or criminal activity. Consequently, in applying this definition and in drafting this report, the Department has sought to locate and discuss programs that, through pattern-based queries, searches, or other analyses, attempt to predict future criminal or terrorist activity.

Data mining initiatives that analyze lawfully acquired information, as is the case with each of the qualifying initiatives, can be extremely valuable tools for investigators. These advanced analytic activities are grounded in traditional investigative techniques, but are designed to process information more efficiently and effectively than can be done by individual investigators conducting those tasks manually. Such initiatives must also be undertaken with deep respect for the privacy and civil liberties of Americans. All of the data mining initiatives undertaken by the Department meet both of these goals.

Federal statutes and internal Department policies and procedures are designed to mitigate potential privacy concerns. For example, as part of the Department's privacy compliance process, the Department instituted an Initial Privacy Assessment (IPA) form, which must be completed for all new or modified information systems and programs in the Department that contain personally identifiable information, including those involving data mining. The IPA permits a determination as to whether additional privacy documentation is necessary under either the Privacy Act (*e.g.*, a System of Records Notice (SORN)) or the E-Government Act of 2002 (*e.g.*, a Privacy Impact Assessment (PIA)). It also permits identification of any other legal or policy privacy issues under those statutes. The deployment of the IPA in the agency privacy compliance process ensures that an opportunity exists to examine all new or modified information systems and programs for potential privacy and civil liberties concerns.

PIAs, completed by Department components pursuant to the E-Government Act of 2002, address the issue of the existing authority for the collection and advanced analysis of information. The goal of a PIA is three-fold: (1) to ensure that handling of information conforms to applicable legal, regulatory, and policy requirements regarding privacy; (2) to determine the risks and effects of collecting, maintaining, and disseminating information in

[1] *See* Statement of Senator Feingold on the Introduction of S.236, 153 Cong. Rec. S359, S360 (Jan. 10, 2007) ("While it can be defined more broadly, for the purpose of this reporting requirement, *data mining is limited to the process of attempting to predict future events or actions* by discovering or locating patterns or anomalies in data." (emphasis added)).

identifiable form via an electronic information system; and (3) to evaluate protections and alternative processes for handling information to mitigate potential privacy risks.

In guidance updated in August 2006, the Department's Office of Privacy and Civil Liberties (OPCL) indicated that PIAs should be conducted when a component is, *inter alia*, (1) developing or procuring any IT systems or projects that collect, maintain, or disseminate information in identifiable form from or about members of the public; (2) initiating, consistent with the Paperwork Reduction Act, a new electronic collection of information in identifiable form for 10 or more persons (excluding agencies, instruments or employees of the federal government); or (3) changing an existing system in a manner that creates new privacy risks (such as when converting from paper-based records to electronic systems or when merging, centralizing, or matching databases that contain information in identifiable form with other databases).[2] The Department is developing additional inquiries as part of the PIA process to provide greater insight and analysis about possible data mining activities. Once the new PIA guidance is reviewed on a Department-wide basis, these inquires will be incorporated into the standard PIA Template.

Moreover, the Department has long been subject to, and is diligent in complying with, the Privacy Act of 1974, 5 U.S.C. § 552a. The Privacy Act's requirements generally apply to records that identify and are about U.S. citizens and legal permanent resident aliens and that are retrieved from a system by reference to an individual's name or other personal identifier. As a result, any information produced as a result of pattern-based data mining that meets these criteria is subject to the Act's requirements. Among these requirements are: (1) that the agency "maintain in its records only such information about an individual as is relevant and necessary to accomplish a purpose of the agency required to be accomplished by statute or by executive order of the President," 5 U.S.C. § 552a(e)(1); (2) that the agency publish descriptive notices in the Federal Register of all records systems about individuals from which information is retrieved by reference to their name or personal identifier, 5 U.S.C. § 552a(e)(4); (3) that the agency "maintain all records which are used by the agency in making any determination about any individual with such accuracy, relevance, timeliness, and completeness as is reasonably necessary to assure fairness to the individual in the determination," 5 U.S.C. § 552a(e)(5); and (4) that the agency "establish appropriate administrative, technical, and physical safeguards to insure the security and confidentiality of records and to protect against any anticipated threats or hazards to their security or integrity which could result in substantial harm, embarrassment, inconvenience, or unfairness to any individual on whom information is maintained," 5 U.S.C. § 552a(e)(10). While the FBI and EOUSA, as law enforcement agencies, have exempted their systems from subsections (e)(1) and (e)(5) pursuant to subsection (j)(2) of the Privacy Act, the

[2] "Privacy Impact Assessments: Official Guidance," Privacy and Civil Liberties Office, Office of the Deputy Attorney General (revised August 7, 2006); *see also* E-Government Act of 2002, Pub. L. No. 107-347, Title II, § 208, Dec. 17, 2002, 116 Stat. 2899, 2921, *codified at* 44 U.S.C.A. § 3501 note; OMB Memorandum 03-22, OMB Guidance for Implementing the Privacy Protections of the E-Government Act of 2002 (Sept. 23, 2003).

agencies nonetheless recognize the need for relevant and accurate information in carrying out their law enforcement missions. Furthermore, an exemption cannot be claimed from (e)(4)(A)-(F) or (e) (10), *inter alia*, nor from subsection (b) of the Act, the very core of the Act that prohibits disclosure of Privacy Act information, except under certain circumstances.

One potential privacy issue with respect to any pattern-based data mining initiative is whether the pattern-based data mining is undertaken for a legitimate purpose. In general, such initiatives have long been recognized as legitimate and permissible law enforcement techniques.[3] In fact, each of the initiatives described below in more detail is grounded in traditional law enforcement techniques designed to discern patterns of criminal or terrorist activity and to appropriately focus available resources. The initiatives are simply designed to accomplish these goals with greater efficiency and accuracy. In addition, in each of the substantive areas in which pattern-based data mining initiatives have or are being developed, the Department has statutory authority to conduct criminal or terrorist investigations, and no law prohibits the data mining initiatives described herein.

A second potential privacy issue relates to the security of the information and how it is retained. In this regard, agencies that administer a pattern-based data mining initiative must ensure that the information is secure and that users utilize the particular tools only for authorized purposes. The Privacy Act requires that agencies "establish appropriate administrative, technical, and physical safeguards to insure the security and confidentiality of records and to protect against any anticipated threats or hazards to their security or integrity which could result in substantial harm, embarrassment, inconvenience, or unfairness to any individual on whom information is maintained." 5 U.S.C. § 552a(e)(10). In addition, the Federal Information Security Management Act of 2002 (FISMA), along with the Federal Information Processing Standards (FIPS) published by the National Institute of Standards and Technology (NIST), define requirements for securing agency information systems. These requirements are implemented at the agency level with information security plans that include certain specific controls designed to ensure that only individuals with proper authorization can access the pattern-based data mining tools and that users have proper training on the use and sensitivity of the system. Controls such as audit logs also help ensure that authorized individuals are only using the data mining tools for official business. Again, where applicable, a PIA, and to some degree a SORN, will include descriptions of such security controls. Moreover, both the FBI and United States Attorneys' Offices (coordinated by EOUSA) are required to comply with the Privacy Act's subsection (b) disclosure prohibition, which, in addition to its general prohibition, restricts disclosure within those components to the officers and employees "who have a need for the record in the performance of their duties," 5 U.S.C. § 552a(b), as well as the Privacy Act's

[3] *See, e.g.*, Data Mining: An Overview, Congressional Research Service, December 16, 2004.

subsection (c)(1) requirement that they keep an accurate accounting of disclosures made outside of their component.[4]

A subsequent potential privacy concern relates to the security of information once the analysis has been undertaken. Again, the protections required by FISMA and implemented in Departmental security policies, requiring strict access controls and audit capabilities, ensure that such data is not accessed by unauthorized users. If information from a data analysis initiative ends up in an investigative file, the data is retained in accordance with the retention schedule of the investigative file; and if that investigative file is subject to the Privacy Act, then that record will also be subject to the protections of the Privacy Act.

As described above, a PIA conducted for a system will require an agency to evaluate the potential privacy risks of a pattern-based data mining initiative and describe mitigation procedures that have been put in place to counter such potential risks. One of the mandatory questions in the Department's standard PIA requires information about the security, quality control, and auditing features of the system, *e.g.*, whether a right of access and amendment exists for records in the system and who has access to the system. The Department's OPCL is fully engaged in the development and analysis of any PIA on a major information system or national security system done by any component within the Department, providing additional insight into the potential privacy concerns at stake and potential for mitigating those concerns. Furthermore, in several of the initiatives described below, personal information is not forwarded to FBI investigators unless it is necessary for opening an investigation pursuant to the Attorney General's Guidelines for Domestic FBI Operations (collectively Attorney General's Guidelines).[5] By minimizing the access to personal information, the risk of a security breach of this data is lessened.

The final privacy issue relates to the accuracy of the data to be searched and the potential for misidentification of innocent persons by a pattern-based data mining initiative. To the extent that a Department component accesses or acquires data from outside sources, initial responsibility for the accuracy of such data rests with the entity that generated the information. Where the source is another government agency, such as the Federal Trade Commission, and the data is in records covered by the Privacy Act of 1974, the attendant accuracy requirements of that statute apply to the agency that generates the records.[6] In some of the initiatives described in this report, the queried data is supplied by individuals who are likely to provide accurate data as they are voluntarily providing information as a victim. With respect to initiatives in which

---

[4] Disclosures made under the Freedom of Information Act, 5 U.S.C. § 552, are excluded from this requirement. 5 U.S.C. § 552a(c)(1).

[5] On October 3, 2008, the Attorney General and Director of the FBI announced the issuance of consolidated Attorney General Guidelines to govern domestic FBI operations. The new guidelines took effect on December 1, 2008.

[6] 5 U.S.C. § 552a(e)(5).

information is obtained from a commercial data aggregator, these private entities have strong business incentives to consistently provide accurate information. There are comparable measures in other initiatives designed to minimize the risk of inaccurate and unreliable results, and the Department is committed to complying with accuracy requirements for all retrieved data regardless of the collection source. Furthermore, in each initiative in which the data comes from victims or other members of the public, an analyst will verify the data with basic analytic tools to correct misspellings and obvious errors before it is used.

As to the accuracy and completeness of data searched by pattern-based data initiatives, search results are routinely checked by the use of the following measures. Leads generated by pattern-based data mining initiatives are not automatically accepted and acted upon, thus reducing the risk of "false positives." Rather, query results from these initiatives are independently evaluated by highly skilled analysts. The results are then passed along to investigators who also closely review results before taking any investigative action. These results are only used for lead purposes, and no action is taken based solely on the analytic products produced by such pattern-based data mining initiatives. Internal Department and FBI procedures, including the Attorney General's Guidelines,[7] set forth the Department's general policy that investigations should be undertaken by non-intrusive means prior to the use of more intrusive investigative means, unless the aforementioned reviews determine further investigation using more intrusive means is relevant and appropriate.

Simply put, no one is labeled a terrorist or a criminal simply because that individual appears in a database or appears as a result of some set of data mining queries. Moreover, the data mining initiatives discussed in this report do not preempt or abrogate other requirements investigators and analysts must satisfy in order to pursue more intrusive techniques. For example, investigators still must have sufficient probable cause in order to obtain a warrant.

The Department realizes that there are privacy risks inherent in the use of pattern-based data mining initiatives, as there are with most law enforcement investigative techniques. As with all law enforcement techniques, the Department strives to mitigate such potential privacy risks through compliance with federal statutes and internal policies and regulations. Through such mitigation, the Department's agencies carry out their law enforcement and terrorism prevention missions while protecting the privacy and civil liberties of our nation's citizens.

II. Description of Programs

A. Federal Bureau of Investigation Data Mining Programs

This report discusses in detail FBI initiatives that arguably meet the criteria for predictive data mining under Section 804. The SAR Initiative is discussed in this report. Other initiatives are discussed in the classified annex to this report.

---

[7] *Supra* note 5.

FBI initiatives that do not fall within Section 804's formal definition of "data mining" include:

1. Durable Medical Equipment Initiative
2. Investigative Data Warehouse
3. Identity Theft Intelligence Initiative
4. Health Care/Medicare Fraud Initiative
5. Housing/Mortgage Fraud Initiative.
6. Automobile Accident Insurance Fraud Initiative
7. Health Care/Prescription Fraud SearchPoint Initiative
8. COPLINK
9. Internet Pharmacy Fraud Initiative

These programs, which were previously summarized in the Department's Data Mining Report to Congress Pursuant to Section 804 of the Implementing Recommendations of the 9/11 Commission Act of 2007, submitted in December 2008, are discussed in summary format in Part III, *infra*.

Suspicious Activity Report (SAR) Initiative

(1) **Description**

On a regular basis, FBI Headquarter Divisions, as well as a number of FBI field offices, analyze Suspicious Activity Reports (SARs) filed by financial institutions, money service businesses, securities firms, and casinos with the U.S. Department of Treasury's Financial Crimes Enforcement Network (FinCEN), pertaining to transactions occurring within their area of responsibility. SARs are required to be filed by financial institutions operating in the United States whenever known or suspected criminal conduct was perpetrated against or took place at the financial institutions. This requirement includes transactions that are indicative of possible criminal activity. SARs are used by FBI field offices for multiple purposes. One goal of this initiative is to cross-reference SAR-related activity with pending FBI investigations in order to de-conflict pending investigations. Another goal is to develop information that may trigger a new investigation into certain finance-related crimes, such as money laundering or terrorism financing. Although the available information is necessarily retrospective, past practices and patterns may suggest a likelihood of future criminal activity, including possible terrorist

financing. SARs may also be used by FBI field offices to identify individuals for possible recruitment as sources.

In various jurisdictions, a working group of law enforcement agencies chaired by the local U.S. Attorney's Office also meets on a regular basis to review SARs filed with FinCEN that pertain to the specific jurisdiction. During these meetings, the SARs are reviewed to identify possible criminal activity. Additional reports may also be generated keyed to a specific dollar amount, or by the number of transactions exceeding a specified dollar amount occurring in a specific jurisdiction. For example, a report may be generated identifying individuals in a specific jurisdiction who have conducted five transactions exceeding a total of $250,000 in the past 12 months. These SARs are then disseminated to working group members for appropriate follow-up. Although the FBI's management of SAR generally meets the definition of "data mining," the activities of these individual working groups do *not* generally qualify as data mining, as they are retrospective in nature. To the extent a United States Attorney's Office takes different steps and engages in data mining, it is discussed in the EOUSA portion of this report.

(2) **Technology and Methodology**

The FBI has three methods of accessing electronic SAR data. The first is by use of the Financial Institution Fraud (FIF) database contained in the FBI's Automated Case Support (ACS) information system. SARs filed by financial institutions are forwarded by FinCEN to the FBI and then uploaded into the FBI's FIF database. This data is then available through the FBI's secure computer system (FBINET) for use by offices investigating money laundering or bank fraud violations. SAR data applicable to a specific geographic location (such as a State) is extracted from ACS and then cross-referenced with information contained in other FBI databases, such as the Investigative Data Warehouse (IDW), National Crime Information Center (NCIC) database, and the Telephone Application (TA) database (containing telephone subscriber data acquired by the FBI through statutorily authorized investigative tools).

A second method of accessing FinCEN SAR data is through FinCEN and the IRS Enterprise Computing Center. The Department of Treasury has developed a web-based system to provide direct access to SAR data through a direct link to their Currency Banking Retrieval System, commonly known as WebCBRS. This system provides authorized users the ability to query and download BSA records for certain enumerated purposes, including criminal, tax, regulatory, and intelligence, and counterintelligence activities to protect against international terrorism. Through WebCBRS, BSA reports can be queried and viewed individually or downloaded to delimited text files, which can then be imported into analytic software applications such as Microsoft Excel and Access. The information in WebCBRS is collected by FinCEN, who is responsible for its accuracy. The FBI is one of many users, along with state, local, and other federal law enforcement and regulatory agencies. To enhance the FBI's use of BSA reports from WebCBRS, a controlled interface (CI) has been developed to allow authorized FBI users direct access from FBINET workstations.

A third method of accessing FinCEN SAR data is through the FBI's Investigative Data Warehouse (IDW). Through a special agreement with FinCEN, SAR data is contained in several databases available in IDW. Using IDW permits more flexible search options than are available through WebCBRS as well as the ability to search multiple data sources with a single query. The data in IDW, while not as current as data available through WebCBRS, is updated regularly and may be accessed directly from FBINET.

After initial analysis of SAR data, analysis may then continue by cross-referencing SAR data information contained in various Internet-based private and public-source databases. These private databases may include AutoTrack, LexisNexis, Accurint, TARGUSinfo, and Dun & Bradstreet. Public source databases may include Google, Yahoo, and government databases of corporate filings, public records, and similar sources of information.

Regardless of the method of initial access to SARs, once possible links between the SAR data and information contained in other FBI databases are identified, the SARs and accompanying links to other FBI information may be placed into a Microsoft Access database and reviewed for patterns or anomalies that might suggest criminal or terrorist financing.

(3) **Data Sources**

As noted above, SARs are the principal source of data being analyzed in this activity. The data in SARs may then be further reviewed against information contained in various databases, including FBI databases such as IDW, ACS, and NCIC; Internet-based private databases such as AutoTrack, LexisNexis, and Accurint; and public-source databases such as Google, Yahoo, and databases of public government records, such as corporate filings.

(4) **Efficacy**

To the extent that some field offices primarily review SARs in order to identify links between recent SARs and pending and past FBI investigations, cross-referencing information from SARs with information contained in FBI databases such as ACS and IDW serves to identify and de-conflict pending FBI investigations. In those field offices using SARs to develop information that may support investigations into financial crimes or terrorism financing, the information contained in SARs has proven to be an effective indicator of potential illegal activity warranting additional investigation.

The information contained in SARs is thought to be generally reliable, as SARs are filed under the authority of the Bank Secrecy Act (BSA), 31 U.S.C. §§ 5311-5332, and, in the case of banks, under Federal bank supervisory agency general rule-making authority contained in Title 12 of the United States Code. The SAR requirements derive from 1992 amendments to the BSA, *see* The Annunzio-Wylie Money Laundering Act, Pub. L. No. 102-550, 106 Stat. 3672, 4044-4074, and regulations first promulgated in 1996. Additionally, although 31 U.S.C. § 5318(g)(3) essentially immunizes reporting institutions from civil liability for submitting SARs, the safe harbor provision requires a "good faith suspicion" that a law or regulation has been violated. *See*

Lopez v. First Nat'l Bank, 129 F.3d 1186 (11th Cir. 1997). As a result, filing a SAR knowing it contains inaccurate information might subject the institution to civil liability at the hands of the subject.

(5) **Privacy and Civil Liberties Impact**

The personal information contained in SARs and provided to FinCEN is required to be reported by statute and regulation. The personal information about individuals required to be reported (name, address, Social Security Number, date of birth, etc.) pertains only to the individual that is the subject of the SAR. SAR information is then reviewed by the FBI and other law enforcement agencies participating in various U.S. Attorneys' Offices' SAR review teams. The FBI also reviews SARs uploaded into the FBI's ACS system.

However, because financial institutions file SARs with the expectation that they will be accorded sensitive treatment, the FBI considers SARs similar to confidential sources of information that, when further investigated, may produce evidence of criminal activity. Consistent with the treatment accorded confidential source information, the existence of SARs relating to conduct being investigated, as well as the content of SARs, is not normally disclosed to persons outside the law enforcement community.

(6) **Law and Regulations**

The legal foundation for the FBI to conduct such investigations is derived from the following:

- 28 U.S.C. § 533 authorizes the FBI to investigate violations of federal law, including money laundering (18 U.S.C. § 1956), bank fraud (18 U.S.C. § 1344), and terrorist financing (18 U.S.C. § 2339A);

- 28 U.S.C. § 534 authorizes the FBI to collect and retain criminal information;

- 28 C.F.R. § 0.85 authorizes the FBI to conduct federal criminal investigations;

- The Attorney General's Guidelines[8] authorize the FBI to conduct investigations of violations of federal criminal laws; to employ all lawful techniques in that pursuit; and to collect and retain information from lawful sources in compliance with the Constitution and federal law; and

[8] *Supra* note 5.

- The Attorney General's Guidelines authorize the FBI to conduct investigations to obtain information or to protect against threats to the national security; to employ all lawful techniques in those endeavors; and to collect, maintain and disseminate information collected pursuant to those Guidelines.[9]

(7) **Privacy and Accuracy Protection Policies**

As noted above, SARs are considered sensitive law enforcement information, and their disclosure to the subject of a SAR might compromise an existing or potential law enforcement investigation. Given the nature of the information contained in SARs and the purposes for which such information is collected, there are therefore strict statutory restrictions governing disclosures of SARs, or the fact that SARs have been filed.

To the extent that personally identifiable information contained in SARs is integrated into FBI files, such information is subject to the full panoply of privacy protections applicable to FBI files and records. First, the FBI complies with the Privacy Act of 1974, FISMA, FIPS published by NIST, and privacy policies established by the Department of Justice in 28 C.F.R. §§ 16.40 and 16.54 and Department of Justice Order 2640.1. Each of these sets forth requirements for securing agency information and information technology systems. With respect to the Privacy Act, the information collected on individuals for whom suspicious financial activity is indicated is entered into the ACS system, which is part of the FBI Central Records System for which there is both a published system of records notice and published exemptions from the notice and personal right of access provisions of the Privacy Act.[10] The FBI has prepared and is reviewing a Privacy Threshold Analysis (the FBI's equivalent of the IPA) for the FBI's access to SAR data through WebCBRS. The Privacy Threshold Analysis will be submitted to the OPCL for a determination of whether a PIA is required by either the E-Government Act or Department policy.

Second, pursuant to the Attorney General Guidelines,[11] no investigative activity is initiated by an FBI field office against any transaction participant identified by this initiative unless the criteria established in the Attorney General Guidelines are met—which includes the logical evaluation of lead information through other non-intrusive, lawful means (FBI record checks, private lender record checks, developed sources, etc.). In addition, the use of more intrusive techniques (grand jury subpoenas, administrative subpoenas, tasking of sources, undercover operations and electronic surveillance, etc.) is regulated by law and procedure designed to ensure that the techniques are lawfully and appropriately employed.

---

[9] *Supra* note 5.

[10] The system of records notices appear at 63 Fed. Reg. 8,671 (Feb. 20, 1998), amended 66 Fed. Reg. 8,425 (Jan. 31, 2001), 66 Fed. Reg. 17,200 (Mar. 29, 2001) and 72 Fed. Reg. 3,410 (Jan. 25, 2007), and the exemption regulations appear at 28 C.F.R. § 16.96 (2009).

[11] *Supra* note 5.

Third, all personally identifying information (PII) from SARs contained in the FBI's FIF database application is accessible only through the FBI secure password-controlled internal computer system, FBINET. FBINET is accessed only from FBI computers located within restricted spaces in FBI field offices and FBI Headquarters. Access is limited by FBI policy to those in the unit and field offices with a need to know. Information from the FIF database of investigative value is entered into FBI ACS, a password-controlled database with a robust audit capability. There are distinct access and security restrictions on the use of ACS in the FBI's Manual of Investigative and Operational Guidelines and the FBI's Security Policy Manual.

Similarly, the PII contained in both WebCBRS and IDW is only available through secure FBI computer systems. WebCBRS may be accessed through UNet, the FBI's unclassified connection to the Internet. Once pertinent data has been identified, however, it is transferred to the FBI's secure internal system, FBINET. Both UNet and FBINET require individual user names and strong passwords. In addition, access to either system is subject to monitoring through user access and use logs.

B. United States Attorneys' Offices Data Mining Programs

1. **Project SEAHAWK, Intermodal and Port Security Pilot Project**

(1) **Description**

Project SeaHawk was a port security pilot project created to enhance maritime and intermodal transportation security in South Carolina's ports. Project SeaHawk was operated under the direction and authority of the U.S. Attorney's Office in the District of South Carolina until September 30, 2009, when it was transferred to the Department of Homeland Security. It served as one of the District's counterterrorism and critical infrastructure initiatives through its Anti-Terrorism Advisory Council. SeaHawk's objective was to prevent and disrupt criminal or extremist activity associated with intermodal transportation by enhancing existing security and the security processes. During the reporting period of January 1, 2008 to September 30, 2009, the SeaHawk Task Force analyzed Unclassified and Law Enforcement Sensitive information pertaining to the international shipping industry. This information was compared against various patterns of potentially illicit activity developed from previous Law Enforcement cases or intelligence information pertaining to extremist or criminal smuggling enterprises. These patterns of activity, known as Suspect Indicator Protocols (SIPs), were then used to screen inbound vessels to identify potential linkages with extremism or other illicit activities. The results of any query generating matches or links were briefed to the SeaHawk Unified Command, comprised of representatives from the U.S. Coast Guard, Customs and Border Protection, Immigration and Customs Enforcement, the FBI and the Joint Terrorism Task Force, South Carolina's State Police, Ports Authority Police, and other Law Enforcement Officers from local jurisdictions associated with the ports of Charleston, Georgetown, and Beaufort, South Carolina.

Project SeaHawk was established in March 2003. The Link Analysis and Data Analysis (LADAS) tool was completed in November 2007 and underwent operational testing until

September 30, 2009, when the SeaHawk Pilot Project was ended.

(2) **Technology and Methodology**

The SeaHawk Task Force collected specific reporting and other information from a variety of federal and commercial sources as part of the intermodal screening process. All information collected under this effort was archived in a dedicated information portal hosted in a secure federal facility located in Charleston, South Carolina. This information was organized in three broad volumes in the SeaHawk portal: Shipping Volume, Ship Management Volume, and Extremist Related Volume (explained below in greater detail).

As part of the maritime screening of vessels intending to call in South Carolina ports, the SeaHawk staff conducted a series of reviews based on SIPs. A SIP was derived from a review of previous cases of known events associated with extremist exploitation of the maritime industry or other criminal smuggling enterprises involving the commercial maritime environment. The screening involved the use of the LADAS tool to conduct a broad based query against the three volumes of information in the SeaHawk portal. All crew members, managers, and owners of a vessel intending to call in a South Carolina port were run through this screening process using the LADAS tool. Results of any query generating matches or links in the SeaHawk Information Portal were briefed to the SeaHawk Unified Command.

LADAS was a software solution consisting of commercial, off-the-shelf (COTS) packages integrated together with custom software code that allowed analysts at Project SeaHawk to consume pre-gathered documents rapidly and transform unstructured narrative text in various documents into structured data that could then be visualized using linkage diagrams.

The purpose of the LADAS tool was to allow analysts to sift rapidly through the continuous flow of commercial maritime information pertinent to international trade as transacted in a U.S. port with the materials they already found to be the most pertinent and useful information derived from published open-source documents.

LADAS was made up of three primary software components: Text Analytics (TA), an Intelligence Analyst's Database (IAD), and Analyst Notebook (ANB). TA extracted data from unstructured text documents and exported it into the IAD. In the IAD, data was stored based on the data type (*e.g.*, person, organization, ship).

In the course of conducting their analyses, analysts queried the IAD to discover links for their analysis. ANB provided a graphical user interface that supported developing linkage diagrams to show relationships between subjects of an analysis. The TA database existed in an Oracle Relational Database Management System (RDBMS). The IAD existed on a Structured Query Language server.

When data was brought into an implementation of the LADAS tool, analysts could use the data loading utilities and text processing utilities to identify objects (*i.e.*, people, places, things) and to establish the relationship between those objects. The identified objects and relationships were used to construct a semantic graph. A semantic graph was a pictograph of the

meaning of a body of words. That picture might have been a link analysis chart, a bar graph, or an information array in a spreadsheet. Attributes about the objects were stored outside the main graph. Project SeaHawk implemented a bibliography of all source documents used for LADAS exploitation which was regularly reviewed.

LADAS could not be used to identify particular pieces of information automatically, nor could it be used to find new data. LADAS was only used to represent particular pieces of information and the relationships between that information.

The security layer of the SeaHawk infrastructure hosting the LADAS tool included access control and authentication services to ensure that only individuals who had received approval could access the system and that their access credentials were authentic. The LADAS tool further provided capabilities to limit access to data to only those with a need to know and in accordance with the policies of the implementing organization. The SeaHawk portal provided the capability to restrict access to data based upon the role(s) assigned to each individual. The SeaHawk portal also implemented the concept of communities of interest, which assigned data to a specific group; authorized users were granted permission to access each group based upon the user's need to know.

(3) **Data Sources**

The screening used a broad based query against three information portals of Unclassified and/or Law Enforcement data: shipping volume information, ship management information, and extremist related information.

- Shipping Volume: Information tagged for the shipping volume was derived from the Coast Guard's Ship Arrival and Notification System (SANS), the U.S. Customs and Border Protection's (CBP) Treasury Enforcement Communications System (TECS), and commercial information acquired from Lloyd's of London. This information included multi-sourced data associated with a commercial ship calling in any South Carolina port. Archived data elements included details on the ship's characteristics, ship photography, ship's voyage history, ship's incident history, and the crew list of that vessel on each occasion the ship called in South Carolina.

- Shipping Management Volume: Information tagged for the Ship Management volume was derived from Lloyd's of London Maritime Information Unit. This data included detailed information associated with the ownership and management of all ships in global commercial service. Archived data elements included details on the ship's owner, manager, associated phone numbers and addresses of managing companies, and other ships managed or owned.

- Extremist Related Volume: This data included published Department of Justice information derived from press reports on terrorism-related convictions. Additionally, other press or academic documents associated with extremism were

collected from the Internet.

The information utilized for Project SeaHawk was either Unclassified or Law Enforcement Sensitive. The Shipping Volume and Shipping Management Volume data was from verified information submitted to the U.S. Government as part of the official record of Customs, Immigration and Ships Security and Safety documentation. The volume of information pertinent to extremist-related reporting was attributed to open-source reporting derived from the Department of Justice or other U.S. Government documents pertinent to terrorism-related convictions. Other sources of extremist-related reporting were sourced to press or academic documents that were verified with all-source information. All open-source non-federal government-originated information was carefully evaluated prior to incorporation of that data into the LADAS environment. This evaluation was accomplished through a tiered review process of source documents that were weighed and judged against other all-source reports. A detailed bibliography was maintained of all documents entered into the LADAS environment for both oversight purposes. Procedures were developed to delete any documents that were subsequently determined to have contained inaccurate information.

(4) **Efficacy**

Project SeaHawk's link analysis program, LADAS, was created to generate potential leads for law enforcement investigations or potential follow-on national intelligence collection using national technical means. The information utilized in LADAS was retrieved from government and commercial databases, such as the Coast Guard's SANS, CBP TECS, and data from Lloyd's of London. Initial responsibility for accuracy with respect to the information lay with the original record owner. By comparing the shipping and management data to existing data about the Department's terrorism-related convictions, there was a high degree of probability that the resulting data set would identify probable offenders.

LADAS was delivered to Project SeaHawk for initial field testing in November 2007. Software and hardware issues were identified and corrected. The tool achieved Initial Operational Capability in approximately February of 2008. LADAS was very effective in helping Project SeaHawk analysts identify potential intermodal activities of interest contained in tens of millions of elements of information.

(5) **Privacy and Civil Liberties Impact**

Project SeaHawk had limited or no impact on the privacy and civil liberties interests of U.S. persons. The only personal information involved in this initiative was information provided to the U.S. Government on the part of a ship as part of its official announcement to call in a U.S. port. Because the information was derived from reports submitted on the part of a visiting ship to the U.S., this information was part of the national record on international commerce, trade, immigration, safety, and customs. This information was retained as part of the security efforts to evaluate all future commercial maritime notifications of arrivals.

(6) **Law and Regulations**

Project SeaHawk was established in March 2003 as a Congressionally-funded pilot project to enhance maritime and intermodal transportation security in South Carolina. In addition, the Department of Justice complied with current laws and regulations regarding privacy, such as the Privacy Act of 1974, FISMA, and FIPS published by NIST, and Department of Justice Order 2640.1. Each of these sets forth requirements for securing agency information and IT systems. No investigative activity was initiated by this initiative unless the criteria established in the relevant Attorney General Guidelines were met – which included the logical evaluation of lead information through other non-intrusive, lawful means. Other applicable guidelines included:

- Homeland Security Presidential Directives (HSPD) 2, 6 and 11- all of which direct the strengthening of screening and analysis program to detect, identify, and interdict individuals entering or within the United States who pose a terrorist threat to national security.

- National Security Presidential Directive 46 (War on Terror) also sets forth strengthening of terrorist screening tools as a major objective of national policy.

- Homeland Security Act of 2002, §§ 201(d)(1), (d)(14); P.L. 107-296 (Nov. 25, 2002).

(7) **Privacy and Accuracy Protection Policies**

Project SeaHawk used the LADAS tool strictly as a means of enhancing the screening of international commercial maritime traffic into the ports of South Carolina. More than 95% of this activity is conducted by foreign national crewmen, companies and vessels. Although 5% of this traffic may be transacted by U.S. citizens serving onboard ships calling in Charleston, file structures were not maintained on U.S. citizens, nor was any personal information collected other than what was provided to the U.S. government by vessels visiting any South Carolina port. No information collected from open-source brokers of public record data was entered directly into the LADAS tool or hosted in LADAS data archives.

2. <u>SAR Review, Western District of New York</u>

(1) **Description**

The Western District of New York (WDNY) SAR Review Team was created to identify and predict criminal activity through proactive analysis of U.S. Bank Secrecy Act data. The SAR Review Team operates under the direction and authority of the U.S. Attorney's Office in the Western District of New York. The primary goal of the WDNY SAR Review Team is to identify finance-related federal crimes and protect the homeland. A second goal is to utilize SAR information related to the WDNY proactively to predict future activities of identified criminals and terrorists.

The WDNY SAR Review Team utilizes FinCEN data to detect, identify, track, and interdict people and organizations committing criminal acts and posing threats to the homeland. FinCEN administers the Bank Secrecy Act (BSA), which requires depository institutions and other industries vulnerable to money laundering to file and report certain data about financial transactions possibly indicative of money laundering and other criminal activity. FinCEN provides direct, electronic access to that data to qualifying law enforcement agencies through the Currency and Bank Retrieval System (CBRS). Law Enforcement agencies acquire access to CBRS by signing a Memorandum of Understanding (MOU) with FinCEN.

The WDNY SAR Review Team consists of representatives from the U.S. Attorney's Office, Drug Enforcement Administration, FBI, Internal Revenue Service (IRS) Criminal Investigation, Secret Service, Immigration and Customs Enforcement, U.S. Postal Inspection Service, IRS-BSA, U.S. Department of Housing and Urban Development, Department of Defense, ATF, New York Joint Terrorism Task Force, New York State Police, New York State Attorney General's Office, New York State Department of Taxation and Finance, New York State Insurance Department, Erie County (New York) District Attorney's Office, and foreign law enforcement from the Canada Revenue Agency. The team independently examines each referral for investigative merit. Upon review, additional FinCEN reports may be generated based on specific subject searches, transaction amounts, or other characteristics such as zip codes.

The WDNY SAR Review Team was formed in late 2004. The proactive use of the FinCEN databases to identify and predict criminal activity was developed during 2006. This program is an ongoing initiative.

During the early part of the reporting period of January 2008 to September 30, 2009, the WDNY SAR Review Team divided its overall initiative into three sub-initiatives focused on (1) Russian and Eastern European criminal organizations; (2) cigarette smuggling operations; and (3) international wire transfers in advance of terrorist operations. Due to resource limitations, however, the SAR Review Team's initiatives were curtailed for a significant portion of the reporting period.

(2) **Technology and Methodology**

The proactive analysis begins by accessing BSA data through a secure connection to the Department of Treasury's FinCEN CBRS. The SAR Review Team queries the CBRS for data relating to the WDNY using specific criteria such as identified subjects, geographic locations, and information filers. Both null and positive results constitute BSA information subject to the FinCEN MOU.

Positive results are manually cross-referenced with information from secondary sources, including open-source public records from commercial data brokers and Internet search engines, to ensure the information is accurate, complete, timely, and relevant. Available Internet search engines, as well as subscription sites, are utilized to track recently acquired targets in an effort to identify additional targets and queries. The WDNY SAR Review Team manually assesses each

individual SAR for investigative merit and patterns of anomalies that may help predict future criminal activity.

(3) **Data Sources**

The SAR Review Team analyzes Unclassified, Law Enforcement Sensitive, and public information. The primary data source is CBRS, which includes BSA information filed in Currency Transaction Reports (CTRs), Foreign Bank Account Reports (FBARs), SARs, and Money Service Business (MSB) registration reports. These reports are filed by covered financial institutions pursuant to the BSA and implementing rules administered by FinCEN.

Secondary sources of data include well known private companies brokering public data, such as LexisNexis, and Internet search engines. When applicable, information and reports from other law enforcement agencies are also utilized.

(4) **Efficacy**

The principal source of data being analyzed is CBRS, which is administered by the Department of Treasury's FinCEN. Non-SAR BSA CBRS data is believed to be reliable. If errors are detected, correspondence letters are mailed to the filer asking for missing and incomplete information. Correspondence replies are posted to the CBRS upon receipt. Note that this process is not utilized for SARs due to the strict confidentiality restrictions in applicable law with respect to SARs. In addition, all open-source non-federal government originated information is carefully evaluated by the SAR Review Team against other all-source information.

The SAR Coordinator also pulls all SAR material that appears to fit each respective assessment. The data is manually reviewed closely to determine the nature of the suspicious activity and whether known and new patterns of activity are occurring. SARs identified as having merit are provided to agents for discussion and further investigation. This method of SAR review has produced accurate indicators of potential illegal activity warranting investigation.

(5) **Privacy and Civil Liberties Impact**

Member agencies of the WDNY SAR Review Team have the authority and responsibility to investigate crimes and to conduct initiatives to detect, deter, and prevent criminal and terrorist activities. The SAR Review Team initiatives are grounded in traditional law enforcement techniques designed to discern patterns of criminal activity and to focus resources appropriately. The WDNY SAR Review initiative is designed to accomplish these goals with greater efficiency and accuracy through the use of electronic data sources, such as CBRS and LexisNexis.

The rules of conduct pertaining to the use of BSA information in the CBRS are established by the Department of Treasury and enforced through FinCEN. Only authorized personnel may access the CBRS. Each authorized user is assigned a unique user identification and password. BSA Information may be utilized by authorized personnel to identify, investigate,

or prosecute possible or actual violations of criminal law that fall within the investigative or prosecutorial jurisdiction of the Agency or Task Force.

In accordance with Department policies, the U.S Attorney's Office in the Western District of New York strives to mitigate potential privacy risks through compliance with Federal statutes and Departmental policies and regulations. Only positive information and leads are disseminated to the SAR Review Team. The information is reviewed for accuracy and completeness before dissemination. To protect privacy and civil liberties further, information derived through the public database searches is provided to agents and stored with the investigative file. In addition, the CBRS searches are conducted in government space with restricted access. Therefore, the impact or likely impact of this data mining initiative on privacy and civil liberties is estimated to be none to minimal.

(6) **Law and Regulations**

The Department of Justice complies with current laws and regulations regarding privacy, such as Privacy Act of 1974, FISMA, and FIPS published by the NIST, as well as Department of Justice Order 2640.1. Each of these sets forth requirements for securing agency information and IT systems. No investigative activity is initiated by the initiative unless the criteria established in the relevant Attorney General Guidelines are met. These criteria include the logical evaluation of lead information through other non-intrusive, lawful means. The Executive and Legislative branches have recognized data mining as a legitimate law enforcement analytic technique.

Other guidelines include:

- HSPDs 2, 6 and 11- all of which direct the strengthening of screening and analysis program to detect, identify and interdict individuals entering or within the United States who pose a terrorist threat to national security.

- National Security Presidential Directive 46 (War on Terror) also sets forth strengthening of terrorist screening tools as a major objective of national policy.

- Homeland Security Act of 2002, §§ 201(d)(1), (d)(14); P.L. 107-296 (Nov. 25, 2002).

(7) **Privacy and Accuracy Protection Policies**

The SAR Review Team does not automatically act upon search results generated from pattern-based data mining initiatives. The results are independently evaluated by U.S. Attorney and SAR Review Team personnel, which may include the SAR Coordinator or Paralegal. In accordance with Department procedures, these initiatives aim to investigate lead information by non-intrusive means. The Attorney General's Guidelines,[12] as well as internal Departmental guidelines, set forth this non-intrusive standard.

[12] *Supra* note 5.

The FinCEN authorizes users, trains them, and monitors their use to ensure that the data, which are considered law enforcement sensitive, are properly used, disseminated, and kept secure. Access to CBRS data is based on an MOU with FinCEN. The MOU requires authorized personnel to make a best effort to limit BSA information queries to those which are immediately useful in connection with the specific matter prompting the query and to destroy all documents or summaries obtained in a timely manner. Access to the FinCEN data is limited to authorized persons with a need to know. SAR Review Team personnel use the FinCEN databases in a secure office environment with password controlled computer access and an effective audit capability. Prior to obtaining access, FinCEN requires users to undergo FinCEN training. Personally identifiable information collected by the SAR Review Team from FinCEN is maintained in the password-controlled computer.

The provisions of the Re-Dissemination Guidelines for Bank Secrecy Act Information govern re-dissemination of CBRS information. Dissemination of the data is limited to SAR Review Team members or other authorized persons with a need to know. In accordance with Bank Secrecy Laws, dissemination of the data within the SAR Review Team, or any other persons determined to have an authorized need to know, is documented. All SAR Review Team members receive and sign re-dissemination memorandums, which are kept on file with the IRS-CI. Copies are kept at the US Attorney's Office.

As for information gleaned from private information service companies, such as LexisNexis, privacy protection and data accuracy are staples of the business practices of these nationally known companies. These companies have privacy policies published on their web sites. The technology used to perform analysis is protected by the various security measures in place in the U.S. Attorney's Office for the Western District of New York.

3. <u>SAR Review, Western District of Michigan</u>

(1) **Description**

The Western District of Michigan (WDMI) SAR Review Team was created to identify and predict criminal activity through proactive analysis of U.S. Bank Secrecy Act data. The SAR Review Team operates under the direction and authority of the U.S. Attorney's Office in the Western District of Michigan. The primary goal of the WDMI SAR Review Team is to identify finance-related federal crimes. A second goal is to utilize SAR information related to the WDMI proactively to predict future activities of identified criminals and terrorists.

The WDMI SAR Review Team utilizes FinCEN data to detect, identify, track, and interdict people and organizations committing criminal acts and posing threats to the homeland. FinCEN administers the Bank Secrecy Act (BSA), which requires depository institutions and other industries vulnerable to money laundering to file and report certain data about financial transactions possibly indicative of money laundering and other criminal activity. FinCEN provides direct, electronic access to that data to qualifying law enforcement agencies through the

Currency and Bank Retrieval System (CBRS). Law Enforcement agencies acquire access to CBRS by signing a Memorandum of Understanding (MOU) with FinCEN.

The WDMI SAR Review Team consists of representatives from the U.S. Attorney's Office, Drug Enforcement Administration, FBI, Internal Revenue Service (IRS) Criminal Investigation (CI), Secret Service, Immigration and Customs Enforcement, U.S. Postal Inspection Service, and ATF. Upon review, additional FinCEN reports may be generated based on specific subject searches, transaction amounts, or other characteristics such as zip codes

The SAR Review Team was formed in 2004. This program is an ongoing initiative.

(2) **Technology and Methodology**

The proactive analysis begins by accessing BSA data through a secure connection to the Department of Treasury's FinCEN CBRS. The SAR Review Team queries the CBRS for data relating to the WDMI using specific criteria such as identified subjects, geographic locations, and information filers. Both null and positive results constitute BSA information subject to the FinCEN MOU. The WDMI SAR Review Team manually assesses each individual SAR for investigative merit and patterns of anomalies that may help predict future criminal activity.

(3) **Data Sources**

The SAR Review Team analyzes Unclassified, Law Enforcement Sensitive, and public information. The primary data source is CBRS, which includes BSA information filed in Currency Transaction Reports (CTRs), Foreign Bank Account Reports (FBARs), SARs, and Money Service Business (MSB) registration reports. These reports are filed by covered financial institutions pursuant to the BSA and implementing rules administered by FinCEN.

Secondary sources of data include well known private companies brokering public data, such as LexisNexis, and Internet search engines. When applicable, information and reports from other law enforcement agencies are also utilized.

(4) **Efficacy**

The principal source of data being analyzed is CBRS, which is administered by the Department of Treasury's FinCEN. Non-SAR BSA CBRS data is believed to be reliable. If errors are detected, correspondence letters are mailed to the filer asking for missing and incomplete information. Correspondence replies are posted to the CBRS upon receipt.

(5) **Privacy and Civil Liberties Impact**

Member agencies of the WDMI SAR Review Team have the authority and responsibility to investigate crimes and to conduct initiatives to detect, deter, and prevent criminal and terrorist activities. The SAR Review Team initiatives are grounded in traditional law enforcement techniques designed to discern patterns of criminal activity and to focus resources appropriately.

The WDMI SAR Review initiative is designed to accomplish these goals with greater efficiency and accuracy through the use of electronic data sources, such as CBRS and LexisNexis.

The rules of conduct pertaining to the use of BSA information in the CBRS are established by the Department of Treasury and enforced through FinCEN. Only authorized personnel may access the CBRS. Each authorized user is assigned a unique user identification and password. BSA Information may be utilized by authorized personnel to identify, investigate, or prosecute possible or actual violations of criminal law that fall within the investigative or prosecutorial jurisdiction of the Agency or Task Force.

In accordance with Department policies, the U.S Attorney's Office in the Western District of Michigan strives to mitigate potential privacy risks through compliance with Federal statutes and Departmental policies and regulations. Only positive information and leads are disseminated to the SAR Review Team. The information is reviewed for accuracy and completeness before dissemination. To protect privacy and civil liberties further, information derived through the public database searches is provided to agents and stored with the investigative file. In addition, the CBRS searches are conducted in government space with restricted access. Therefore, the impact or likely impact of this data mining initiative on privacy and civil liberties is estimated to be none to minimal.

(6) Law and Regulations

The Department of Justice complies with current laws and regulations regarding privacy, such as Privacy Act of 1974, FISMA, and FIPS published by the NIST, as well as Department of Justice Order 2640.1. Each of these sets forth requirements for securing agency information and IT systems. No investigative activity is initiated by the initiative unless the criteria established in the relevant Attorney General Guidelines are met. These criteria include the logical evaluation of lead information through other non-intrusive, lawful means. The Executive and Legislative branches have recognized data mining as a legitimate law enforcement analytic technique.

Other guidelines include:

- HSPDs 2, 6 and 11- all of which direct the strengthening of screening and analysis program to detect, identify and interdict individuals entering or within the United States who pose a terrorist threat to national security.

- National Security Presidential Directive 46 (War on Terror) also sets forth strengthening of terrorist screening tools as a major objective of national policy.

- Homeland Security Act of 2002, Section 201(e)(14); P.L. 107-296 (Nov. 25, 2002); Data Mining: An Overview, Congressional Research Services (Dec 16, 2004) recognized data mining as a legitimate law enforcement analytic technique.

(7) Privacy and Accuracy Protection Policies

The SAR Review Team does not automatically act upon search results generated from pattern-based data mining initiatives. The results are independently evaluated by U.S. Attorney and SAR Review Team personnel, which may include the SAR Coordinator or Paralegal. In accordance with Department procedures, these initiatives aim to investigate lead information by non-intrusive means. The Attorney General's Guidelines, as well as internal Departmental guidelines, set forth this non-intrusive standard.

The FinCEN authorizes users, trains them, and monitors their use to ensure that the data, which are considered law enforcement sensitive, are properly used, disseminated, and kept secure. Access to CBRS data is based on an MOU with FinCEN. The MOU requires authorized personnel to make a best effort to limit BSA information queries to those which are immediately useful in connection with the specific matter prompting the query and to destroy all documents or summaries obtained in a timely manner. Access to the FinCEN data is limited to authorized persons with a need to know. SAR Review Team personnel use the FinCEN databases in a secure office environment with password controlled computer access and an effective audit capability. Prior to obtaining access, FinCEN requires users to undergo FinCEN training. Personally identifiable information collected by the SAR Review Team from FinCEN is maintained in the password-controlled computer.

The provisions of the Re-Dissemination Guidelines for Bank Secrecy Act Information govern re-dissemination of CBRS information. Dissemination of the data is limited to SAR Review Team members or other authorized persons with a need to know. In accordance with Bank Secrecy Laws, dissemination of the data within the SAR Review Team, or any other persons determined to have an authorized need to know, is documented.

As for information gleaned from private information service companies, such as LexisNexis, privacy protection and data accuracy are staples of the business practices of these nationally known companies. These companies have privacy policies published on their web sites. The technology used to perform analysis is protected by the various security measures in place at the United States Attorney's Office for the Western District of Michigan.

III. Advanced Analytical Tools That Do Not Meet the Definition in Section 804

The Department of Justice has developed additional initiatives that do not meet the definition set forth in Section 804, but may be perceived as involving "data mining" based on some understandings of that term. Information as to some of these initiatives is provided below, in the interests of providing full and useful information. In addition, information on certain systems that have the capacity to allow advanced analysis is also included below. Where applicable, components have completed or are in the process of completing PIAs for these programs, and Privacy Act compliance issues have been addressed.

A. Drug Enforcement Administration (DEA) Initiatives

1. **Automation of Reports of Consolidated Orders System (ARCOS)**: Under applicable DEA regulations, manufacturers and distributors of Schedule I, II, or III narcotic

controlled substances must report the sale, purchase, loss, or inventory adjustment of these controlled substances to DEA. This data, which is collected in the ARCOS database, enables DEA to monitor the flow of these controlled substances from their point of manufacture through commercial distribution channels to point of sale or distribution at the dispensing and retail level. DEA reviews this data to ensure that purchase, sale, and other transaction reports match. It also reviews the data for suspicious activity, such as massive or recurrent losses of controlled substances. Such suspicious activity could lead DEA to investigate a target previously unknown to DEA. These reviews are both retrospective and subject-based in nature. Consequently, they do not satisfy Section 804's definition of "data mining."

2. **Drug Theft Loss (DTL) Database**: Similar to ARCOS reporting, DEA registrants at all levels (including practitioners and pharmacies) must report all losses of controlled substances to the DEA. This information is maintained in the DTL database. As with ARCOS, this database is reviewed for suspicious activities, which may lead DEA to investigate a previously unknown target. DTL does not qualify as "data mining" for the same reasons as ARCOS.

3. **SearchPoint**: SearchPoint was a DEA project which utilized information obtained commercially from ChoicePoint, a private data aggregator. ChoicePoint procured prescription data, both insurance and cash transactions. The information provided to DEA consisted of only filled prescriptions for controlled substances, including the prescribing official (practitioner); the dispensing agent (pharmacy, clinic, hospital, etc.); and the name and quantity of the controlled substance (drug information). No patient information was made available to DEA. The program ceased in 2008.

DEA utilized the SearchPoint database to conduct queries on practitioners, pharmacies, and controlled substances. The database enabled DEA to identify the volume and type of controlled substances a practitioner was prescribing or the volume and type of controlled substances a pharmacy was dispensing. For example, through the use of the SearchPoint database, DEA could quickly corroborate a complaint raised about a practitioner (*i.e.*, the practitioner prescribes only pain medications). Similarly, DEA could use SearchPoint to determine whether a pharmacy was operating as an Internet pharmacy by looking at indicators such as use of all cash transactions, only one or two drugs being dispensed, and the prescribing official being located in a different state than the pharmacy. DEA, utilizing the SearchPoint database, could also identify in which regions of the country sales of a particular type of controlled substance(s) were increasing in volume (*e.g.*, OxyContin).

With the SearchPoint database, DEA was able to identify current trends, prescribing and dispensing practices, and other patterns of activity, thus enabling the agency to identify probable anomalies outside of the normal prescribing practices, either locally or nationally. Using this tool, DEA was capable of quickly identifying potential violations of the Controlled Substances Act and could more effectively deploy its resources and manpower to those situations demanding the greatest and most urgent attention.

As SearchPoint queries were subject-based and as this initiative was designed to locate and solve past crimes, these efforts would not qualify as "data mining" for purposes of Section 804. As noted above, the program ceased in 2008.

4. **Online Investigative Project (OIP)**: OIP was a tool used to identify Internet pharmacies. This program enabled DEA to scan the Internet using search terms that might indicate the operation of an illegal Internet pharmacy (*e.g.*, "Vicodin," and "no prescription necessary"). Leads developed through the OIP could be further examined by investigative personnel to determine whether the website was, indeed, operating as an illegal Internet pharmacy. The OIP was an effort to identify targets through a search of databases using terms that could be, but not necessarily were, indicative of criminal activity. All OIP searches were conducted of "information publicly available to any member of the public without payment of a fee," which Congress specifically excluded from the definition of "database" in Section 804(b)(2). The program ceased in 2008.

B. Bureau of Alcohol, Tobacco, Firearms and Explosives (ATF) Initiatives

1. **Bomb Arson Tracking System (BATS)**: BATS is an Internet-accessible system that permits state, local, and other federal law enforcement agencies to share information related to bomb and arson investigations and incidents. ATF owns the BATS database, but each participating agency manages and controls its own information. The type of information queried via BATS is similarities of components, targets, or methods, and can be used, for example, to make connections between multiple incidents with the same suspect. As BATS is used to solve prior criminal events, it does not attempt to "discover or locate a predictive pattern or anomaly." Consequently, it does not satisfy Section 804's definition of data mining.

2. **GangNet**: GangNet is an Internet-accessible COTS system owned by ATF. GangNet tracks gang members, gangs, and gang incidents in a granular fashion and allows for sharing of this information across departments, agencies, states and regions. This system provides gang, gang members, and gang incident tracking and also provides for gang intelligence analysis to discern trends, relationships, patterns and demographics with respect to gangs. Again, as this program is retrospective and uses subject-based queries, it does not attempt to locate predictive patterns as required by Section 804.

C. Federal Bureau of Investigation (FBI) Initiatives

1. **Durable Medical Equipment (DME) Initiative**: This initiative was designed to help set investigative priorities for the FBI based on preliminary analysis of suspicious claims (submitted by DME suppliers) by contractors for the Centers for Medicare and Medicaid Services (CMS). These analyses, which identified DME suppliers engaged in the most egregious fraud and providers who have abnormal results from CMS billing audits, were provided to the FBI where they were analytically compared (using COTS software) by FBI analysts to FBI databases, as well as other complaints submitted to CMS and the Office of Inspector General for the U.S. Department of Health and Human Services. The results (analyses, provider lists and billing information) were forwarded to the relevant FBI field office for further investigation as

enclosures to an electronic communication that becomes part of FBI case files. In each case, the search was conducted in a manner that falls outside the scope of Section 804 because the queries were subject-based, rather than pattern-based. This activity ceased during the most recent reporting period.

The current DME Initiative focuses on the significant fraud problem that CMS, the National Health Care Anti-Fraud Association, and the FBI have identified with regard to DME providers. The goals of the initiative will be to supply information to field offices concerning suspect DME providers, to provide training on DME investigations, and to obtain media exposure regarding the FBI's investigation of DME fraud.

The primary source of information regarding DME suppliers are audits and investigations completed by the Supplier Audit and Compliance Unit (SACU), a contractor for CMS. Spreadsheets of information identifying DME suppliers operating under questionable accreditation or suspected of fraudulent activities against the Medicare program are forwarded to the FBI. The intelligence-gathering portion of the initiative includes the FBI completing analysis of those providers and providing associated referrals to the field with substantial predication to open a criminal case. In each case, the search is conducted in a manner that falls outside the scope of Section 804 because the analysis conducted by the FBI is subject-based, rather than pattern-based.

2. The **Investigative Data Warehouse (IDW)** is an FBI-managed system that enables investigators to search many FBI data sources across organizations within the FBI. The IDW has become a vital, robust analytic tool used by both analysts and agents within the FBI and uses data from more than 45 sources including the FBI, Department of State, and FinCEN. IDW users search data contained in intelligence reports, suspicious activity reports, watch lists, and FBI investigative files. The IDW provides capability for distributed search and presentation of integrated results to the agents and analysts that use its capabilities. Prior to the deployment of IDW, each of the sources of information would have to have been searched independently, which was inefficient.

By contrast, an IDW user today signs on to a single system and enters a search across the sources specified by the user with integrated search results provided to the user. The integration of such search results allows IDW users to examine the relationships efficiently between items of interest, including persons, places, communication devices, organizations, financial transactions, and case-related information across significantly larger amounts of data.

IDW is not pattern-based data mining within the meaning of Section 804 because it is not automated to conduct pattern based searches. Although there is access to databases such as ChoicePoint or Accurint through IDW, in order to query those particular commercial databases, the analyst must use specific subject based identifiers such as a name. The IDW is an ongoing initiative.

3. The **Identity Theft Intelligence Initiative** used Microsoft Excel, Microsoft Access, and Analyst Notebook I2 to extract consumer complaints from the Federal Trade Commission's

Identity Theft Clearinghouse into an FBI database to develop clusters of common identities, phone numbers, and e-mail addresses of subjects of complaints in a given geographical area. This initiative was used to generate leads for field offices to pursue beginning in late 2003. The activity described above ended sometime after November 2007 and prior to April 2009.

Responsibility for ID theft matters was transferred from the FBI's Criminal Investigative Division to the Cyber Division in late 2007. Within the Cyber Division, the Internet Crime Complaint Center (IC3) currently has responsibility for investigation of ID theft matters. IC3 does access FTC complaints of Internet fraud, but the nature of the activity has changed from the initiative described above. IC3 now provides the FTC with information about complaints of computer crime (including identity theft) received by IC3. In the course of investigating a specific complaint of identify theft, IC3 may search the FTC complaint database, but such a search is subject-based, using the name of a specific individual. The subject-based and retrospective nature of this program takes it outside of Section 804's definition of "data mining."

4. The **Health Care/Medicare Fraud Initiative** enables FBI analysts to research and investigate health care providers who may be continually over-billing Medicare for patient care. This technology was introduced in its present form in 2003. This initiative uses Microsoft Excel and Microsoft Access to examine Medicare summary billing records extracted from the Centers for Medicare and Medicaid Services (CMS), supported by the CMS Fraud Investigative Database and the National Health Care Anti-Fraud Association Special Investigative Resource and Intelligence System (private insurance data). As this program focuses on the "detection of fraud, waste, or abuse in a Government agency or program," it is exempted from Section 804. *See* Section 804(C)(i). This initiative formally ceased in approximately May 2008, when the Financial Crimes Intelligence Unit (FCIU) completed preparing spreadsheets, etc., of possible Medicare fraud activity for all 56 FBI field offices. However, upon request, FCIU still has the ability to search the data sources described above and provide a field office with an updated report.

5. The **Housing/Mortgage Fraud Initiative** uses public source data containing buyer, seller, lender, and broker identities and property addresses originally purchased by the FBI from ChoicePoint in order to uncover housing purchases that may constitute mortgage fraud. This initiative began in 1999. However, updated information continues to be provided to the FBI by LexisNexis (the successor to ChoicePoint) as new real estate transactions meeting the criteria take place.

Data purchased for this initiative contains real estate transactions in which properties were purchased and sold within a short-time period with a significant differential price ("property flipping"). This data is exported to a Microsoft Access database and includes buyer, seller, lender, address, and values. The information is available for access by field office analysts assigned to economic crime matters. The analyst reviews pertinent information from the database and may also check FBI databases to identify related transactions. These connections are researched and developed solely by an analyst, not by a program.

SARs are also analyzed by the FBI for evidence of mortgage fraud. After a specific company or individual has been identified, FBI Headquarters reviews SARs originating from the relevant geographic locations and integrates the information with data in internal FBI data sources, such as IDW or the property flip database. The data is then compiled into a package that is sent to the pertinent field office to assist the office in investigating the possible mortgage fraud. This activity is conducted pursuant to an open assessment of mortgage fraud activity.

FBI analytic activity with respect to mortgage fraud focuses on investigating and responding to past illegal activity. Consequently, its retrospective nature takes it out of Section 804's definition of "data mining."

6. The **Automobile Accident Insurance Fraud Initiative** was designed to identify and analyze information regarding possible staged automobile accident cases as well as other automobile insurance fraud schemes. The analysis is expected to reveal the national scope of staged accident frauds, identify the major perpetrators and organized groups, and identify multi-city clusters where the staged accidents are occurring. The goal is to use this initiative to target staged automobile accidents in major metropolitan areas throughout the United States. The initiative, which has begun to be deployed nationwide, is currently used in approximately 15 field offices. As this initiative responds to past suspicious activity, its retrospective nature takes it out of Section 804's definition of "data mining."

7. **Health Care/Prescription Fraud SearchPoint Initiative**: The FBI's Health Care Fraud Unit (HCFU), in conjunction with a limited number of field offices, utilizes SearchPoint, to analyze data regarding possible prescription drug fraud or pharmaceutical diversions. Analyses are completed on an as-needed basis to provide FBI field offices with intelligence regarding health care providers and pharmacies possibly engaging in prescription fraud or diversions of pharmaceuticals. The FBI first began using ChoicePoint for this initiative through an existing Department contract during Fiscal Year 2006, but now has a separate contract with SearchPoint effective in Fiscal Year 2008. This initiative is ongoing. As this program is retrospective in nature, it does not attempt to locate patterns of criminal activity. Consequently, it does not qualify as "data mining" for purposes of Section 804.

8. **COPLINK**: The FBI's Tampa (Florida) Division has access to, and analyzes data from, the COPLINK database operated by the Hillsborough County (Florida) Sheriff's Office (HCSO) to identify possible criminal or terrorist activity. COPLINK is COTS software used to warehouse information from approximately ten local police departments in the Tampa and surrounding area. HCSO received a grant from the U.S. Department of Homeland Security to establish this system. An MOU was signed between the Tampa Division and the HCSO to permit FBI access to COPLINK. Pursuant to that MOU, the Tampa Division began accessing COPLINK in October 2007. As this system is neither prospective nor pattern-based, it is not "data mining" for purposes of Section 804. This program is an ongoing initiative.

9. **Internet Pharmacy Fraud Initiative:** This initiative uses COTS, such as Microsoft Access and I2 Analyst Notebook, to search consumer complaints involving alleged fraud by Internet pharmacies. This initiative was created in December 2005 to identify and prosecute

licensed and unlicensed Internet pharmacies involved in the illegal distribution of diverted, counterfeit, or unapproved pharmaceuticals to consumers in the U.S. Although this program was initially intended to engage in data mining, it is currently used only as a de-confliction tool to determine whether a specific Internet pharmacy is the subject of either pending or closed criminal investigations. As this program does not prospectively attempt to locate patterns, it does not satisfy Section 804's definition of "data mining."

D. U.S. Attorneys' Offices Initiatives

Health Care Fraud

Several U.S. Attorneys' Offices conduct pattern-based queries for purposes of identifying and confirming health care-related criminal activity.

The U.S. Attorney's Office for the Western District of Michigan's (WDMI) Health Care Fraud Unit, working with Federal and State investigative agencies and private insurance companies, conducts pattern-based inquiries seeking to identify criminal conduct or confirm criminal conduct. Typically, the office or agency makes a request of the fiscal intermediaries or subpoena private insurance carriers who are responsible for maintaining such data. The queries involve data mining on certain procedures to identify potential targets of fraud based on extremely high or anomalous billing patterns. The queries are also used to further investigate identified targets to corroborate information or identity other types of fraud not yet reported. Information gathered by the participating agencies and the WDMI is retained by the agencies and the USAO for analysis. The participating agencies use the information to investigate and prevent crimes occurring in the WDMI and to prove the crimes and corresponding losses in any resulting cases. As this program focuses on the "detection of fraud, waste, or abuse in a Government agency or program," it is exempted from Section 804. *See* Section 804(C)(i).

The U.S. Attorney's Office for the Western District of North Carolina's (WDNC) Zone Program Integrity Contracts contractor, AdvanceMed, conducts pattern-based searches in support of the WDNC's Affirmative Civil Enforcement and criminal health care fraud units. AdvanceMed uses Medicare and Medicaid data to determine which providers are statistical outliers by virtue of the codes they bill and merit further investigation. As this program focuses on the "detection of fraud, waste, or abuse in a Government agency or program," it is exempted from Section 804. *See* Section 804(C)(i).

The U.S. Attorney's Office for the Western District of Oklahoma (WDOK) uses pattern-based data mining in health care fraud cases to develop and track billing trends. In addition, the WDOK used data mining in investigations involving an occupational rehabilitation company and a pharmaceutical company. In these investigations, the WDOK, employing a contractor, conducted pattern-based searches of Department of Labor claims and Medicare claims databases. As this program focuses on the "detection of fraud, waste, or abuse in a Government agency or program," it is exempted from Section 804. *See* Section 804(C)(i).

C. U.S. Department of Justice – Asset Forfeiture and Money Laundering Section

National SAR Review

In addition to local SAR review teams chaired by U.S. Attorneys' Offices, the National SAR Review Team meets on a monthly basis to review SARs filed with FinCEN that do not have a particular jurisdiction. The National SAR Review Team is led by the Department of Justice's Asset Forfeiture and Money Laundering Section, and consists of agents and investigators from the IRS, Federal Reserve, FBI, ATF, U.S. Secret Service, FinCEN, Domestic Security Section, and the Securities and Exchange Commission. The primary mission of the National SAR Review Team is to review a broad range of SAR classifications evidencing a significant international or multi-district nexus. This approach captures complex schemes that are multi-jurisdictional in nature and involve foreign nationals or international activity—activity that often overlaps with terrorism intelligence.

The National SAR Review Team does not use a predictive analysis model or attempt to identify future criminal activity. Accordingly, the activity of the National SAR Review Team does not come under the qualifying activity enumerated in Section 804.

F. Additional Department of Justice Systems

The Department also has systems or data warehouses that could be capable of supporting advanced analytic tools, but do not themselves fall within the requirements set forth in Section 804. Distinct technical and operational differences exist when comparing, on the one hand, a data warehouse that utilizes search tools and, on the other hand, a warehouse that is part of an initiative within the meaning of Section 804. Department law enforcement components employ numerous search tools and databases to help accomplish a variety of missions. Various groups collect data, others analyze data, and still others report data to Department law enforcement entities, as well as trusted federal, state, local and tribal law enforcement partners. These systems are used to save time and enable law enforcement properly and accurately to connect the dots, as prescribed by the 9/11 Commission, the Markle Foundation, and others.

The systems listed below are data systems with search and analytic tools used to conduct investigations, but they do not perform data mining, as defined in Section 804. In addition, several of the systems mentioned below have either completed or are in the process of completing PIAs. Because these systems are national security systems, a PIA is not required under the eGov Act; however, the Department still requires certain projects to complete PIAs as a risk mitigation step (although the PIA is not publicly available), and this policy is enforced by the Chief Information Officers in each component. Of course, the Department also ensures that it complies with the requirements of the Privacy Act where applicable to these systems.

1. The **Organized Crime and Drug Enforcement Task Force (OCDETF) Fusion Center** maintains a database named Compass that contains relevant drug and related financial intelligence information from numerous law enforcement organizations. The Department

centrally manages the group and its contributors in the OCDETF Fusion Center. These contributors include DEA, FBI, IRS, ICE, U.S. Marshals Service, Bureau of Prisons, FinCEN, Coast Guard, ATF, and the Department of Justice-Joint Automated Booking System.

The goal of the database is to use cross-case analysis tools to transform multi-agency information into actionable intelligence in order to support major investigations across the globe. As Compass is used to identify and solve past criminal activity, it does not satisfy the definition contained in Section 804.

2. **Internet Crime Complaint Center (IC3)** is a partnership between the FBI and the National White Collar Crime Center (NW3C). The mission of IC3 is to serve as a vehicle to receive, develop, and refer criminal complaints regarding the rapidly expanding arena of cybercrime. The IC3 gives the victims of cybercrime a convenient and easy-to-use reporting mechanism to provide authorities with tips on suspected criminal or civil violations. For law enforcement and regulatory agencies at the federal, state, local, and international level, IC3 provides a central referral mechanism for complaints involving Internet related crimes. The IC3 database contains complaints, submitted by the public, crossing the spectrum of cybercrime matters, to include online fraud in its many forms, including intellectual property rights matters, computer intrusions (hacking), economic espionage (theft of trade secrets), online extortion, international money laundering, identity theft, and a growing list of Internet-facilitated crimes. The FBI maintains the database with all of the cybercrime complaints, and if a complaint turns into a case, that information is loaded into the FBI's central case management system, ACS. As IC3 is merely a referral system, data is not mined under the definition in Section 804.

3. **Computer Analysis and Response Team (CART) Family of Systems (FOS)** include the tools needed to support computer forensics work across the country. CART maintains its own Storage Area Network to handle the large amount of data that it processes. The data obtained and stored is data covered under a valid search warrant, as a result of a criminal investigation. CART takes all data from the hard drive of a computer and makes an evidence-ready copy of the data. Advanced analytic tools are used to search the data on each system and to look for similarities across properly confiscated hard drives. Because these searches are retrospective and typically subject-based, the CART tools and capabilities do not meet the definition of data mining under Section 804.

IV. Conclusion

As set forth above, the Department of Justice takes very seriously its obligation to prevent terrorism and investigate criminal conduct using all available and lawful tools, while also respecting the privacy and civil liberties of Americans. The use of advanced analytic tools is extremely valuable and is only undertaken with due regard for the privacy concerns of individuals. The Department's use of advanced analytic tools meets these standards.