This report was obtained (via FOIA), OCRed, and posted by AltGov2

www.altgov2.org
Mr. Russ Kick  
P.O. Box 36914  
Tucson, AZ 85740-6914  
russkick@gmail.com  

Dear Mr. Kick:

This is our final response to your Freedom of Information Act (FOIA) request made via the Office of Information Policy FOIAonline Portal and received in this Office on July 16, 2016, in which you requested the two most recent reports submitted by the Department of Justice as required by Federal Agency Data Mining Reporting Act. This response is made on behalf of the Office of the Attorney General.

By letter dated August 15, 2016, we provided a copy of a previously-processed report and advised the remaining report you requested required consultations with other Offices. Those consultations are now complete, and I have determined that twenty-four pages of material are appropriate for release with an excision made pursuant to Exemption 7(E) of the FOIA, 5 U.S.C. § 552(b)(7)(E), which pertains to records or information compiled for law enforcement purposes, the release of which would disclose a technique or procedure for law enforcement investigations or prosecutions.

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) (2015) (amended 2016). 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.

You may contact our FOIA Public Liaison, Douglas Hibbard, for any further assistance and to discuss any aspect of your request at: Office of Information Policy, United States Department of Justice, Suite 11050, 1425 New York Avenue, NW, Washington, DC 20530-0001; telephone at 202-514-3642; or facsimile at 202-514-1009.

Additionally, you may contact the Office of Government Information Services (OGIS) at the National Archives and Records Administration to inquire about the FOIA mediation services they offer. The contact information for OGIS is as follows: Office of Government Information Services, National Archives and Records Administration, Room 2510, 8601 Adelphi Road, College Park, Maryland 20740-6001; e-mail at ogis@nara.gov; telephone at 202-741-5770; toll free at 1-877-684-6448; or facsimile at 202-741-5769.

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 OIP’s FOIAonline portal at [https://foiaonline.regulations.gov/foia/action/public/home](https://foiaonline.regulations.gov/foia/action/public/home). Your appeal must be postmarked or electronically submitted within ninety days of the date of my response to your request. If you submit your appeal by mail, both the letter and the envelope should be clearly marked “Freedom of Information Act Appeal.”

Sincerely,

[Signature]

Vanessa R. Brinkmann
Senior Counsel

Enclosure
Dear Mr. President:

We are pleased to transmit the unclassified portion of the report required by the Federal Agency Data Mining Reporting Act of 2007 for the period covering October 1, 2009 through September 30, 2012. This reporting requirement is contained within section 804 of the Implementing Recommendations of the 9/11 Commission Act of 2007, Public Law 110-53. We regret our delay in submitting this portion of the report.

We hope this information is helpful. Please do not hesitate to contact this office if we may be of additional assistance to you. The Office of Management and Budget has advised us that from the perspective of the Administration’s program, there is no objection to submission of this report.

Sincerely,

Peter J. Kadzik
Assistant Attorney General

Enclosure

IDENTICAL LETTERS SENT TO THE HONORABLE MITCH McCONNELL, MAJORITY LEADER, UNITED STATES SENATE; THE HONORABLE HARRY REID, MINORITY LEADER, UNITED STATES SENATE; THE HONORABLE CHARLES E. GRASSLEY, CHAIRMAN, COMMITTEE ON THE JUDICIARY, UNITED STATES SENATE; THE HONORABLE PATRICK J. LEAHY, RANKING MINORITY MEMBER, COMMITTEE ON THE JUDICIARY, UNITED STATES SENATE; THE HONORABLE RICHARD M. BURR, CHAIRMAN, SELECT COMMITTEE ON INTELLIGENCE, UNITED STATES SENATE; THE HONORABLE DIANNE FEINSTEIN, VICE CHAIRMAN, SELECT COMMITTEE ON INTELLIGENCE, UNITED STATES SENATE; THE HONORABLE JOHN A. BOEHNER,
SPAKER, HOUSE OF REPRESENTATIVES; THE HONORABLE KEVIN MccARTHY, MAJORITY LEADER, HOUSE OF REPRESENTATIVES; THE HONORABLE NANCY PELOSI, MINORITY LEADER, HOUSE OF REPRESENTATIVES; THE HONORABLE BOB GOODLATTE, CHAIRMAN, COMMITTEE ON THE JUDICIARY, HOUSE OF REPRESENTATIVES; THE HONORABLE JOHN CONYERS, JR., RANKING MINORITY MEMBER, COMMITTEE ON THE JUDICIARY, HOUSE OF REPRESENTATIVES; THE HONORABLE DEVIN NUNES, CHAIRMAN, PERMANENT SELECT COMMITTEE ON INTELLIGENCE, HOUSE OF REPRESENTATIVES; AND THE HONORABLE ADAM SCHIFF, RANKING MINORITY MEMBER, PERMANENT SELECT COMMITTEE ON INTELLIGENCE, HOUSE OF REPRESENTATIVES.
I. Legislative Requirements

The Federal Agency Data Mining Reporting Act of 2007\(^1\) 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.

The Act 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.\(^2\)

In construing the definition of "data mining" for purposes of this report, three elements of the Act’s definition warrant further discussion. First, the Act’s definition covers all electronic


\(^2\) Id. § 2000ee-3(b)(1).
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). As a result,
programs dedicated exclusively to solving past crimes or incidents do not fall within this
definition. Third, the Act 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.

For each initiative and program that meets the “data mining” statutory definition above,
the head of the agency must provide the following information:

1. A thorough description of the data mining activity and its goals.
2. The target dates, where appropriate, for the deployment of the data mining activity.
3. 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.
4. A thorough description of the data sources that are being or will be used.
5. 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.
6. 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.
7. 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.
8. 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

3 Id. § 2000ee-3(b)(1)(A); see also 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).
to: (a) protect the privacy and due process rights of individuals, such as redress procedures; and (b) ensure that only accurate and complete information is collected, reviewed, gathered, analyzed, or used, and guard against any harmful consequences of potential inaccuracies.4

This report discusses the Department’s qualifying programs conducted during the period from October 1, 2009, to September 30, 2012, in response to each of the requested pieces of information listed above and set forth by the Act. Specifically, the report discusses qualifying programs carried out during this period by the Federal Bureau of Investigation (FBI) and certain United States Attorneys’ Offices, as reported by the Executive Office for United States Attorneys (EOUSA). This report also provides information about certain advanced analytic activities conducted by the Department during this period that do not meet the definition of “data mining” set forth by the Act, but that may nonetheless be perceived as “data mining” based on some understandings of that term. Finally, the Department has submitted an annex that addresses data mining in connection with certain classified activities.

II. Protection of Privacy and Civil Liberties

The Department recognizes the importance of carefully considering privacy and civil liberties issues raised by programs that conduct data mining activities. Data mining activities that analyze lawfully acquired information, as is the case with each of the qualifying activities, can be extremely valuable tools for law enforcement, counterterrorism, and intelligence purposes. 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 and analysts conducting those tasks manually.

While the use of advanced analytics on large datasets can produce significant benefits, it can also pose certain privacy and civil liberties concerns if such activities are conducted without proper protections, controls, and careful decision-making processes in place. Privacy issues that can be implicated include concerns regarding the legal authority for the collection and use of the information, security risks of maintaining and accessing the information, and the potential misuse of, and inaccurate determinations made based on, information obtained through such data mining activities.

The Department’s programs that conduct data mining activities serve the Department’s important law enforcement and national security missions and are undertaken in a manner consistent with policies and procedures that promote and preserve fundamental privacy and civil liberties principles. This section discusses the Department’s privacy compliance process, which is designed to identify and mitigate the risks to privacy that may arise in the development and implementation of a Department system, program, or operation, including those that involve data mining activities. Specific privacy and civil liberties assessments on the programs included in this report are discussed below in Part III and in the classified annex.

The Department's Chief Privacy and Civil Liberties Officer (CPCLO) and the Office of Privacy and Civil Liberties (OPCL) were established to oversee the Department's privacy compliance process and to ensure that appropriate privacy and civil liberties protections are incorporated into the Department's systems, programs, and operations. The CPCLO position was created by Section 1174 of the Violence Against Women and Department of Justice Reauthorization Act of 2005 and has additional statutory responsibilities under Section 803 of the Implementing Recommendations of the 9/11 Commission Act. The CPCLO's duties include: advising on privacy policy with regard to the Department's collection, storage, use, and disclosure of personally identifiable information; advising Department leadership on privacy issues when the Department proposes, develops, and implements laws, regulations, policies, procedures, or guidelines related to its national security efforts; and ensuring the Department complies with privacy-related laws and policies. The CPCLO and OPCL also work closely with the privacy officials within each component to identify and review potential privacy issues and concerns at the outset of a program.

As the first step of its privacy compliance process, the Department established the Initial Privacy Assessment (IPA) template, which consolidates various privacy compliance requirements into a single, unified, and comprehensive process. The IPA is a tool designed to assist Department components identify privacy issues during the development of an information system in order to help resolve and mitigate any such concerns, and also helps Department components determine whether additional privacy documentation is required by either the Privacy Act of 1974 (e.g., a System of Records Notice (SORN)) or Section 208 of the E-Government Act of 2002 (e.g., a Privacy Impact Assessment (PIA)). The Department has incorporated the IPA process into its information technology (IT) security framework, and IPAs must be completed for all new or modified Department information systems and programs that contain personally identifiable information, including those that involve data mining activities. The use of IPAs in the 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, required by Section 208 of the E-Government Act of 2002, address the existing authorities permitting the collection and advanced analysis of information. PIAs provide an


analysis of how information is handled to ensure compliance with applicable legal, regulatory, and policy requirements regarding privacy; to determine the risks and effects of collecting, maintaining, and disseminating information in identifiable form via an electronic information system; and to examine and evaluate protections and alternative processes for handling information to mitigate potential privacy risks. 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). The Department’s PIA template includes a specific inquiry about certain data mining activities, providing greater insight, analysis, and description of mitigation procedures that have been put in place for such activities. As with the IPA, PIAs are an integral part of the Department’s privacy compliance process and have been incorporated into the Department’s IT security framework, which ensures that all IT systems that require PIAs are identified and allows the Department to resolve or mitigate privacy risks they may pose.

Moreover, the Department has long been subject to the Privacy Act of 1974. Through the privacy compliance process, components work internally and with OPCL to review and resolve Privacy Act issues as they arise. The Privacy Act’s requirements generally apply to records that identify and are about U.S. citizens and lawful permanent resident aliens, and that are retrieved from a system by reference to the 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 Privacy Act’s requirements. Although exemptions exist for law enforcement purposes, among these requirements are that: (1) 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”; (2) 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

---


14 See id. § 552a(a) (defining, *inter alia*, “record,” “system of records,” and “individual”).

15 Id. § 552a(e)(1).
identifier,\(^{16}\) (3) 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”;\(^{17}\) and (4) 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.”\(^{18}\) Moreover, at the very core of the Privacy Act is its prohibition on disclosure of Privacy Act-protected information absent written consent or under certain statutorily defined circumstances.\(^{19}\)

In addition to the privacy compliance process, internal Department and FBI procedures, including The Attorney General’s Guidelines for Domestic FBI Operations (AGG-DOM),\(^{20}\) set forth the Department’s general policies regarding how investigations should be conducted and provide further guidelines on how data about individuals should be treated in the investigative process. For example, in certain initiatives described below, personally identifiable information is not forwarded to FBI investigators unless it is necessary for opening an investigation pursuant to the AGG-DOM. Limiting access to personally identifiable information in this manner substantially reduces the risk of misuse of information. 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, criminal investigations must still be predicated on a proper basis, and investigators still must have sufficient probable cause in order to obtain a warrant or arrest an individual. In each data mining initiative described in this report and its classified annex, no adverse actions are taken against individuals solely on the basis of data mining results.

The Department recognizes that there are privacy risks inherent in the use of pattern-based data mining initiatives. The Department strives to mitigate these potential risks through compliance with federal statutes and internal policies and regulations. Adherence to such mitigation procedures ensures that the collection, maintenance, and use of information about individuals by the Department in its data mining activities incorporates appropriate privacy and civil liberties protections while advancing its fundamental law enforcement and national security missions.

\(^{16}\) Id. § 552a(e)(4).

\(^{17}\) Id. § 552a(e)(5).

\(^{18}\) Id. § 552a(e)(10).

\(^{19}\) Id. § 552a(b).

III. Description of Department of Justice Data Mining Programs

This report discusses qualifying programs conducted during the period from October 1, 2009, to September 30, 2012. Specifically, the report discusses the analysis of Financial Crimes Enforcement Network (FinCEN) Suspicious Activity Reports (SARs) by the FBI and certain U.S. Attorneys' Offices. Other initiatives meeting the data mining definition are discussed in the classified annex to this report.

This report also includes descriptions of certain advanced analytic activities and programs conducted by the Department that do not fall within the statutory definition of “data mining,” but which may be perceived as involving data mining based on some understandings of that term. These activities and programs include:

Drug Enforcement Administration
1. Automation of Reports and Consolidated Orders System (ARCOS)
2. Drug Theft Loss Database (DTL)

Bureau of Alcohol, Tobacco, Firearms, and Explosives
1. Bomb Arson Tracking System (BATS)
2. GangNet
3. Semantica

Federal Bureau of Investigation
1. Durable Medical Equipment Initiative (DME)
2. Data Integration and Visualization System (DIVS)
3. Identity Theft Intelligence Initiative
4. Health Care/Prescription Fraud SearchPoint Initiative
5. Housing/Mortgage Fraud Initiative
6. Automobile Accident Insurance Fraud Initiative
7. COPLINK
8. Internet Pharmacy Fraud Initiative

United States Attorneys' Offices
1. Health Care Fraud

U.S. Marshals Service

1. Firearm Incident Risk Examination (FIRE)

These activities and programs, many of which were discussed in the Department's previous data mining reports to Congress pursuant to the Act, are summarized in Part IV, infra. Finally, this report discusses various other systems or data warehouses supporting advanced analytic tools that do not meet the "data mining" statutory definition.

A. Analysis of FinCEN Suspicious Activity Reports (SARs)

1. Description

The Currency and Foreign Transaction and Reporting Act of 1970, as amended by Title III of the USA PATRIOT Act of 2001 and other legislation, is commonly referred to as the Bank Secrecy Act (BSA). The BSA requires U.S. financial institutions to assist federal agencies in detecting and preventing criminal activity by maintaining records of cash purchases of negotiable instruments; filing reports of cash transactions exceeding $10,000; and reporting suspicious activity that might indicate money laundering, tax evasion, or other criminal activity. Treasury Department regulations implementing the BSA require financial institutions to file such Suspicious Activity Reports (SARs) with the Financial Crimes Enforcement Network (FinCEN). Entities filing such reports fall into several general categories, such as deposit institutions (e.g., banks and credit unions), casinos, money services businesses (e.g., check cashing and wire transfer companies), and businesses involved in the securities and futures industries.

The FBI analyzes and conducts pattern-based queries involving FinCEN SARs for several purposes. One purpose is to develop information regarding transactions that may involve funds derived from illegal activities, that are designed to evade BSA regulations, that have no apparent business or other lawful purpose, or that are not the sort of transactions typically engaged in by the customer. Various patterns located in SAR data may indicate possible criminal activity, such as the frequency with which an individual appears in connection with a reported transaction, her association with a particular business named in multiple SARs, or her familiarity with its particular business practices. Similarly, data contained in SARs may assist in placing a subject in a particular geographical location at a certain time, establish that a subject has regular contact with a specific financial establishment, or reveal inconsistent representations.

---


22 See id. §§ 5313(a), 5316(a), 5318(g).

23 31 C.F.R. Part 1010, subpart C (2013). Financial institutions may also voluntarily report suspicious transactions that are not otherwise required to be reported. Banks, for example, may voluntarily report suspicious activity to FinCEN pursuant to 31 C.F.R. § 1020.320(a) (2013).
made by an individual on different SARs. Such transactions may trigger a new investigation into certain financial-related crimes, such as money laundering or terrorism financing. Information contained in SARs may also be used to identify individuals as potential human sources. Finally, SAR-related activity is cross-referenced with pending FBI investigations in order to de-conflict investigations into the same activity.

In various jurisdictions, a working group of law enforcement agencies chaired by the local U.S. Attorney’s Office meets on a regular basis to review FinCEN SARs pertaining to that specific jurisdiction. These SAR review teams may consist of representatives from the U.S. Attorney’s Office, Drug Enforcement Administration (DEA), FBI, Internal Revenue Service (IRS) Criminal Investigation, Secret Service, Immigration and Customs Enforcement (ICE), U.S. Postal Inspection Service, U.S. Department of Housing and Urban Development (HUD), Department of Defense, Bureau of Alcohol, Tobacco, Firearms and Explosives (ATF), the Joint Terrorism Task Force (JTTF), state police and other state-level agencies, local law enforcement, and, when appropriate, accredited foreign partners. During these meetings, the SARs are reviewed to identify possible criminal activity, and additional reports may be generated, based on specific subject searches, transaction amounts, or other characteristics, such as zip codes. 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 twelve months. These SARs are then disseminated to working group members for appropriate follow-up. Although the FBI’s management of SARs generally meets the definition of “data mining,” the activities of these individual working groups do not generally qualify as data mining, as they are typically retrospective in nature.

2. Technology and Methodology

The FBI has four methods of accessing electronic SAR data. First, the FBI uses the Financial Institution Fraud (FIF) database, an application operating on the Bureau’s investigative mainframe. The FIF serves as an internal central repository for SARs filed by federally insured financial institutions, as well as other entities such as mortgage companies. The SARs are forwarded by FinCEN to the FBI and then uploaded into the FIF application by Criminal Investigative Division (CID) personnel at FBI headquarters. The data is then available through the FBI’s secure computer system, the SECRET Enclave, for use by field offices investigating money laundering or bank fraud violations. SAR data applicable to a specific geographic location, such as a state, is extracted from the FIF and then cross-referenced with information contained in other FBI databases, such as the Data Integration and Visualization System (DIVS), the National Crime Information Center (NCIC) database, and the Telephone Application (TA) database, which contains telephone subscriber data acquired by the FBI through statutorily authorized investigative tools.

A second method of accessing FinCEN SAR data is through a secure Treasury Department Internet portal, using the FinCEN Query tool, which will replace the Currency Banking Retrieval System (known as WebCBRS). The FBI and various U.S. Attorneys’ Offices are among the many users of FinCEN Query, along with state, local, and other federal law enforcement and regulatory agencies. This system provides authorized users the ability to query
and download BSA reports for specific purposes, including criminal, tax, regulatory, and intelligence and counterintelligence activities to protect against international terrorism. FinCEN Query allows users to easily access, query, and analyze more than eleven years of FinCEN data; apply filters and narrow search results; and import lists of data, such as names, identification numbers, and addresses. These capabilities were not available using WebCBRS. In addition, FinCEN Query includes an audit tool allowing FinCEN to audit agencies’ use of this data.

Third, the FBI accesses FinCEN SAR data through its Data Integration and Visualization System (DIVS), which incorporates the capabilities of the former Investigative Data Warehouse (IDW). The DIVS is a data-integration capability sitting on top of FBI data repositories and provides information to FBI personnel based on user-defined queries. DIVS provides the capability to search multiple data sources with a single query, so that personnel spend more time performing high value tasks rather than conducting multiple data searches across data systems that are not interconnected. Through an agreement with FinCEN, SAR data is available in several datasets accessed through DIVS. Using DIVS provides more flexible search options than are available solely through the FinCEN Portal, enabling personnel to analyze data at a more granular level.

Finally, the FBI accesses FinCEN SAR data through its Enterprise Palantir system, which accesses the same SAR data as DIVS. This system is a knowledge management application, accessed from the SECRET Enclave, which combines multiple capabilities needed for the effective evaluation of intelligence, including search, analysis, knowledge management, and collaboration. This system allows users to import large amounts of data and organize it visually to facilitate analysis, such as depicting data in charts or graphs or depicting links between data. Users in this system use keyword and parametric filters to identify SARs related to counterterrorism, counterintelligence, and cyber programs, as well as complex financial crimes.

After initial analysis of SAR data, analysts may then cross-reference the SAR data with information contained in various Internet-based public- and private-source databases. Such public-source databases may include Google, Yahoo, and government databases of corporate filings, property tax records, and similar sources of information. Private data sources include commercial data aggregators such as LexisNexis—Accurint, Thomson Reuters, and Dun & Bradstreet.

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

FinCEN SARs are the principal source of data analyzed in this activity. The data in SARs may also be compared to information contained in various databases, including FBI
databases accessed through DIVS, ACS, and NCIC; Internet-based private databases such as LexisNexis and Dun & Bradstreet; and public-source databases such as Google and Yahoo, as well as databases of public government records, such as corporate filings.

4. Efficacy

The financial transaction information contained in FinCEN SARs has proven to be an effective indicator of potential financial crimes such as money laundering. Information in SARs may also be an indicator of financial irregularities related to national security threats, such as terrorist financing. The FBI also reviews SARs in order to identify links between recent SARs and pending and past FBI investigations, to cross-reference information from SARs with information contained in FBI databases such as DIVS, and to identify and de-conflict pending FBI investigations.

The information contained in SARs is generally deemed reliable for purposes of initiating an investigation since SARs are a required regulatory filing, pursuant to the BSA, and, in the case of banks, under Federal bank supervisory agency general rule-making authority contained in Title 12 of the United States Code. Moreover, although 31 U.S.C. § 5318(g)(3) essentially immunizes reporting financial institutions from civil liability for submitting SARs, the safe harbor provision requires a “good faith suspicion” that a law or regulation has been violated. As a result, financial institutions have both a regulatory and civil liability interest in ensuring that SARs contain accurate information.

5. Law and Regulations

The FBI’s legal authority to conduct such investigations derives from the following statutes, regulations, and Department of Justice guidelines:

- 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 for Domestic FBI Operations (AGG-DOM)

---

24 The Automated Case Support (ACS) system is the legacy FBI case management system. Some data sets that could not be migrated to Sentinel, the FBI’s current electronic information and case management system, remain available through ACS, and Sentinel data compared to SARs is accessed through DIVS.


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

- the AGG-DOM also authorizes the FBI to conduct investigations to obtain information about 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.

6. Privacy and Civil Liberties Impact and Protections

The BSA requires certain financial institutions to report to FinCEN suspicious transactions which may indicate money laundering or other criminal activity. The SARs reporting such activity include certain personally identifiable information about individuals who are the subject of the report, such as their name, address, Social Security Number, and date of birth.

To help mitigate the privacy and civil liberties concerns inherent in the collection and review of such information by law enforcement agencies, there are established privacy protections governing the collection and use of information within FinCEN. The FBI treats SARs as law enforcement sensitive information, and 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 under investigation, as well as the content of individual SARs, is not normally disclosed to persons outside the law enforcement community. Access to the reports is restricted, and SARs are also exempt from public access under the Freedom of Information Act. Once the SARs are reported to FinCEN, the personally identifiable information that is maintained in the system is protected under Department of Treasury privacy rules. These rules and protections are described in the Department of Treasury’s Privacy Act system of records notice, regulations, and PIAs for FinCEN.

Access to FinCEN reports by the FBI and other Department components is governed by the disclosure rules and restrictions that apply to the Department of Treasury and each accessing entity. The Department of Treasury and each Department of Justice component accessing FinCEN have entered into memoranda of understanding to ensure that protections and safeguards are delineated and understood before any access to the information is permitted. For example, as


28 See id. § 1010.960.

indicated in the PIA for FinCEN, all external users, including the FBI, must provide a justification for each query of the system. As mentioned above, the FBI accesses FinCEN data through certain FBI systems, and the information collected on individuals for whom suspicious financial activity is indicated is entered into the Sentinel system. The Sentinel system is part of the FBI Central Records System, which is covered by and thus afforded the protections of the Privacy Act. In addition, the FBI has conducted PIAs or is in the process of updating PIAs for each FBI system that accesses or maintains FinCEN data. These PIAs provide evaluations and analyses of the information collected and used in these systems and help ensure that the FBI applies appropriate privacy protections and safeguards to such systems.

In addition, all FBI investigative and intelligence activities must comply with the AGG-DOM and the FBI's Domestic Investigations and Operations Guide (DIOG), both of which were intended, in significant part, "to ensure that the FBI conducts its investigations and other activities in a lawful and reasonable manner that respects liberty and privacy and avoids unnecessary intrusions into the laws of law-abiding people." AGG-DOM at 5. For example, part I.C.3 of the AGG-DOM provides:

All activities under these Guidelines must have a valid purpose consistent with these Guidelines, and must be carried out in conformity with the Constitution, and all applicable statutes, executive orders, Department of Justice regulations and policies, and Attorney General Guidelines. These Guidelines do not authorize investigating or collecting or maintaining information on United States persons solely for the purpose of monitoring activities protected by the First Amendment or the lawful exercise of other rights secured by the Constitution or laws of the United States. These Guidelines also do not authorize any conduct prohibited by the Guidance Regarding the Use of Race by Federal Law Enforcement Agencies [issued by the Attorney General].

In addition, section 4.1.1 of the DIOG provides:

The FBI is responsible for protecting the security of our nation and its people from crime and terrorism while maintaining rigorous obedience to the Constitution. The Attorney General's Guidelines for Domestic FBI Activities (AGG-DOM) establish a set of basic principles that serve as the foundation for all FBI mission-related activities. When these principles are applied, they demonstrate respect for civil liberties and privacy as well as adherence to the Constitution and laws of the United States. These principles are as follows:

A) Protecting the public includes protecting their rights and liberties. FBI investigative activity is premised upon the fundamental duty of government to protect the public, which must be performed with care to protect individual rights and to ensure that investigations are confined to matters of legitimate government interest.

B) Only investigate for a proper purpose. All FBI investigative activity must have an authorized law enforcement, national security, or foreign intelligence purpose.

C) Race, ethnicity, religion, or national origin alone can never constitute the sole basis for initiating investigative activity. Although these characteristics may be taken into account under certain circumstances, there must be an independent authorized law enforcement or national security purpose for initiating investigative activity.

D) Only perform authorized activities in pursuit of investigative objectives. Authorized activities conducted as part of a lawful assessment or investigation include the ability to: collect criminal and national security information, as well as foreign intelligence; provide investigative assistance to federal, state, local, tribal, and foreign agencies; conduct intelligence analysis and planning; and retain and share information.

E) Employ the least intrusive means that do not otherwise compromise FBI operations. Assuming a lawful intelligence or evidence collection objective, i.e., an authorized purpose, strongly consider the method (technique) employed to achieve that objective that is the least intrusive available (particularly if there is the potential to interfere with protected speech and association, damage someone’s reputation, intrude on privacy, or interfere with the sovereignty of foreign governments) while still being operationally sound and effective.

F) Apply best judgment to the circumstances at hand to select the most appropriate investigative means to achieve the investigative goal. The choice of which investigative method to employ is a matter of judgment, but the FBI must not hesitate to use any lawful method consistent with the AGG-DOM when the degree of intrusiveness is warranted in light of the seriousness of the matter concerned.

Pursuant to the AGG-DOM and the DIOG, no investigative activity is initiated by an FBI field office against any individual identified in SARs unless specific criteria are met. These criteria include the logical evaluation of lead information through other non-intrusive, lawful means, such as FBI record checks, private lender record checks, and developed sources. In addition, the use of more intrusive techniques, such as grand jury subpoenas, administrative subpoenas, tasking of sources, undercover operations, and electronic surveillance, is regulated by law and procedures designed to ensure that the techniques are lawfully and appropriately employed.

All FBI personnel receive extensive training on the need to protect personally identifiable information. All FBI employees must complete a ninety minute Virtual Academy course, Privacy: It’s Every Employee’s Business, within three months of entering on duty or reporting to their first duty station after graduation from the FBI Academy. The course provides basic information on the roles and responsibilities of FBI employees to protect personally identifiable information contained in FBI records. In addition, all personnel with access to FBI IT systems – virtually the entire workforce – must complete annual information security (INFOSEC) training, which includes material on the Privacy Act and the proper handling of material containing
personally identifiable information. Failure to complete the INFOSEC training in a timely manner results in suspension of access to IT systems.

Finally, to the extent that personally identifiable information contained in SARs is integrated into any Department of Justice investigative files, such information is subject to the full panoply of privacy protections applicable to the Department of Justice. These include the Privacy Act of 1974, the Federal Information Security Management Act (FISMA), Federal Information Processing Standards (FIPS) published by the National Institute of Standards and Technology (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. These statutes and policies set forth requirements for securing and protecting agency information and IT systems, including those involving data mining.

\[^{31}\text{44 U.S.C. §§ 3541 et seq. (2012).}\]
IV. Advanced Analytical Tools That Do Not Meet the Definition in the Act

The Department of Justice has developed additional initiatives that do not meet the definition of “data mining” set forth in the Act, but that may be perceived as such based on other understandings of that term. In the interests of providing full and useful information, descriptions of some key initiatives are provided below. In addition, information about certain systems that have the capacity to perform other forms of advanced analysis is also included below. Where applicable, components have completed or are conducting PIAs for these programs, and Privacy Act compliance issues have been or are being addressed.

A. Drug Enforcement Administration (DEA) Initiatives

1. **Automation of Reports and Consolidated Orders System (ARCOS):** Under applicable DEA regulations, manufacturers and distributors of Schedule I, II, or III narcotic controlled substances must report to the DEA the sale, purchase, loss, or inventory adjustment of these controlled substances. This data, which is collected in the ARCOS database, enables the DEA to monitor trends in 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. The 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 the DEA to investigate a target previously unknown to the DEA. These reviews are both retrospective and subject-based in nature. Consequently, they do not satisfy the Act’s definition of “data mining.”

2. **Drug Theft Loss Database (DTL):** Similar to ARCOS reporting, DEA registrants at all levels, including practitioners and pharmacies, must report to the DEA all thefts and significant losses of controlled substances. This information is maintained in the DTL database. As with ARCOS, this database is reviewed for suspicious activities, which may lead the DEA to investigate a previously unknown target. These reviews are both retrospective and subject-based in nature. Consequently, they do not satisfy the Act’s definition of “data mining.”

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. The ATF owns and hosts the BATS database, but each participating agency manages and controls its own information. BATS is used to identify similar components, targets, or methods across investigations and incidents. BATS is used to solve prior criminal events and does not attempt to “discover or locate a predictive pattern or anomaly.” Thus, the queries conducted in BATS do not satisfy the Act’s definition of “data mining.”

2. **GangNet:** GangNet is an Internet-accessible commercial-off-the-shelf (COTS) system owned by ATF. GangNet tracks gang members, gangs, and gang incidents in a manner that 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 gang demographics. As GangNet is retrospective in nature and uses subject-based queries, it is outside the scope of the Act.

3. **Semantica**: Semantica allows users to integrate data from disparate ATF data sources and systems into a single application, and to display links between items of interest, such as entities, persons, addresses, Social Security Numbers, telephone numbers, and email addresses. The user can view this data in various ways, including as a network, histogram, by entity, or plotted on a map using geo-coding technology. Semantica streamlines data collection and allows the user to quickly identify leads within and external to their investigations and networks of interest. Semantica does not conduct pattern-based queries, and therefore falls outside the scope of the Act.

### C. Federal Bureau of Investigation (FBI) Initiatives

1. **Durable Medical Equipment (DME) Initiative**: The DME initiative discussed in the Department’s previous reports ceased prior to FY 2012. Present FBI investigative efforts focus on the significant fraud problem that the Centers for Medicaid and Medicare Services (CMS), the National Health Care Anti-Fraud Association, and the FBI have identified regarding DME providers. The goals of the initiative are to supply information to FBI 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 sources of information regarding DME suppliers are audits and investigations completed by the Supplier Audit and Compliance Unit (SACU), a contractor for CMS. Information identifying DME suppliers operating under questionable accreditation or suspected of fraudulent activities against the Medicare program is forwarded to the FBI. The FBI thoroughly evaluates these providers and refers those with sufficient predication to open a criminal case to field offices. The queries and analysis conducted during the course of an investigation of Medicare or Medicaid fraud are excluded from the Act’s definition of “data mining” because they are aimed at “the detection of fraud, waste, or abuse in a Government ... program.”\(^\text{32}\) In addition, the queries and analysis regarding DME suppliers are subject-based, rather than pattern-based, and are therefore outside the scope of the Act for that reason as well.

2. **Data Integration and Visualization System (DIVS)**: The FBI’s Investigative Data Warehouse (IDW) was discussed in the Department’s previous reports. The IDW’s capabilities were merged into the Data Integration and Visualization System (DIVS) and IDW was retired in December 2012, creating a single Bureau-wide investigative environment.

   DIVS is a vital investigative tool and is now the FBI’s second-most-used IT capability behind Sentinel, the Bureau’s case management system. DIVS includes data from more than 50

---

data repositories and makes all FBI intelligence and investigative data accessible and searchable by FBI agents and analysts through a single, secure user interface.

Users can query across multiple data repositories within DIVS, filter results, and quickly obtain integrated search results. The integration of these search results allows efficient examination of relationships between items of interest, such as persons, places, communications, communication devices, organizations, financial transactions, and case-related information.

The FBI’s use of DIVS does not constitute pattern-based data mining under the Act because these queries must include one or more subject-based identifiers.

3. **Identity Theft Intelligence Initiative**: The Identity Theft Intelligence Initiative, previously discussed in the FY 2009 report, concluded prior to the end of CY 2009. This activity involved subject-based analysis outside the scope of the Act’s definition of “data mining.”

The FBI’s Internet Crime Complaint Center (IC3) continues to receive and refer complaints related to identity theft for possible investigation. The IC3 also provides the Federal Trade Commission (FTC) with information about complaints received alleging computer crime, including identity theft. In the course of investigating a specific complaint of identity theft, IC3 personnel may search the FTC complaint database, but these searches are subject-based, and therefore outside the scope of the Act’s definition of “data mining.”

4. **Health Care/Prescription Fraud SearchPoint Initiative**: The Health Care/Prescription Fraud SearchPoint Initiative, described in previous reports, is continuing. The FBI’s Health Care Fraud Unit (HCFU), in conjunction with a limited number of field offices, continues to utilize SearchPoint to analyze transaction data regarding possible prescription drug fraud or pharmaceutical diversions. Analyses are prepared on an as-needed basis to provide FBI field offices with information regarding health care providers and pharmacies possibly engaging in prescription fraud or diversions of pharmaceuticals. As this program is retrospective in nature, the queries are not intended to identify or utilize patterns predicting future criminal or terrorist activity. The activity therefore does not constitute “data mining” within the meaning of the Act.

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

The real estate transaction data analyzed identifies properties purchased and resold within a short period of time and with a significant difference between the selling prices. 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 then researched and developed
by an FBI analyst, and are not generated by an automated program.

FBI analytic activity with respect to mortgage fraud focuses on investigating and
responding to past illegal activity rather than attempting to utilize a pattern to identify and
predict which individuals might commit illegal acts in the future. The retrospective nature of this
initiative places it outside the scope of the Act’s definition of “data mining.”

6. **Automobile Accident Insurance Fraud Initiative:** The Automobile Accident
Insurance Fraud Initiative, described in previous reports, was discontinued in 2011, although the
FBI continues to investigate staged automobile accident cases, as well as other automobile
insurance fraud schemes. Since this initiative targeted crime that had already occurred, its
retrospective nature placed it outside of the Act’s definition of “data mining.”

7. **COPLINK:** Previous reports discussed COPLINK, a program under which the FBI’s
Tampa field office accesses a regional network containing law enforcement information from
multiple local law enforcement agencies. Numerous FBI field offices have access to similar
localized collections of law enforcement information.

Information in localized law enforcement networks such as COPLINK may be available
on a national level, for purposes of criminal and national security investigations, in the FBI’s
National Data Exchange (N-DEx), operated by the FBI’s Criminal Justice Information Services
(CJIS) Division. The N-DEx is a national information-sharing system available to law
enforcement agencies (including the FBI), through a secure Internet connection, permitting
queries and analysis of data spanning the entire criminal justice cycle, including crime reports
and investigative records, arrest, booking and incarceration records, and probation and parole
records. As a central repository of information contributed by federal, state, regional, local, and
tribal criminal justice entities, the N-DEx supports queries to quickly identify investigative links
between crimes, criminal investigations, and related events contained in that information. The
N-DEx contains more than 150 million records contributed by more than 4,000 criminal justice
agencies throughout the United States.

Users may query N-DEx to identify correlations between investigative information and
information already contained in N-DEx. These correlations are based on matches to personal
identifiers such as names, Social Security Numbers, and driver’s license numbers, rather than
matches to non-specific crime data such as race, gender, or nationality. Such predicated queries,
relying on subject-based identifiers, are outside the scope of “data mining” as defined in the Act.

8. **Internet Pharmacy Fraud Initiative:** The Internet Pharmacy Fraud Initiative,
described in previous reports, was discontinued in April 2011, although the FBI continues to
investigate allegations of fraudulent activity by Internet pharmacies. As discussed in previous
reports, the program was a deconfliction tool to determine whether a specific Internet pharmacy
is the subject of a criminal investigation, and therefore was outside the scope of the Act’s
definition of “data mining.”

-19-
D. U.S. Attorneys’ Offices (USAO) Initiatives

1. Health Care Fraud

Several U.S. Attorneys’ Offices conduct pattern-based queries for purposes of identifying and confirming health care-related criminal activity. These queries typically seek to identify criminal conduct based on anomalous billing patterns. As these programs focus on the “detection of fraud, waste, or abuse in a Government agency or program,” the Act, they are exempted from the Act.

E. U.S. Marshals Service (USMS) Initiatives

1. Firearm Incident Risk Examination (FIRE): The Behavioral Analysis Unit’s (BAU) FIRE program began in March 2011. This program uses data from the National Incident-Based Reporting System (NIBRS), USMS Justice Detainee Information System (JDIS), and internal critical incident reports to analytically examine shooting incidents between law enforcement and criminal offenders. The analytic products focus on identifying criminal and demographic attributes associated with the initiation of firearm violence during apprehension. The system is used to assess the relative risk between different scenarios, and the results of this analysis have been used to better inform those responsible for risk-mitigation training as well as those who assess shooting incidents when they occur, such as the USMS Shooting Review Board. This program does not predict shooting incidents, nor does it retrieve information based on any search criteria. Consequently, it falls outside the scope of the Act.

F. Additional Department of Justice Systems

The Department also has other systems or data warehouses supporting advanced analytic tools that do not meet the “data mining” statutory definition. Department law enforcement components employ numerous search tools and databases to help accomplish a variety of missions. Various groups collect, analyze, and 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 and the Markle Foundation, among 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 the Act. The Department has either completed or is in the process of completing privacy documentation for these systems, and

33 Id.
the Department also ensures that it complies with the requirements of the Privacy Act where applicable to such systems.

1. **Organized Crime and Drug Enforcement Task Forces (OCDETF):** As part of the OCDETF Fusion Center (OFC), OCDETF maintains a database named Compass that contains relevant drug and related financial intelligence information from numerous law enforcement organizations. Compass is a relational database that accesses data on subject-based queries or searches, and therefore does not satisfy the data mining definition contained in the Act.

   OCDETF also maintains the OCDETF Management Information System (MIS), which is used by OCDETF member agencies throughout the country. Because the MIS uses subject-based queries, it also falls outside the scope of the Act.

2. **FBI Computer Analysis and Response Team (CART) Family of Systems (FOS):** CART includes 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 stored in CART is lawfully obtained in the course of criminal, national security, and administrative investigations, such as through search warrants or with consent. 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 criminal or national security evidence. Because these searches are retrospective and typically subject-based, the CART tools and capabilities do not meet the definition of “data mining” under the Act.
V. Conclusion

As set forth above, the Department of Justice takes very seriously its mission to prevent terrorism and investigate criminal conduct using all available and lawful tools, while also respecting the privacy and civil liberties of individuals. The use of "data mining" and other advanced analytic tools is extremely valuable and is only undertaken with due regard for the privacy concerns of individuals. The Department conducts the activities and programs described in this report in accordance with its privacy compliance process and all other applicable laws and guidelines, and remains committed to ensuring that all necessary and appropriate protections have been and will continue to be implemented in any of its data mining activities.