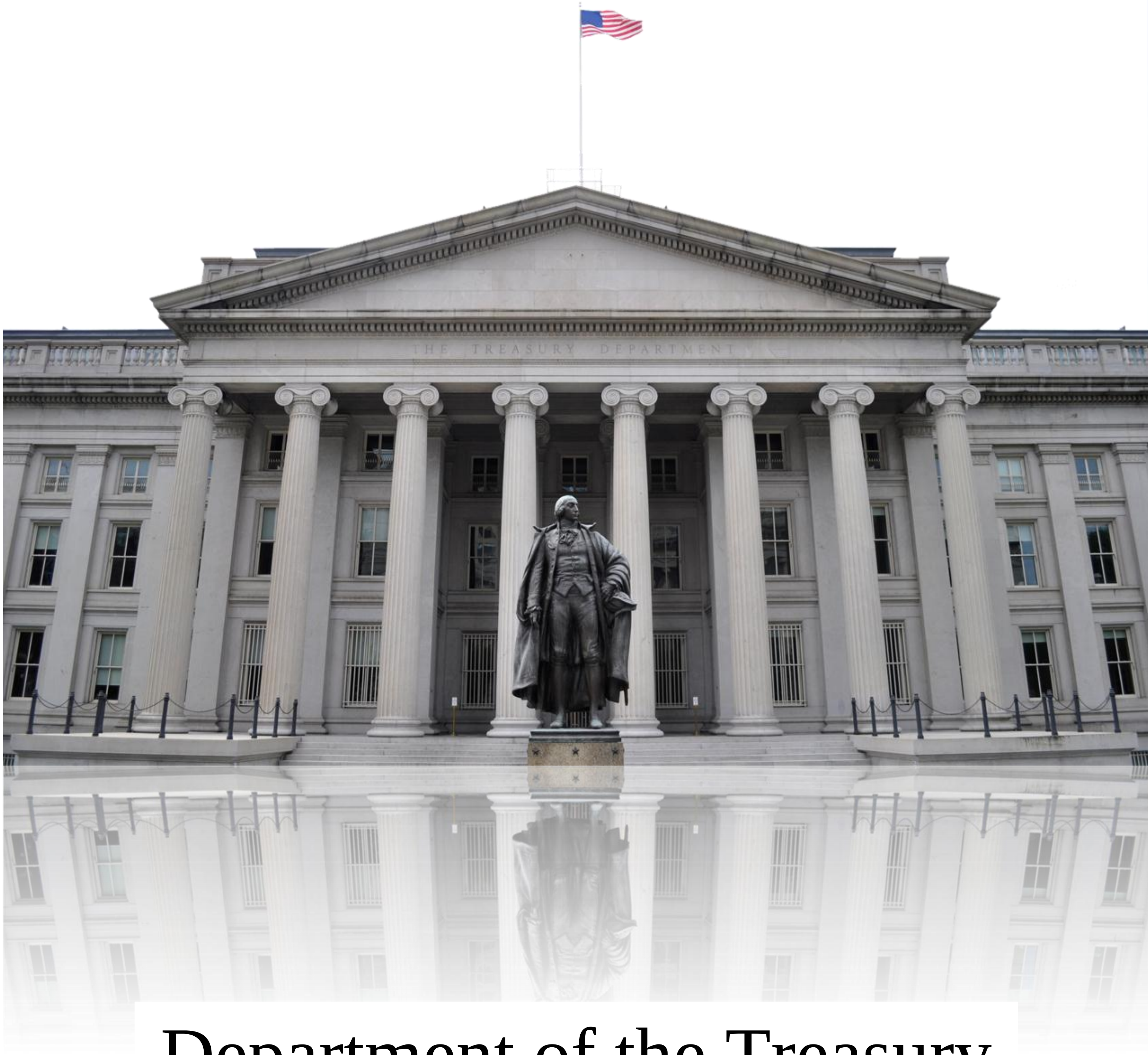

Department of the Treasury

2013 Annual Privacy and Data Mining Reports

MESSAGE FROM THE CHIEF PRIVACY AND CIVIL LIBERTIES OFFICER

I am pleased to present the Department of the Treasury's Annual Privacy and Data Mining Reports for Fiscal Year 2013, as required by Section 522 of the Consolidated Appropriations Act of 2005 and the Federal Agency Data Mining Reporting Act of 2007. For the first time, Treasury is combining these two separate reporting requirements into a single report.

This year was one of change for Treasury's Office of Privacy, Transparency, and Records. In June, Helen Goff Foster joined our team as Deputy Assistant Secretary for Privacy, Transparency, and Records. Under her leadership, the Office of Privacy, Transparency, and Records continues its tradition of proactively safeguarding individual privacy and civil liberties in all Treasury activities.

Inquiries about this report may be directed to privacy@treasury.gov. This report, as well as previous annual reports, can be found on the Department's Privacy Act website at www.treasury.gov/privacy/annual-reports.

Nani A. Coloretti
Senior Agency Official for Privacy & Chief Privacy and Civil Liberties Officer
Assistant Secretary for Management

2013 Annual Privacy and Data Mining Reports

TABLE OF CONTENTS

Message from the Chief Privacy and Civil Liberties Officer 1

Table of Contents 2

Notice of Consolidated Reports and Legislative Language 4

Section One: 6

Department of the Treasury 2013 Annual Privacy Report 6

Background 6

The Role of the Treasury Chief Privacy and Civil Liberties Officer 6

The Role of the Office of Privacy, Transparency, and Records (OPTR).................................... 8

Oversight and Compliance 9

System of Records Notices (SORNs) 9

Privacy Impact Assessments 9

Federal Information Security Management Act of 2002 10

Section 803 of the Implementing Recommendations of the 9/11 Commission Act of 2007.... 10

PII Holdings and Reduction.. 11

Elimination of the Unnecessary Use of Social Security Numbers 11

Privacy Awareness and Training 12

A Culture of Privacy Awareness 12

Leadership and Coordination within Treasury 12

Treasury Directive 25-10: Information Sharing Environment Privacy and Civil Liberties Policy... 12

Do Not Pay 12

Executive Order (E.O.) 13636: Improving Critical Infrastructure Cybersecurity 13

Privacy Act Website.. 13

Terrorist Finance Tracking Program 13

Treasury Computer Matching Programs 14

Section Two: 15

Department of the Treasury 2013 Data Mining Report 15

Background 15

Definitions 15

FinCEN’s Data Mining Activities 16

IRS Data Mining Activities 25

TTB Data Mining Activities 32

Conclusion 36

NOTICE OF CONSOLIDATED REPORTS AND LEGISLATIVE LANGUAGE

In this report, Treasury consolidates the following two reporting requirements to reduce duplication and to provide Congress and the public with a more comprehensive overview of Treasury's privacy compliance and oversight activities:

> (1) The annual privacy report required by Section 522(a) of the Consolidated Appropriations Act of 2005; and
> (2) the Data Mining Reporting Act requirement contained in the Section 803 of the Implementing Recommendations of the 9/11 Commission Act of 2007, 42 U.S.C. § 2000ee–3.

THE ANNUAL PRIVACY REPORT

The Annual Privacy Report has been prepared in accordance with Section 522(a) of the Consolidated Appropriations Act of 2005, which includes the following requirement:

Privacy Officer—

> Each agency shall have a Chief Privacy Officer to assume primary responsibility for privacy and data protection policy, including—
>
> * * *
>
> > (6) preparing a report to Congress on an annual basis on activities of the Department that affect privacy, including complaints of privacy violations, implementation of section 552a of title 5, 11 United States Code, internal controls, and other relevant matters;
>
> * * *

THE DATA MINING REPORT

The Federal Agency Data Mining Reporting Act of 2007, 42 U.S.C. § 2000ee-3, includes the following requirement:

> (c) Reports on data mining activities by Federal agencies
>
> > (1) Requirement for report - The head of each department or agency of the Federal Government that is engaged in any activity to use or develop data mining shall submit a report to Congress on all such activities of the department or agency under the jurisdiction of that official. The report shall be produced in coordination with the privacy officer of that department or agency, if applicable, and shall be made available to the public, except for an annex described in subparagraph (3).

(2) Content of report - Each report submitted under subparagraph (A) shall include, for each activity to use or develop data mining, the following information:

(A) A thorough description of the data mining activity, its goals, and, where appropriate, the target dates for the deployment of the data mining activity.

(B) 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.

(C) A thorough description of the data sources that are or will be used.

(D) 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.

(E) 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.

(F) 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.

(G) 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—

(i) protect the privacy and due process rights of individuals, such as redress procedures; and

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

SECTION ONE:
DEPARTMENT OF THE TREASURY 2013 ANNUAL PRIVACY REPORT

BACKGROUND

The Role of the Treasury Chief Privacy and Civil Liberties Officer

Section 522 of the Consolidated Appropriations Act of 2005[1] requires the Department of the Treasury (Treasury or Department) to appoint a Chief Privacy Officer to assume primary responsibility for privacy and data protection policy. Similarly, Section 803 of the Implementing Recommendations of the 9/11 Commission Act of 2007[2] requires the Department to appoint a senior officer to serve as its Privacy and Civil Liberties Officer. In addition, Office of Management and Budget (OMB) Memorandum 05-08, *Designation of Senior Agency Officials for Privacy*, February 11, 2005 (OMB M-05-08), directs agency heads to designate a Senior Agency Official for Privacy (SAOP) with agency-wide responsibility for ensuring implementation of information privacy protections and full compliance with information privacy laws, regulations, and policies.

Consistent with these requirements, Treasury Directive 25-09, *Privacy and Civil Liberties Activities Pursuant to Section 803 of the Implementing Recommendations of the 9/11 Commission Act of 2007* (TD 25-09), assigns all of these responsibilities to the Treasury Chief Privacy and Civil Liberties Officer (CPCLO). TD 25-09 designates the Assistant Secretary for Management (ASM) as the Department's CPCLO.[3]

In her role as Treasury's CPCLO, the ASM is responsible for:

- assuring that the use of technologies sustain, and do not erode, privacy protections relating to the use, collection, and disclosure of information in identifiable form;[4]
- assuring that technologies used to collect, use, store, and disclose information in identifiable form allow for continuous auditing of compliance with stated privacy policies and practices governing the collection, use and distribution of information in the operation of the program;[5]
- assuring that personal information contained in Privacy Act systems of records is handled in full compliance with fair information practices as defined in the Privacy Act of 1974; [6]
- evaluating legislative and regulatory proposals involving collection, use, and disclosure of personal information by the federal government;[7]

[1] Pub. L. No. 108-447, Section 522(a)(1).
[2] Pub. L. No. 110-53, § 803(a)(1).
[3] TD 25-09 (and all TDs and TOs) are available at http://www.treasury.gov/about/role-of-treasury/orders-directives/.
[4] Consolidated Appropriations Act of 2005, Pub. L. No. 108-447, Section 522(a)(1).
[5] *Id*. at Section 522(a)(2).
[6] *Id*. at Section 522(a)(3).
[7] *Id*. at Section 522(a)(4).

- conducting a privacy impact assessment of proposed rules of the Department on the privacy of information in an identifiable form, including the type of personally identifiable information (PII) collected and the number of people affected;[8]
- preparing a report to Congress on an annual basis on activities of the Department that affect privacy, including complaints of privacy violations, implementation of section 552a of title 5, 11 United States Code, internal controls, and other relevant matters;[9]
- ensuring that the Department protects information in an identifiable form and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction;[10]
- training and educating employees on privacy and data protection policies to promote awareness of and compliance with established privacy and data protection policies;[11]
- ensuring compliance with the Department's established privacy and data protection policies;[12]
- assisting the head of such department, agency, or element and other officials of such department, agency, or element in appropriately considering privacy and civil liberties concerns when such officials are proposing, developing, or implementing laws, regulations, policies, procedures, or guidelines related to efforts to protect the Nation against terrorism;[13]
- periodically investigating and reviewing department, agency, or element actions, policies, procedures, guidelines, and related laws and their implementation to ensure that such department, agency, or element is adequately considering privacy and civil liberties in its actions;[14]
- ensuring that such department, agency, or element has adequate procedures to receive, investigate, respond to, and redress complaints from individuals who allege such department, agency, or element has violated their privacy or civil liberties;[15]
- considering certain factors when providing advice on department, agency, or element proposals to retain or enhance a particular governmental power, including whether the proponent has established[16]
    a. that the need for the power is balanced with the need to protect privacy and civil liberties;
    b. that there is adequate supervision of the use by such department, agency, or element of the power to ensure protection of privacy and civil liberties; and
    c. that there are adequate guidelines and oversight to properly confine its use.
- reviewing and updating privacy procedures;[17]

[8] *Id*. at 522(a)(5).
[9] *Id*. at 522(a)(6).
[10] *Id*. at § 522(a)(7).
[11] *Id*. at § 522(a)(8).
[12] *Id*. at § 522(a)(9).
[13] Implementing Recommendations of the 9/11 Commission Act of 2007, Pub. L. No. 110-53, § 803(a)(1).
[14] *Id*. at § 803(a)(2).
[15] *Id*. at § 803(a)(3).
[16] *Id*. at § 803(a)(4).
[17] OMB M-05-08, Designation of Senior Agency Officials for Privacy, February 11, 2005.

- ensuring implementation of information privacy protections, including compliance with applicable information privacy laws, regulations, and policies;[18]
- ensuring employees and contractors receive privacy training;[19]
- having a role in Treasury's development and evaluation of legislative, regulatory, and other policy proposals that implicate information privacy issues.[20]

The Role of the Office of Privacy, Transparency, and Records (OPTR)

To assist the ASM with the aforementioned responsibilities, TD 25-04, *The Privacy Act of 1974, As Amended*, January 27, 2014 (TD 25-04), designates the Deputy Assistant Secretary for Privacy, Transparency, and Records (DASPTR) as the ASM's principal advisor on issues related to privacy and civil liberties. The DASPTR ensures Treasury collects, maintains, and discloses PII in a manner consistent with legal and policy requirements. The DASPTR leads OPTR and provides the ASM with day-to-day support in executing the privacy and civil liberties duties entrusted to her in her capacity as the Department's CPCLO.

OPTR supports Treasury's missions through three core functions:

- **Safeguarding the privacy and civil liberties** of individuals when Treasury collects, maintains, and discloses personal information;
- Providing **transparency and accountability** to the public with respect to Treasury policies, activities, and functions; and
- Preserving and providing access to Treasury's **institutional knowledge, records, and information resources**.

OPTR is responsible for monitoring and overseeing privacy and civil liberties compliance throughout the Department. This includes working closely with Treasury leadership and Treasury bureaus to develop, implement, and monitor agency-wide privacy policies and procedures in compliance with the U.S. Constitution and relevant federal statutes, Executive Orders, OMB memoranda and guidance, as well as other relevant policy, standards, and regulations. Some of OPTR's operations include:

- Promulgating Treasury-wide policy to ensure consistent application of privacy and civil liberties safeguards in all Treasury activities;
- Reviewing and publishing system of records notices;
- Conducting and analyzing Privacy Impact/Threshold Assessments;
- Reviewing and reporting on paper and electronic incidents involving PII;
- Developing and conducting privacy and civil liberties training;
- Drafting and updating Treasury directives that address privacy and civil liberties issues; and
- Reviewing and implementing Executive Orders and OMB guidance.

[18] *Id.*
[19] *Id.*
[20] *Id.*

OVERSIGHT AND COMPLIANCE

For Treasury to accomplish its mission, it must collect PII from its employees and the public, as well as acquire it from various organizations and other government agencies. The Department is responsible for managing and protecting the information it collects, maintains, and discloses. Federal law, regulations, and policies regulate these activities and are designed to maintain the public's trust while collecting, maintaining, and using PII.

System of Records Notices (SORNs)

A system of records is a grouping of paper or electronic records maintained by a federal agency from which information about an individual is retrieved by the name of the individual or another unique identifier assigned to the individual (e.g., Social Security number). Pursuant to 5 U.S.C. § 552a (e)(4), agencies are required to publish SORNs in the *Federal Register* for each system of records. Treasury has published regulations describing how it collects, maintains, and discloses records about individuals that are maintained in a system of records. These regulations provide procedures by which individuals may request access to their information maintained by Treasury.[21]

In FY 2013, OPTR completed its biennial review of its SORN inventory to help ensure that the existing SORNs accurately describe Treasury's systems of records. As a result, OPTR decommissioned Treasury SORN .008, *Treasury Emergency Management System*, as it was no longer in use. OPTR also made minor updates and edits to many other SORNs.

In addition, in FY 2013 the Department published two new SORNs in the *Federal Register*: Treasury .014, *Department of the Treasury User Profile Services*, and United States Mint .013, *United States Mint National Electronic Incident Reporting System of Records*.

Treasury maintains approximately 200 systems of records, nearly 60 percent of which are maintained by the Internal Revenue Service (IRS). A complete list of the Department's SORNs is available online at: http://www.treasury.gov/privacy/issuances.

Privacy Impact Assessments

A Privacy Impact Assessment (PIA) is an invaluable tool for helping Treasury document its information safeguarding practices while establishing public trust and ensuring the flow of mission-critical information. Section 208 of the E-Government Act of 2002 (E-Gov Act) requires agencies to conduct PIAs for electronic information systems and collections that involve the collection, maintenance, or dissemination of information in identifiable form from or about members of the public. In FY 2013, Treasury reviewed 100 PIAs. Pursuant to the E-Gov Act,

[21] *See* 31 C.F.R. §§ 1.20-1.36.

agencies are required to make PIAs publicly available through the agency website, the *Federal Register*, or other means. The Department's PIAs are available online at: http://www.treasury.gov/privacy/PIAs.

Federal Information Security Management Act of 2002

The Federal Information Security Management Act of 2002 (FISMA) requires each agency to develop, document, and implement an agency-wide program to provide security for the information and information systems that support its operations. In addition, FISMA requires agency Chief Information Officers, Inspectors General, and SAOPs to report to OMB on information security questions that address areas of risk. Federal agencies must report performance metrics related to the management of their privacy programs. This entails tracking and reporting the number of Treasury systems that contain PII, and the number of systems that require and/or have completed a PIA and/or SORN.

For FY 2013, the Department reported a total inventory of 282 FISMA systems containing personal information in identifiable form. One-hundred percent of the Treasury systems known to require a PIA or SORN completed the required documentation and publication requirements.

Number of FISMA Systems containing PII	Percentage with SORN	Percentage with PIA
282	100%	100%

Section 803 of the Implementing Recommendations of the 9/11 Commission Act of 2007

Pursuant to Section 803 of the Implementing Recommendations of the 9/11 Commission Act of 2007, agencies must ensure that adequate processes exist to receive, investigate, respond to, and redress complaints from individuals who allege privacy or civil liberties violations. To meet the requirement, Treasury issued TD 25-09, which directs heads of bureaus and relevant offices to establish internal procedures to ensure accurate and complete reporting to OPTR.

The ASM, with the support of OPTR, continues to provide timely Section 803 metrics to Congress on behalf of the Department. For FY 2013, Treasury performed 496 reviews, provided advice and responses 15 times, and responded to 13 privacy and civil liberties complaints. The types of reviews the Department and its bureaus conducted included:

Types of Reviews	FY2013
Privacy Threshold Assessments/Privacy Impact Assessments	155/126
System of Records Notices and Routine Uses	11
Exhibit 300 Process	3
Social Security Number Redaction/Elimination	162
Computer Matching Agreement Program	26

PII HOLDINGS AND REDUCTION

Treasury maintains an inventory of its PII holdings in its PII Holdings Database. Treasury completed the PII Holdings Database in FY 2012 to comply with OMB Memorandum 07-16, *Safeguarding Against and Responding to the Breach of Personally Identifiable Information*, (OMB M-07-16). In FY 2013, OPTR began working with employees in the Office of the Chief Information Officer's (OCIO) Office of Cybersecurity to develop the Treasury FISMA Inventory Management System (T-FIMS), a new system to enhance Treasury's compliance with FISMA. Once T-FIMS is operational, information regarding particular FISMA systems will be linked to information that OPTR maintains in the PII Holdings Database. This new system will enable both OCIO and OPTR to identify Treasury systems that contain PII easily, and ensure that privacy compliance documentation is reviewed and updated when systems are modified. In addition, OPTR will be able to update the PII Holdings Database rapidly when a new system is in development that maintains PII. This will allow OPTR to get involved in the process earlier to identify and remediate privacy risks.

In addition, Treasury adopted a policy requiring the Office of Human Resources to maintain a single repository of the PII contained in employee Personal Identification Verification (PIV) cards. The central control of employee PII on PIV cards reduces the cost associated with operating separate repositories and enhances privacy and information security by allowing operators to grant and monitor access to one system.

Finally, the IRS Office of Privacy, Governmental Liaison, and Disclosure issued IRS-wide guidance on the practice of including PII in Outlook calendars. The guidance, *Interim Guidance on Personally Identifiable Information (PII) on Outlook Calendars*, details what PII can be included in Outlook calendars and applies to all IRS employees, contractors, and vendors.

ELIMINATION OF THE UNNECESSARY USE OF SOCIAL SECURITY NUMBERS

Throughout FY 2013, Treasury continued to make steady gains in reducing the use of Social Security numbers (SSNs). For example, the IRS implemented a system that masks the SSN by displaying it as a two-dimensional barcode. As a result, IRS masked SSNs on 220 informational and nonpayment notices sent to taxpayers. Of the total number of notices and letters, 58 were nonpayment notices, impacting 20.4 million taxpayers.

FinCEN upgraded and widened deployment of a data leakage protection application to scan for SSNs in e-mail messages, including the attachments to those messages. In so doing, FinCEN has made significant strides in their effort to eliminate the unnecessary use of SSNs.

PRIVACY AWARENESS AND TRAINING

A Culture of Privacy Awareness

OMB M-07-16 requires agencies to train employees on their privacy and security responsibilities before granting them access to agency information and information systems. Additionally, agencies must provide at least annual refresher training to ensure employees continue to understand their responsibilities. In FY 2013, 99 percent of all Treasury employees completed annual privacy awareness training.

In FY 2013, pursuant to OMB A-130, *Management of Federal Information Resources*, Appendix I, Section 3.a(6), OPTR conducted a biennial review of the Department's training practices and updated the departmental course "A Culture of Privacy Awareness."

LEADERSHIP AND COORDINATION WITHIN TREASURY

Treasury Directive 25-10: Information Sharing Environment Privacy and Civil Liberties Policy

The Intelligence Reform and Terrorism Prevention Act of 2004 created the Information Sharing Environment (ISE), which creates a process and structure for the sharing of terrorism information that is needed to enhance national security. As a participant in the ISE, Treasury was required to develop and implement a written ISE privacy and civil liberties protection policy that sets forth the procedures its personnel will follow to implement the privacy guidelines issued by ISE program manager.

In FY 2013, OPTR published TD 25-10, *Information Sharing Environment Privacy and Civil Liberties Policy*, to assist Treasury in implementing ISE privacy and civil liberties requirements. TDs outline departmental policy objectives as well as the roles, responsibilities, and processes for implementing legal and policy obligations. TD 25-10 is available at: http://www.treasury.gov/privacy/directives.

Do Not Pay

In accordance with the Improper Payment Elimination and Recovery Improvement Act of 2012 (IPERIA), OMB designated Treasury's Bureau of the Fiscal Service to host the Do Not Pay Working System. The working system will strengthen and enhance financial management controls to better enable the detection and prevention of improper payments.

OMB Memorandum 13-20, *Protecting Privacy while Reducing Improper Payments with the Do Not Pay Initiative*, authorized Treasury to establish a system of records to carry out activities outlined in IPERIA. In FY 2013, OPTR worked closely with Do Not Pay Business Center leadership to draft a SORN for the system. The notice will appear in the *Federal Register* in FY

2014. As the Privacy Act requires,[22] notice of the new system of records will be provided to the House of Representatives Committee on Oversight and Government Reform, the Senate Committee on Homeland Security and Governmental Affairs, and OMB.

Executive Order (E.O.) 13636: Improving Critical Infrastructure Cybersecurity

On February 12, 2013, the President signed E.O. 13636, *Improving Critical Infrastructure Cybersecurity*, stating: "[i]t is the policy of the United States to enhance the security and resilience of the Nation's critical infrastructure and to maintain a cyber-environment that encourages efficiency, innovation, and economic prosperity while promoting safety, security, business confidentiality, privacy, and civil liberties."

To ensure the inclusion of privacy and civil liberties protections in activities under the Order, section 5(a) of the E.O. required federal agencies to coordinate E.O. 13636-related cybersecurity activities with their SAOP. Section 5(b) further required the SAOP to conduct an assessment of their agency's activities under the Order. As required, OPTR conducted a privacy and civil liberties assessment of the Department's cybersecurity activities under the E.O. As directed under the E.O., Treasury submitted its assessment to the Department of Homeland Security for inclusion in a consolidated public report.

Privacy Act Website

In FY 2013, OPTR made substantial improvements to the Privacy Act page on the Treasury website. The website, www.treasury.gov/privacy, now provides visitors with easier access to PIAs, SORNs, computer matching notices, reports, and more. In addition to making information easier to find on the website, OPTR leveraged technical upgrades to the *Federal Register* website, which now allows users to review notices online more easily.

Terrorist Finance Tracking Program

In FY 2013, OPTR continued to participate in the implementation of the *Agreement Between the European Union (EU) and the United States of America on the Processing and Transfer of Financial Messaging Data from the European Union to the United States for the Purposes of the Terrorist Finance Tracking Program* (TFTP Agreement).[23] The TFTP Agreement allows Treasury to use EU-stored data to track terrorism financing while protecting the privacy interests of EU citizens. OPTR also participated in annual privacy reviews by which EU and U.S. officials review the TFTP program to ensure U.S. compliance under the agreement.

[22] 5 U.S.C. § 552a(r).

[23] The TFTP Agreement and a full discussion of its provisions can be found on the Treasury website at: http://www.treasury.gov/resource-center/terrorist-illicit-finance/Terrorist-Finance-Tracking/Pages/tftp.aspx

U.S.-EU Data Privacy and Protection in Law Enforcement, Criminal Justice, and Public Security Matters

In FY 2013, OPTR continued to participate in the negotiations between the United States and the EU on the Data Protection and Privacy Agreement (DPPA). The DPPA negotiations are an effort to establish standard data protection provisions for future personal information sharing for law enforcement purposes between the United States, the EU, and EU member states.

TREASURY COMPUTER MATCHING PROGRAMS

Pursuant to the Computer Matching and Privacy Protection Act of 1988,[24] Treasury maintains a Data Integrity Board (DIB) to oversee Treasury computer matching programs. Computer matching programs provide a direct benefit to the public by assisting in the elimination of errors and in monitoring waste, fraud, and abuse.

In FY 2013, the Treasury DIB reviewed and approved renewals and extensions for six of the Department's ongoing computer matching programs as well as one new computer matching program – the IRS Data Loss Prevention Project. The purpose of the IRS Data Loss Project is to detect Sensitive But Unclassified information that is transmitted in violation of IRS security policy. Notice of the new matching program, along with the published notices for all of Treasury ongoing computer matching programs, is available online through the Treasury Privacy Act page at: http://www.treasury.gov/privacy/computer-matching-programs.

[24] Pub. L. No. 100-503.

SECTION TWO:

DEPARTMENT OF THE TREASURY 2013 DATA MINING REPORT

BACKGROUND

The Role of the Treasury Chief Privacy and Civil Liberties Officer (CPCLO)

The Department of the Treasury (Treasury or the Department) is providing this report to Congress pursuant to Section 804 of the *Implementing Recommendations of the 9/11 Commission Act of 2007* (9/11 Commission Act), entitled the *Federal Agency Data Mining Reporting Act of 2007* (Data Mining Reporting Act or the Act). This report discusses activities currently deployed or under development in the Department that meet the Data Mining Reporting Act's definition of data mining. The report also provides the information the Act requires with respect to each data mining activity.

DEFINITIONS

(1) DATA MINING. The term "data mining" means a program involving pattern-based queries, searches, or other analyses of one 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) DATABASE. The term "database" does not include telephone directories, news reporting, information publicly available to any member of the public without payment of a fee, or databases of judicial and administrative opinions or other legal research sources.[25]

Three Treasury bureaus maintain systems using applications that meet the definition of data mining: the Financial Crimes Enforcement Network (FinCEN), the Internal Revenue Service (IRS), and the Alcohol and Tobacco Tax and Trade Bureau (TTB). The IRS and FinCEN

[25] 42 U.S.C § 2000ee-3(b)(1). "[E]lectronic telephone directories, news reporting, information publicly available to any member of the public without payment or a fee, or databases of judicial and administrative opinions or other legal research sources" are not "databases" under the Act.

systems were discussed in previous Treasury data mining reports. TTB is the only new addition to this year's report.

FINCEN'S DATA MINING ACTIVITIES

(A) *A thorough description of the data mining activity, its goals, and, where appropriate, the target dates for the deployment of the data mining activity.*

FinCEN is statutorily obligated to analyze information to "determine emerging trends and methods in money laundering and other financial crimes."[26] These trend analyses typically involve querying a database FinCEN maintains containing information reported by financial institutions under the Bank Secrecy Act (BSA).[27] This information (BSA information or BSA reports) is collected where it has a "high degree of usefulness in criminal, tax, or regulatory investigations or proceedings, or in the conduct of intelligence or counterintelligence activities, including analysis, to protect against international terrorism."[28]

FinCEN conducts analyses to determine emerging trends and methods in money laundering by: (1) examining reports filed on specific violations (e.g., terrorism financing) or filed on specific industries or geographic areas, and conducting analyses on these subsets to determine whether they contain any identifiable trends, patterns or methods; (2) conducting statistical analyses of currency flows over time to determine whether the data contains anomalous trends, patterns or methods; and (3) identifying trends, patterns or specific activities indicative of money laundering or financial crimes through the review and evaluation of reports as part of ongoing review processes.

FinCEN also engages in proactive efforts to identify subjects for investigation using trend, statistical or strategic analyses. FinCEN also identifies subjects for investigation via link-analysis software systems (see item B below). This includes searching for unknown subjects by establishing search criteria based on previously established suspicious or illicit patterns. Other proactive methods include identifying subjects connected through the same addresses or telephone numbers, and searching for subjects with the largest number of BSA reports filed on their financial activities.

(B) *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*

FinCEN's "Advanced Analytics" system is designed to allow users to query across large data sets based on user-defined text patterns or data parameters. The following data sets are available for query within that system: (1) all BSA reports authorized by statute or regulation maintained

[26] 31 U.S.C. § 310(b)(2)(C)(v).
[27] 31 U.S.C. § 5311, *et seq.*
[28] 31 U.S.C. § 5311.

in report-specific files; (2) FinCEN's case management system; and (3) third party data sets (e.g., the Death Master File).[29] This system also enables users to create scheduled queries on user-defined data parameters.

In FY 2013, the majority of users of BSA data were transitioned from the IRS hosted BSA data system, the Currency & Banking Retrieval System (WebCBRS), to the FinCEN hosted system FinCEN Query. Users with access to FinCEN Query are able to query the BSA data set based on user-defined patterns or data parameters.

The basis for determining whether particular patterns or anomalies are indicative of terrorist or criminal activity varies. Because many BSA reports do not reveal the potential underlying criminal activity leading to the reported financial activity, FinCEN attempts to identify any illicit cause for suspicious trends, patterns or methods by querying law enforcement databases on subjects or by identifying other financial or commercial records that may reinforce indications of anomalous or illicit activities.

(C) *A thorough description of the data sources that are being or will be used*

Reports provided under the BSA as administered by FinCEN, e.g., a report by a financial institution of a suspicious transaction relevant to a possible violation of law or regulation,[30] form the underlying data for FinCEN's manual and automated proactive search methods and trend analysis activities. Commercially available databases are used to support or further identify information and to aid in the identification of an illicit cause based on suspicious trends, patterns, or methods. FinCEN's trend analysis utilizes any records available to FinCEN in fulfilling its mission, including subpoenaed financial records, public source information, commercial database information, Census bureau data, and Federal Reserve data. This analysis is used to support or amplify conclusions or hypotheses derived from the analysis of BSA data.

(D) *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*

FinCEN provides strategic and tactical products for its law enforcement clients. These proactive products generally fall within two categories: (1) referrals based on review and evaluation of Suspicious Activity Reports (SARs); and (2) investigative lead information that complemented, or arose from, strategic assessments of geographic areas, industries, or issues.

FinCEN also produces strategic-level proactive (self-initiated) threat assessments of geographic areas, violation types, industries, and terrorism financing issues. FinCEN received feedback demonstrating that these types of products are useful to law enforcement and the public. For example, FinCEN continues to research SARs proactively in order to identify significant

[29] The Death Master File is Social Security Administration (SSA) information used by medical researchers, hospitals, medical programs, and law enforcement agencies and other government agencies to verify a person's death and to prevent fraud. Although it is SSA information, the National Technical Information Service in the Department of Commerce maintains the database. *See* http://www.ntis.gov/products/ssa-dmf.aspx.

[30] 31 U.S.C. § 5318(g).

suspicious activity in certain foreign countries, particularly with respect to corrupt foreign officials. FinCEN received feedback indicating that these alerts have generated significant interest within foreign governments, leading to the initiation of several investigations into high-level corruption within those countries.

FinCEN does receive positive feedback on its products generated in support of law enforcement and regulatory efforts to combat terrorism financing; healthcare, mortgage and government programs fraud; and southwest border narcotics and bulk cash smuggling. For example:

- FinCEN has continued its partnership with the Department of Health and Human Services (HHS) and its Health Care Fraud Prevention and Enforcement Action Teams (HEAT). Together, through the use of FinCEN data analysis for specific geographic locations, FinCEN and HHS HEAT are able to identify complex large-scale fraud schemes and the most egregious individual perpetrators and organized groups defrauding the health care system. The teams, which include investigators and prosecutors from the Department of Justice (DOJ) and the HHS, are working to strengthen existing programs, investigate fraud, and invest in new resources and technology to prevent future fraud, waste, and abuse.

- In FY 2013, FinCEN produced 26 analytical/financial reports, involving analysis of over 5,019 BSA records. In addition, FinCEN provided case support to two federal and two state and local agencies. FinCEN’s support in these matters was referenced in DOJ press releases about health care fraud takedowns.

- FinCEN continues to support U.S. law enforcement counter-narcotics initiatives on the southwest border by employing advanced technology to systematically detect anomalous or suspicious activity in FinCEN data and other financial records. FinCEN collaborates with key partners in law enforcement and the private sector to share and exploit information useful for understanding emerging money laundering trends, and to identify nodes of illicit activity for potential investigation. In addition to informing senior leaders in the U.S. and Mexican public and private sectors about bank vulnerabilities associated with these cross-border financial flows, FinCEN was also able to provide law enforcement field offices with specific information for targeting purposes.

- FinCEN queried its SAR databases for reports related to several mortgage securities fraud cases under investigation by DOJ, and provided nearly 4,200 SARs to Assistant United States Attorneys (AUSAs) in several regions. These AUSAs commented that the SARs strengthened their cases by providing additional sources of corroborating evidence.

- Since 2009, FinCEN has queried its BSA databases for SARs reflecting suspicious use of proceeds from government financial support programs such as the Trouble Asset Relief Program (TARP). In FY 2013, FinCEN sent the Special Inspector General for TARP approximately17,388 SARs referencing 22,390 subjects of possible interest indicating more than $117 billion in suspicious financial activity.

FinCEN has experienced some difficulty in assessing the efficacy of its proactive work due to insufficient feedback from law enforcement. This makes it difficult to quantify the number of investigations opened and the quality of the targets identified, i.e., whether the identified activity was in fact related to illicit activities.

(E) *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 or will be taken with regard to the property, privacy, or other rights or privileges of any individual as a result of implementing the data mining activity*

The impact of FinCEN's congressionally-mandated mission on the privacy and civil liberties of individuals has been and will continue to be minimal. As a threshold matter, the Supreme Court has ruled that the financial information that banks and other financial institutions hold, and that FinCEN collects and analyzes pursuant to its authority in 31 U.S.C. § 310 and the BSA (discussed in more detail in item (F) below), carries no constitutionally protected "expectation of privacy."[31] Moreover, the Right to Financial Privacy Act of 1978[32] expressly provides that it gives no protection for financial records or information required to be reported in accordance with any federal statute or regulation, which includes information contained in BSA reports.[33]

Significantly, FinCEN takes no adverse actions against individuals based on the existence of, or information contained in, BSA data. Since a BSA report itself is not necessarily indicative of criminal activity, it is only useful when viewed in conjunction with other evidence. Therefore, FinCEN provides the data, or analytical products analyzing the data, to outside agencies where the information may be relevant to current or potential investigations or proceedings under the jurisdiction of those agencies.

The BSA provides standards for proper use of the financial data collected by FinCEN. The collected information is also generally subject to the Privacy Act of 1974,[34] discussed in more detail under item (F) below. FinCEN has developed extensive policies and procedures to ensure, to the extent reasonably possible, that: (1) the analyzed information is used for purposes authorized by applicable law; and (2) the security of the information is adequately maintained. Analytical products produced by FinCEN are subject to clearly specified restrictions regarding use and further dissemination of the products to ensure that the products will only be used by appropriate agencies for statutorily authorized purposes. To the extent such products reference information collected pursuant to the BSA, FinCEN has issued guidelines requiring user agencies to attach warning language to such products and to follow specific procedures for further dissemination of the BSA information. These procedures aim to ensure that: (1) only appropriate agencies will have access to the information; (2) the information will be used for statutorily authorized purposes; (3) agencies with access are aware of the sensitivity of the

[31] *United States v. Miller*, 425 U.S. 435, 442 (1976).

[32] 12 U.S.C. § 3401, *et seq.*

[33] 12 U.S.C. § 3413(d) ("Disclosure pursuant to Federal statute or rule promulgated thereunder nothing in this chapter shall authorize the withholding of financial records or information required to be reported in accordance with any Federal statute or rule promulgated thereunder.")

[34] 5 U.S.C. § 552a.

material; and (4) FinCEN will be able to track which agencies have such materials in their possession.

(F) *A list and analysis of the laws and regulations that govern the information being collected, reviewed, gathered, analyzed, or used in the data mining activity*

1. The Bank Secrecy Act, 31 U.S.C. § 5311, ***et seq*****. (BSA) and Implementing Regulations, 31 C.F.R. Chapter X,** ***et seq*****:**

31 U.S.C. § 5311— Declaration of Purpose

This section specifies that the purpose of the recordkeeping and reporting requirements in the BSA is to "require certain reports where they have a high degree of usefulness in criminal, tax, or regulatory investigations or proceedings, or in the conduct of intelligence or counterintelligence activities, including analysis, to protect against international terrorism." FinCEN strives to ensure that all uses of information are consistent with this purpose.

31 C.F.R. § 1010.301— Determination by the Secretary

This regulation provides the determination that the reports collected pursuant to the BSA have a "high degree of usefulness" in the areas covered by 31 U.S.C. § 5311.

31 U.S.C. § 5319 — Availability of Reports

This section makes it clear that, upon request, the Secretary (as delegated to FinCEN) is required to provide BSA information for the purposes specified in 31 U.S.C. § 5311, to agencies including state financial institutions supervisory agencies, United States intelligence agencies, or self-regulatory organizations registered with the Securities and Exchange Commission or the Commodity Futures Trading Commission. This list of types of agencies is not exhaustive, but those listed are clearly covered. This section also provides that reports collected pursuant to the BSA are exempt from disclosure under the Freedom of Information Act, 5 U.S.C. § 552.

31 C.F.R. § 1010.950 — Availability of Information

This section authorizes the Secretary to make BSA information available to appropriate agencies for purposes specified in the BSA, and specifies that the requesting agency is to receive the information "in confidence."

31 U.S.C. § 5313 — Reports on domestic coins and currency transactions

This section provides for the reporting by financial institutions of reports of certain currency transactions involving more than an amount specified by the Secretary (as delegated to FinCEN).

31 C.F.R. §§ 1010.311; 1021.311 — Reports of transactions in currency

These regulations implement the reporting requirement of 31 U.S.C. § 5313 and specify the amount of reportable transactions in currency at more than $10,000.

31 U.S.C. § 5316 — Reports on exporting and importing monetary instruments

This section requires reports by those that transport currency or other monetary instruments of more than $10,000 at one time from outside the United States into the United States, or from the United States outside the United States.

31 C.F.R. § 1010.340 — Reports of transportation of currency or monetary instruments

This regulation implements the reporting requirement of 31 U.S.C. § 5316 with respect to currency or other monetary instruments of more than $10,000 imported into the United States or exported outside the United States.

31 U.S.C. § 5314 — Records and reports on foreign financial agency transactions

This section authorizes the Secretary (as delegated to FinCEN) to prescribe regulations requiring the reporting of certain types of foreign transactions and relationships with foreign institutions.

31 C.F.R. § 1010.350 — Reports of foreign financial accounts

This regulation, implementing 31 U.S.C. § 5314, requires that U.S. persons file reports of foreign bank accounts.

31 U.S.C. § 5318(g) — Reporting of suspicious transactions

This section authorizes the Secretary (as delegated to FinCEN), to require the reporting of suspicious transactions relevant to a possible violation of law. The section also provides for the confidentiality of such reports, barring financial institutions from notifying anyone involved in the transaction that the transaction has been reported. Government employees are subject to the same confidentiality restrictions, except as “necessary to fulfill the official duties” of such employees. The policies and procedures detailed above in response to item (E) are aimed, in large part, at maintaining the confidentiality of these reports.

31 C.F.R. §§1010.320;1020.320; 1021.320; 1022.320; 1023.320; 1024.320; 1025.320; 1026.320 — Reports of Suspicious Transactions

These regulations implement 31 U.S.C. § 5318(g), requiring covered financial institutions to file suspicious activity reports and requiring the maintaining of strict confidentiality of the reports.

31 U.S.C. § 5331— Reports relating to coins and currency received in nonfinancial trade or business

This section provides for the reporting of currency transactions of more than $10,000 by businesses other than financial institutions.

31 C.F.R. § 1010.330 — Reports related to currency in excess of $10,000 received in a trade or business

This regulation implements 31 U.S.C. § 5331.

2. The Privacy Act of 1974 (Privacy Act), 5 U.S.C. § 552a

Generally, the Privacy Act protects reports that FinCEN collects pursuant to the BSA in that the reports are "records" contained in a "system of records."[35] The Privacy Act provides that covered records may be disclosed without the written permission of the individual to whom the record pertains if they are disclosed pursuant to a "routine use."[36] FinCEN includes sets of routine uses in its published Systems of Records Notices (SORNs) as the Privacy Act requires. These routine uses identify the individuals and organizations external to Treasury with which FinCEN routinely shares BSA information. Sharing with these specified recipients is consistent with the purposes for which the information is collected, as specified in the BSA.

FinCEN has three SORNs that cover the information it collects under the BSA: (1) Treasury/FinCEN .001, *FinCEN Investigations and Examinations System*[37], (2) Treasury/FinCEN .002, *Suspicious Activity Report (SAR) System*[38], and (3) Treasury/FinCEN .003, *Bank Secrecy Act (BSA) Reports System*.[39]

FinCEN followed Privacy Act procedures (including appropriate public notice and comment periods) to exempt certain records maintained in the SARs and BSA systems of records from specific provisions of the Privacy Act, including those allowing for subject's access to the reports, notification to the subject when reports are shared, requests for correction of the contents of such reports by the subject, and the civil remedies covering these areas. These exemptions prevent individuals who are planning crimes from avoiding detection or apprehension or structuring their operations to avoid detection or apprehension.

3. Other Relevant Provisions

31 U.S.C. § 310— Financial Crimes Enforcement Network

[35] 5 U.S.C. § 552a (b)(3) ("Records maintained on individuals . . . (4) the term "record" means any item, collection, or grouping of information about an individual that is maintained by an agency, including, but not limited to, his education, financial transactions, medical history, and criminal or employment history and that contains his name, or the identifying number, symbol, or other identifying particular assigned to the individual, such as a finger or voice print or a photograph; (5) the term "system of records" means a group of any records under the control of any agency from which information is retrieved by the name of the individual or by some identifying number, symbol, or other identifying particular assigned to the individual");

[36] 5 U.S.C. § 552a (b)(3).

[37] 77 Fed. Reg. 60016 (October 1, 2012).

[38] *Id.* at 60017.

[39] *Id.* at 60020.

This section establishes FinCEN as a bureau in the Department of the Treasury, sets out the duties and powers of the Director, and empowers the Director to administer the BSA to the extent delegated by the Secretary of the Treasury.[40] This section also requires FinCEN to maintain a "government-wide data access service" for the information collected under the BSA, as well as records and data maintained by other government agencies and other publicly and privately available information.[41] FinCEN is required to "analyze and disseminate" the data for a broad range of purposes consistent with the law.[42] These purposes include identifying possible criminal activity; supporting domestic and international criminal investigations (and related civil proceedings); determining emerging trends and methods in money laundering and other financial crimes; supporting the conduct of intelligence and counterintelligence activities, including analysis, to protect against international terrorism; and supporting government initiatives against money laundering.

The section further requires that FinCEN furnish research, analytical, and informational services to financial institutions and domestic and foreign law enforcement agencies for the "detection, prevention, and prosecution of terrorism, organized crime, money laundering and other financial crimes" and provide "computer and data support and data analysis to the Secretary of the Treasury for tracking and controlling foreign assets."[43] The section also provides for the establishment of standards for making the information available through efficient means, and to screen appropriate users and appropriate uses.[44] The activities and procedures described in this report adhere to the requirements of this statute.

(G) *A thorough discussion of the policies, 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:*

(i) Protect the privacy and due process rights of individuals, such as redress procedures

A description of the policies, procedures, and guidance in place to ensure the privacy and due process rights of individuals that are the subject of FinCEN data mining activities is provided in subsection (E) above.

(ii) Ensure that only accurate and complete information is collected, reviewed, analyzed, or used and guard against any harmful consequences of potential inaccuracies

FinCEN, through its data perfection procedures, ensures that information contained in the database of BSA reports is accurate and complete. In addition, as discussed in item (E) above, FinCEN does not take adverse actions against individuals (outside the context of enforcing the requirements of the BSA itself) based on the information contained in BSA reports. In addition, because user agencies only use BSA information in conjunction with other evidence, a BSA

[40] Treasury Order 180-01, *Financial Crimes Enforcement Network* (March 4, 2003) (establishing FinCEN as a bureau in the Department of the Treasury and delegating authority to administer, implement, and enforce the BSA to the Director of FinCEN).

[41] 31 U.S.C.§ 310(b)(2)(B)

[42] *Id*. at § 310(b)(C)(i)-(vii).

[43] *Id.* at § 310(b)(2)(E), (G).

[44] *Id*. at § 310(c)(1-2).

report in itself is not used as the sole basis for adverse actions by user agencies. Accordingly, there is an inherent system of “checks and balances” in the use of BSA information that greatly reduces the risk of harmful consequences from inaccuracies that may be contained in BSA reports.

IRS DATA MINING ACTIVITIES

(A) *Data mining activity, goals, and target dates for the deployment of data mining activity, where appropriate*

Three divisions of the IRS are engaged in data mining activities covered by the Act: IRS Criminal Investigation organization (IRS-CI); the IRS Small Business/Self-Employed Division (SB/SE); and the IRS Wage and Investment Division (W&I). Each of these IRS divisions uses one or more of four available data mining applications to search for specific characteristics that are indicators of potential criminal activity:

- Reveal;
- Investigative Data Analytics (IDA);
- Electronic Fraud Detection System (EFDS); and
- Web-CBRS.

IRS-CI Data Mining Activities

IRS-CI is tasked with protecting IRS revenue streams by detecting fraudulent activity and preventing recurrences. IRS-CI uses the Reveal, IDA, and EFDS systems to support this work. Data uncovered using these systems may be reflected in indictments and criminal prosecutions.

Reveal is a data query and visualization tool that allows CI analysts and agents to query and analyze large and potentially disparate sets of data through a single access point. This enhances the analyst's ability to develop a comprehensive picture of suspicious or criminal activity. The program presents information to the user visually, exposing associations between entities in the data that might otherwise remain undiscovered. The VisuaLinks tool within Reveal builds visualization diagrams. IRS-CI Lead Development Centers (LDC),[45] Scheme Development Centers (SDC),[46] and field offices all use the system to identify and develop leads for counterterrorism, money laundering, offshore abusive trust schemes, and other financial crime.

IDA is a data query tool currently in use at the LDCs, SDCs, and field offices, and it provides CI analysts and special agents with the ability to query and analyze large and potentially disparate sets of electronic data through a single access point. IDA enhances these search results by linking relationships and exposing associations with events and other individuals. By using the IDA application, special agents and investigative analysts can proactively identify patterns indicative of illegal activities. This tool enhances investigation selection and supports investigative priorities in tax law enforcement, counterterrorism, and other high-priority criminal investigations.

[45] The LDCs focus on other cases such as international and terrorism cases and are responsible for the voluntary disclosure program.

[46] IRS-CI has eight SDCs. These are groups of CI Investigative Analysts who develop Questionable Refund Program and Return Preparer Program schemes for the field offices. These SDCs are part of IRS-CI.

The IDA application uses data for both reactive and proactive queries. Reactive queries are a result of specific, targeted investigations; proactive queries are the result of pattern matching to generate leads. Data available in the IDA application enable users to detect suspicious financial transactions indicative of money laundering, terrorism, and other financial crimes. IDA query results are used exclusively for the purpose of generating leads. Any investigative process that results from these leads uses the corresponding data from the originating systems.

IRS-CI and W&I both use EFDS to maximize detection of tax return fraud. EFDS compiles, cross-references, and verifies information indicative of potentially fraudulent tax returns. As EFDS receives returns, it loads and assigns a score to each tax return. Scores range from 0.0 to 1.0, with a higher score indicating a greater potential for fraud. IRS-CI does not directly examine the scores, but does use returns that W&I determines to be potentially fraudulent as a basis for its criminal investigations.

Web-CBRS is a web-based application that accesses a database containing BSA forms and information. IRS-CI has directed its users to access the FinCEN Query system (described in the previous section of this report) as the system of record for all BSA data, rather than Web-CBRS. Accordingly, IRS-CI no longer accesses the Web-CBRS database for research in tax cases, tracking money-laundering activities, investigative leads, intelligence for the tracking of currency flows, or corroborating information and probative evidence.

IRS SB/SE will continue to use Web-CBRS until early 2014. IRS plans to retire the data mining portion of Web-CBRS by the third quarter of FY 2014, at which time it will transition to FinCEN. In addition, the information in the system will be transferred to another system of records and maintained by FinCEN.

(B) *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*

Reveal, IDA, and Web-CBRS do not provide IRS with the ability to determine indicators of terrorist or criminal activity. Special agents and investigative analysts use "canned queries" based on experience. For example, an analyst might use the Reveal database to search for individuals that have had five or more SARs filed on them by financial institutions in a six-month period. Agents and analysts determine indicators of fraudulent activity based on previous successful investigations of money laundering, counterterrorism, and BSA violations.

IRS-W&I uses EFDS to identity potentially fraudulent activity. IRS-CI uses the fraudulent tax returns identified by IRS-W&I as a basis for its criminal investigations. Paper refund returns come to EFDS from the Generalized Mainline Framework[47] and Questionable Refund

[47] The Generalized Mainline Framework is a service center pipeline processing system that validates and perfects data from a variety of input sources. Tax returns, remittances, information returns, and adjustment and update transactions in the system are controlled, validated, corrected, and passed on for master file posting.

Program.[48] This allows IRS-W&I and SDC employees to review those returns for suspicious activities.

EFDS employs a data mining technology called IBM SPSS Modeler. Using this tool, EFDS creates rule sets using a standard built-in algorithm called C5.0. Using examples of current and prior year verified fraud and non-fraud data, the machine-learning model discerns patterns or rules indicative of fraud. The output of the model is a score where a higher score (in the range of 0.0 to 1.0) represents a higher risk or a higher likelihood of a return being fraudulent.

If a return meets designated score tolerances and other criteria, IRS-W&I personnel examine the return for fraudulent activity. Once a return is verified to be false via the wage verification process, EFDS adds fraudulent returns to its Scheme Tracking and Retrieval System (STARS) component. IRS-CI investigators examine the returns in STARS to find possible schemes, or fraudulent patterns, which may result in a referral to a CI field office for investigation.

(C) ***Data sources that are or will be used:***

The IRS-CI applications Reveal and IDA leverage the following data sources.

- **Taxpayer:** The source is the electronically filed return (as transmitted through the MeF or a paper filed tax return.
- **Employers/Payers:** Information from employers/payers captured on various forms as stored in the Information Returns Master File (IRMF).
- **Other Treasury sources:** BSA data provided by FinCEN, Specially Designated Nationals' data provided by the Office of Foreign Assets Control.
- **Other IRS sources:** Tax Exempt Organizations data, Voluntary Disclosures, Criminal Investigations data.

The EFDS application leverages the following data sources.

- **Taxpayer**: The source is the electronically filed return (as transmitted through the MeF) or a paper filed tax return.
- **Employers/Payers**: Information from employers/payers captured on Form W-2 and/or form 1099 as stored in the IRMF.
- **Other federal agencies**: Federal Bureau of Prisons for prisoner information; BSA data; HHS Services for information on new hires; Social Security Administration for National Accounts Profile data for dates of births and deaths.
- **State and local agencies**: All states and the District of Columbia prisons deliver prisoner-listing information annually to IRS-W&I in electronic format.
- **Other third party sources**: IRS-W&I purchases commercial public business telephone directory listings/databases to contact employers for employment and wage information.

[48] The Questionable Refund Program (QRF) is a subsystem of EFDS. QRF is a nationwide multifunctional program designed to identify fraudulent returns, to stop the payment of fraudulent refunds and to refer identified fraudulent refund schemes to CI field offices.

(D) *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*

The data uncovered during the query searches are only leads and require additional investigative steps for quality verification. There is no empirical data on the efficacy of searches by the Reveal and IDA applications.

The efficacy of the data mining on EFDS can be measured in terms of fraud prevention. A key overall measure of efficacy is "hit: scan," which represents the number of returns selected for verification that, upon inspection by IRS employees, are found to be fraudulent. The overall "hit: scan" for the EFDS system is 1:1.2 for FY 2013. This means that the data mining program accurately predicts fraudulent returns in 10 of 12 cases.

As discussed previously, Web-CBRS users are migrating to the FinCEN Query system. The efficacy of the FinCEN Query system is discussed in Section (D) of that report.

(E) *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 or will be taken with regard to the property, privacy, or other rights or privileges of any individual as a result of implementing the data mining activity*

Once evidence of fraud is discovered, laws and administrative procedures, policies, and controls govern the ensuing actions. Reveal and IDA applications use personally identifiable information (PII) for pattern matching but the results of a query are used for further investigation. IRS-CI follows the IRS security and privacy IRM standards and regulations for the use and protection of PII.

The impact or likely impact of the EFDS data mining activities on privacy and civil liberties of individuals is governed by 26 U.S.C. § 6103, which provides general rules of maintaining confidentiality and permissible disclosures. Under this statute, all taxpayer data are private and confidential and protected from disclosure except under specific conditions. Additional laws provide for civil and criminal penalties for any unauthorized disclosure of taxpayer data. The penalties include (1) felony for the willful unauthorized disclosure of tax information, (2) misdemeanor for the unauthorized inspection of tax information, and (3) civil cause of action for the taxpayer whose information has been inspected or disclosed in a manner not authorized by Section 6103. The CI Special Agents receive periodic training on maximum sentencing and penalties for each criminal violation. Access to the system requires a background check. IRS has a system, Online 5081, that governs program access authorization. The only users with access to Online 5081 are current CI personnel who are granted access on a need-to-know basis.

Further, EFDS data mining activities, including its machine learning and scoring process, do not use any PII in determining whether a return is likely to be fraudulent. Scoring occurs on the characteristics of the return in question, not on the PII. When performing investigative techniques, PII associated with the return is pulled in to assist in validating the return was filed using the taxpayer account in question and to determine venue of the investigation.

The tax returns that IRS-CI reviews are the subjects of criminal investigations and actions based on tax laws, policies, and criminal procedures. Other tax returns are subjected to IRS civil treatments and examination procedures that provide for due process and redress procedures through taxpayer notification, appeals, and tax court options.

(F) *A list and analysis of the laws and regulations that govern the information being collected, reviewed, gathered, analyzed, or used in the data mining activity*

The use of all tax data is governed by 26 U.S.C. § 6103. Subsection (a) sets out the general rule of confidentiality. Subsection (b) sets forth definitions of terms commonly used throughout Section 6103. Subsections (c) through (o) of Section 6103 contain exceptions to the general rule of confidentiality. These subsections permit disclosures as described generally below:

- Section 6103(c) – Disclosures to taxpayer's designees (consent);
- Section 6103(d) – Disclosures to state tax officials;
- Section 6103(e) – Disclosures to the taxpayer and persons having a material interest;
- Section 6103(f) – Disclosures to committees of Congress;
- Section 6103(g) – Disclosures to the President and White House;
- Section 6103(h) – Disclosures to Federal employees and the courts for tax administration purposes;
- Section 6103(i) – Disclosures to Federal employees for non-tax criminal law enforcement purposes and to combat terrorism, as well as the Government Accountability Office;
- Section 6103(j) – Disclosures for statistical purposes;
- Section 6103(k) – Disclosures for certain miscellaneous tax administration purposes;
- Section 6103(l) – Disclosures for purposes other than tax administration;
- Section 6103(m) – Disclosures of taxpayer identity information (generally for Federal debt collection purposes);
- Section 6103(n) – Disclosures to contractors for tax administration purposes; and
- Section 6103(o) – Disclosures with respect to wagering excise taxes.

In addition to disclosures permitted under the provisions of Section 6103, other provisions of the Code also authorize disclosure of tax information. For example, Section 6104 authorizes disclosure of certain tax information regarding tax-exempt organizations, trusts claiming charitable deductions, and qualified pension plans. Section 6110 authorizes disclosure of certain written determinations and their background files. The information contained in Web-CBRS is gathered under the guidelines dictated by the Bank Secrecy Act[49].

(G) *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:*

[49] 31 U.S.C. § 5311.

(i) *Protect the privacy and due process rights of individuals, such as redress procedures*

All tax information is protected as required in 26 U.S.C. § 6103 (see E and F above). All employees who interact with tax return and other protected information are required to undergo yearly refresher training that details their responsibilities with respect to information protection and disclosure. In addition to covering 26 U.S.C. § 6103 disclosure provisions, this training module also includes information on the Privacy Act, E-Government Act, Freedom of Information Act, and policies related to protecting PII and other sensitive information. The use of BSA information is strictly controlled under the statute that directs its collection.

The data uncovered during query in Reveal and IDA applications are used as a lead and requires additional investigative steps to verify the quality of the information, as discussed above. IRS maintains an audit trail on all users' access to case data. In addition, a full system log is maintained for any system level activities, including new data loads to the IDA application.

EFDS does not determine whether a return is fraudulent or whether a person is going to be subject to criminal prosecution. Once fraud is suspected, laws and administrative procedures, policies, and controls govern criminal investigations or any other ensuing actions. Due process is provided during any ensuing criminal investigation or civil action.

(ii) *Ensure that only accurate and complete information is collected, reviewed, analyzed, or used, and guard against any harmful consequences of potential inaccuracies*

An individual/entity self-reports tax data when submitting the information to the government. Web-CBRS data are gathered from information compiled by the reporter based on information provided by their customer or based on the reporter's personal experience. Investigators scrutinize the Suspicious Activity Reports filed by the subject companies and request grand jury subpoenas for the underlying documentation. The supporting records are examined and individuals of interest are identified.

The Reveal and IDA applications are not the authoritative owners of data. However, the data is used for investigative analysis purposes under the IRS Internal Revenue Manual (IRM) standards and guidelines. The data uncovered during query searches are only used as a lead and require additional investigative steps to verify the quality of the information. Therefore, IRS-CI uses this data for generating leads and the special agents verify this information through further investigative work.

The tax return information and other information stored in EFDS used for data mining are based on outside data sources. The only data generated directly in EFDS are the processing steps and the results of examinations of possibly fraudulent returns. Through a series of test case procedures executed through Application Qualification Testing (AQT), Systems Acceptability Testing (SAT), and Final Integration Test (FIT), the IRS verifies that the data loaded into EFDS matches the data from the input source and that the system accurately displays the data in the EFDS end user applications. AQT, SAT, and FIT perform verification with each release of the system.

IRS applications are required to have internal auditing capabilities. The internal audits track user access and queries performed with checks against misuse.

TTB Data Mining Activities

(A) ***Data mining activity, goals, and target dates for the deployment for data mining activity, where appropriate.***

TTB collects federal excise taxes on alcohol, tobacco, firearms, and ammunition, and assures compliance with federal tobacco and alcohol permitting, labeling, and marketing requirements to protect consumers. The purpose of the Tobacco Importer Risk Model (TIRM) is to identify tobacco import anomalies warranting TTB field attention. The model identifies both civil revenue issues and potential criminal activities through the analysis of tobacco import transactions. The model uses specified risk criteria to identify companies that may be importing tobacco products illegally or to identify importers that may be involved in tax evasion. The risk criteria are based on TTB field experience and knowledge of the industry and schemes to evade tax.[50]

(B) ***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.***

The TIRM uses risk factors based on predefined trends and patterns to identify companies' import transactions that may violate the law. Each risk factor is weighted by relative perceived risk and summarized to produce an overall risk score for each importer. The model also allows users to create customized risk-weighted queries based on user-defined data parameters. TTB field staff is able to view detailed information for all tobacco importers sorted by risk score, including the Department of Homeland Security's Customs and Border Protection (CBP) entry data and TTB permit information.

(C) ***Data sources that are being or will be used.***

The risk model uses four data sources: CBP's Automated Commercial System (ACS) Import Data, CBP's Automated Commercial Environment (ACE) Import Data, and two internal TTB systems (Integrated Revenue Information System (IRIS) and AutoAudit). ACS and ACE provide tobacco import information such as type of tobacco products, taxes, quantity type of import transaction, port of entry, and date of import. IRIS records excise tax and permit data and Auto Audit tracks historical case information.

(D) ***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.***

The leads identified by the TIRM require subsequent research and audit work to verify the efficacy of the risk scoring information. The results of the model are measured in terms of

[50] TIRM appears in Treasury's data mining report for the first time in 2013. Although TIRM was in initial development before 2013, it had limited data sources and functions.

correctly identifying potential criminal or fraudulent activity or revenue deficiencies. In FY 2013, TIRM led to positive results in 10 of 16 cases.

(E) ***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 or will be taken with regard to the property, privacy, or other rights or privileges of any individual as a result of implementing the data mining activity.***

The impact or likely impact of the TIRM on privacy and civil liberties of individuals is minimal. The TIRM analyses of existing data are not subject based but are based on identified risk criteria to locate anomalies that warrant field attention. The analysis of a data subject is limited to existing taxpayer data provided by tobacco importers. All taxpayer data are considered private and confidential and protected from disclosure by 26 U.S.C. § 6103, as described in the previous section of this report.

Further, TTB does not take adverse actions based solely on TIRM analyses. The purpose of the TIRM is limited to identifying entities that may merit further attention, such as an audit or investigation. Any subsequent audit or investigation is governed by existing laws and administrative procedures, policies, and controls that ensure due process for criminal and administrative investigations.

System operators are notified of the requirements and legal consequences of accessing the TIRM when they log in. The message states:

> 26 U.S.C. 6103 Data Warning. Information contained in this report is tax return information protected from disclosure by 26 U.S.C. 6103. By accessing this report, you hereby certify that your official duties require you to inspect such information for tax administration purposes.

(F) ***A list and analysis of the laws and regulations that govern the information being collected, reviewed, gathered, analyzed, or used in the data mining activity.***

Importation of tobacco products, cigarette papers and tubes, and processed tobacco is regulated by TTB pursuant to its statutory authority in Chapter 52 of the Internal Revenue Code (26 U.S.C.) and applicable agency regulations.[51] The authority to collect excise taxes on imported alcohol and tobacco products was originally retained by the Secretary of the Treasury through the Homeland Security Act of 2002.[52] Through Treasury Order 100–16, the Secretary of the Treasury delegates authority over customs revenue functions to the Secretary of the Department of Homeland Security. The Homeland Security Act of 2002 defines "customs revenue functions" as "[a]ssessing and collecting customs duties (including antidumping and countervailing duties and duties imposed under safeguard provisions), excise taxes, fees, and

[51] 27 C.F.R. pt. 41.

[52] *See* 6 U.S.C. §§ 212 and 215.

penalties due on imported merchandise, including classifying and valuing merchandise for purposes of such assessment."[53]
TTB is authorized to access data within CBP data systems necessary to fulfill its statutory mission.[54] TTB is working in conjunction with CBP to fulfill its statutory mission as it relates to imported products subject to various taxes and to ensure taxpayers understand their tax responsibilities related to these products. Cooperation among federal agencies will accommodate the collection of data as it relates to imported commodities subject to federal taxes, including but not limited to, retail, excise, manufacturers, and environmental taxes. In addition to the Federal Alcohol Administration Act, TTB enforces laws on the taxation and importation of tobacco products found in Chapters 51, and 52 of the Internal Revenue Code (IRC) and regulations issued pursuant to the IRC found in 27 CFR. TTB is also responsible for administering and enforcing the provisions of Chapters 61 through 80 of the IRC. This includes any issues related to taxes imposed by IRC sections 4181 or 4182, Chapter 51, or Chapter 52.

Insofar as the data analyzed by the TIRM consist of taxpayer information, the use of all tax related data is governed by 26 U.S.C. § 6103. Subsection (a) sets out the general rule of confidentiality. Subsection (b) sets forth definitions of terms commonly used throughout Section 6103. Subsections (c) through (o) contain exceptions to the general rule of confidentiality. The use of confidential commercial, financial, or trades secrets information is governed by the Trade Secrets Act, 18 U.S.C. § 1905, which prohibits the unlawful disclosure of this information by any federal official, employee, or contractor

(G) ***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:***

(i) ***Protect the privacy and due process rights of individuals, such as redress procedures.***

The TIRM does not determine whether a person or entity is going to be subject to administrative enforcement action or criminal prosecution. Any audit or investigation that is initiated based, in part, upon data from the TIRM is governed by the laws, administrative procedures, policies, and controls that govern criminal investigations or any other ensuing actions.

Information generated and accessed by the TIRM is protected by internal controls[55] that limit TIRM access to persons whose official duties require inspection of such information for tax administration purposes. The information is further protected by 26 U.S.C. § 6103 governing the confidentiality of return, return information or taxpayer return information and the Trade Secrets

[53] 6 U.S.C. § 215(a) (1).

[54] Multiple authorities support TTB's mission: the Homeland Security Act of 2002, Pub. L. No. 107-296; Executive Order 13439, *Establishing and Interagency Working Group on Import Safety*, July 18, 2007; the Internal Revenue Code of 1986; and the Federal Alcohol Administration Act, 27 U.S.C. chapter 8.

[55] All of TTB's information collections go through the OMB process and any forms that request personal information include a Privacy Act statement. In addition, TTB's privacy policy is posted on TTB.GOV (http://www.ttb.gov/about/privacy_policy.shtml) and is referenced from TTB's Online Applications. TTB's systems of record notice can be found in the *Federal Register* at http://www.gpo.gov/fdsys/pkg/FR-2011-12-01/pdf/2011-30898.pdf.

Act, 18 U.S.C. § 1905, which protects confidential commercial, financial, or trade secrets information collected by the federal government.
TIRM users receive training in a number of ways to ensure proper handling of information available via the TIRM. Users receive TIRM system demonstrations and have access to a user guide. Field Operations staff receive initial, basic 26 U.S.C. § 6103 and disclosure training. A government warning is displayed when accessing the TIRM, reinforcing handling and disclosure restrictions. In addition, all TTB employees complete annual privacy and cybersecurity awareness training. Finally, system sponsors and IT staff supporting development, maintenance, and operations of IT Systems are required to take additional specialized security training each year.

(ii) ***Ensure that only accurate and complete information is collected, reviewed, analyzed, or used and guard against any harmful consequences of potential inaccuracies.***

The TIRM relies on information collected through systems that have their own accuracy-related checks and balances. TTB does not rely solely on information gathered through the TIRM to take any adverse action against any individual or entity. Rather, the TIRM is the first step in gathering data that TTB then verifies through subsequent research and audits of tobacco importers before it takes any adverse action.

All data sets associated with TTB's systems, including the TIRM, are documented and managed using the TTB Systems Development Life Cycle (SDLC). Checks and balances are inherent with the data correction process ensuring different teams handle different steps of the effort and includes oversight by the Quality Assurance (QA) team. When the system owner identifies inconsistencies with data, a data correction process managed by the TTB Office of the Chief Information Officer QA team may be initiated. All changes are documented via the Request for Change process managed by the configuration management team. Work orders track the correction through its life cycle from request to development and through implementation along with confirmation of successful completion by the system owner. The process includes specific identification of the data to be corrected along with the rationale for the change. SDLC artifacts (e.g., database scripts) supporting data corrections conform to Data Management (DM) standards required by TTB's SDLC and are documented in TTB's Data Management Handbook. Analysis, development, and testing by the software maintenance team are verified through a quality review process conducted by the DM team to ensure the data correction is thoroughly documented. Once the DM team has approved the data correction, the operations team executes the correction, which is verified by the system owner.

The Memorandum of Understanding with CBP contains language that both parties will notify one another if data issues are discovered. Also, the ACS and ACE data import processes in support of the TIRM were documented and tested using TTB's SDLC.

CONCLUSION

The Department of the Treasury is pleased to provide to Congress its report on Treasury data mining activities. Congress authorized the Department to engage in data mining in furtherance of the Treasury mission while ensuring the protection of privacy. OPTR has reviewed the programs described in this report and worked with the reporting bureaus to implement the necessary privacy protections. OPTR will continue to provide oversight of all Treasury programs and systems, including those involved in data mining.