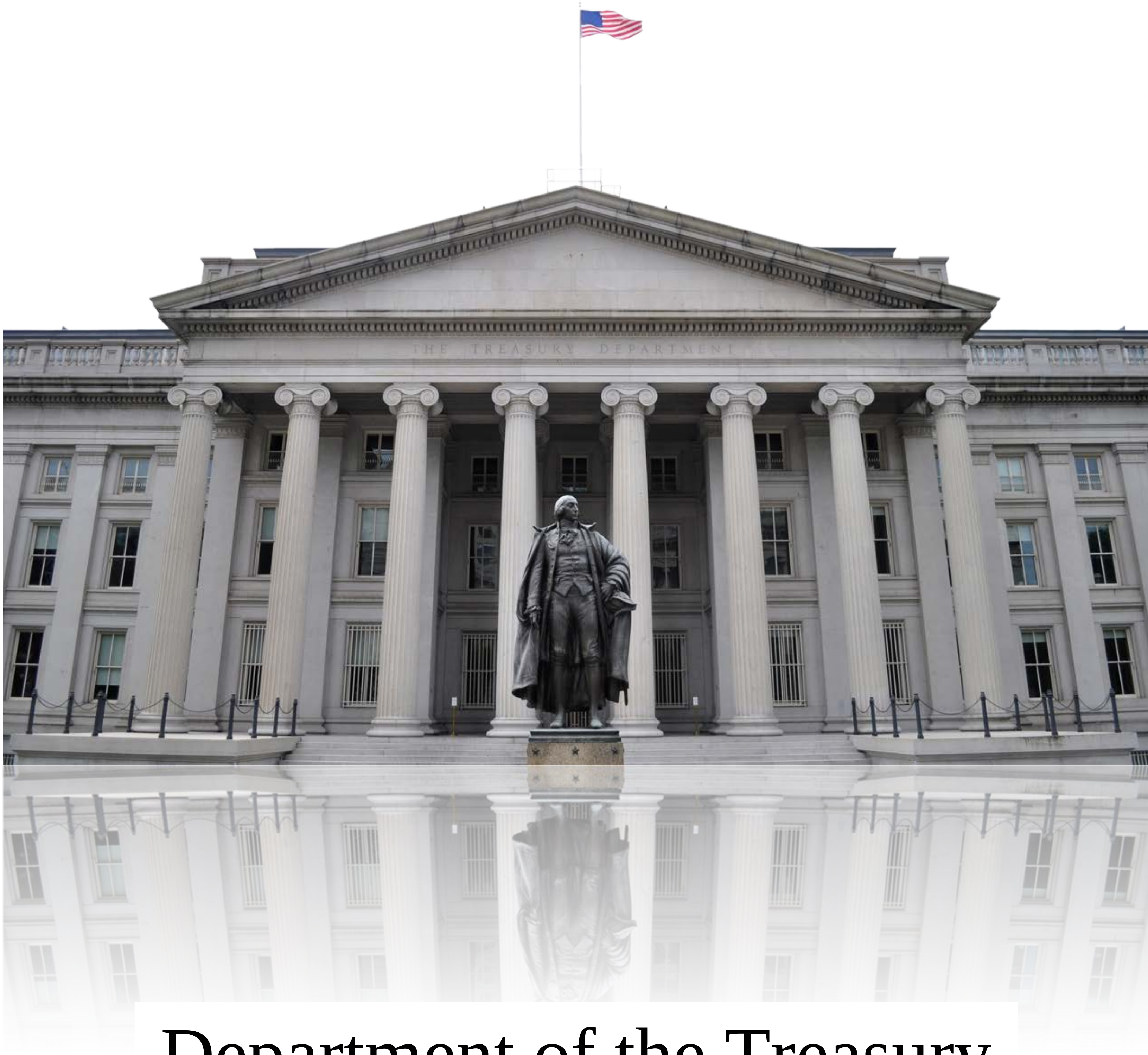

Department of the Treasury

2014 Annual Privacy and Data Mining Reports

Message from the Deputy Assistant Secretary for Privacy, Transparency, and Records

On behalf of the Department of the Treasury Senior Agency Official for Privacy and Chief Privacy and Civil Liberties Officer, I am pleased to present Treasury's Annual Privacy and Data Mining Reports for Fiscal Year 2014, as required by Section 522 of the Consolidated Appropriations Act of 2005 and the Federal Agency Data Mining Reporting Act of 2007. For the second year in a row, Treasury is combining these two separate reporting requirements into a single report.

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.

Helen Goff Foster
Deputy Assistant Secretary
for Privacy, Transparency, and Records
U.S. Department of the Treasury

2014 Annual Privacy and Data Mining Reports

TABLE OF CONTENTS

STATUTORY REQUIREMENTS

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 REPORTING PERIOD

This report covers Treasury activities within the 2014 fiscal year (the reporting period).

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 2014 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 his 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]
- 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]

[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).

- 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 the Department 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 the Department has adequate procedures to receive, investigate, respond to, and redress complaints from individuals who allege the Department has violated their privacy or civil liberties;[15]
- considering certain factors when providing advice on Department 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]
- 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]

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

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

Treasury Directive 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 him in his capacity as the Department's CPCLO.

OPTR supports Treasury's mission 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 applicable federal statutes, Executive Orders, OMB memoranda and guidance, as well as other relevant policy, standards, and regulations. Some of OPTR's operations include:

- Ensuring consistent application of privacy and civil liberties safeguards in all Treasury activities through Treasury-wide policy and oversight;
- Reviewing and publishing system of records notices;
- Conducting and reviewing Privacy Impact/Threshold Assessments;
- Analyzing and reporting on paper and electronic incidents involving PII; and
- Developing and conducting privacy and civil liberties training.

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.

[20] *Id.*

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 a SORN 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]

During FY 2014 the Department published three new SORNs in the *Federal Register*: Treasury .016, *Department of the Treasury Reasonable Accommodations*, Bureau of Engraving and Printing .049, *Tour Scheduling*, and Fiscal Service .023, *Payment Verification Records*. On January 2, 2014, the Departmental Offices republished systems of records notices in the *Federal Register* at 79 Fed. Reg. 209, http://www.gpo.gov/fdsys/pkg/FR-2014-01-02/html/2013-31356.htm. This publication included DO .120, *Records Related to Office of Foreign Assets Control Economic Sanctions;* DO .191, *Human Resources and Administrative Records System*; DO .196, *Treasury Information Security Program*; DO .218, *Making Home Affordable Program*; DO .225, *Troubled Asset Relief Program Fraud Investigation Information System*; and DO .226, *Validating EITC Eligibility with State Data Pilot Project Records*. The notices were last published in their entirety on April 20, 2010, beginning at 75 FR 20675.

In addition, during FY 2014 Treasury terminated one system of records, Treasury .008 *Emergency Management System* (75 FR 54435). Records from this system of records were retained and disposed of in accordance with the appropriate National Archives and Records standards. Files were purged in accordance with Treasury Directive 80-05, ‘‘Records and Information Management Program.’’

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) documents information safeguarding practices while 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 2014, Treasury reviewed 281 PIAs. Pursuant to the E-Gov Act, 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.

[21] *See* 31 C.F.R. §§ 1.20-1.36.

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 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 2014, the Department reported a total inventory of 280 FISMA systems containing personal information in identifiable form.

Number of FISMA Systems containing PII	Percentage Requiring SORN	Percentage Requiring PIA
280	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 Treasury Directive (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 2014, Treasury performed 192 reviews (excluding Social Security Number Redactions and Eliminations), 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	FY 2014
Privacy Threshold Assessments/Privacy Impact Assessments	109/17
System of Records Notices and Routine Uses	49
Section 508	86%
Social Security Number Redaction/Elimination	59
Computer Matching Agreement Program	2
Treasury Requested Scan	5

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 2013 to comply with OMB Memorandum 07-16, *Safeguarding Against and Responding to the Breach of Personally Identifiable Information*, (OMB M-07-16).

ELIMINATION OF THE UNNECESSARY USE OF SOCIAL SECURITY NUMBERS

In FY 2014, the IRS launched the third phase of its Social Security Number (SSN) Elimination Project. This phase included two major events: the masking of SSNs on payment notices and the addition of two dimensional (2D) barcodes to four types of installment agreement notices. This will affect an estimated 33 million notices annually. The 2D barcode will provide the IRS with information needed to process notices and mask the full display of the SSN in the body of the notices by exposing only the last four digits of the SSN. In addition, the IRS enhanced its payment-processing systems to accommodate the 2D barcode in notices processed by these systems.

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. Ninety nine percent of all Treasury employees completed annual privacy awareness training during the reporting period.

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 during the reporting period. This review resulted in an updated version of the departmental privacy training course, entitled "A Culture of Privacy Awareness."

LEADERSHIP AND COORDINATION WITHIN TREASURY

Treasury Order 102-25: Delegation of Authority Concerning Privacy and Civil Liberties

In July 2014, OPTR completed a revision of Treasury Order (TO) 102-25, *Delegation of Authority Concerning Privacy and Civil Liberties*. The revised TO formally designates the Assistant Secretary for Management as the "Senior Agency Official for Privacy" and the "Information Sharing Environment (ISE) Privacy Official," pursuant to OMB M-14-04 and related implementation guidance in National Institute for Standards and Technology Special Publication 800-53, Appendix J. TO 102-25 is available at: http://www.treasury.gov/about/role-of-treasury/orders-directives/Pages/to102-25.aspx. Similarly, Treasury Directive 25-04, *The*

Privacy Act of 1974, As Amended, updated in January 2014, provides Treasury guidance regarding the responsibilities of the Senior Agency Official for Privacy, and is available at: http://www.treasury.gov/about/role-of-treasury/orders-directives/Pages/td25-04.aspx.

Treasury Directive 85-01: Department of the Treasury Information Technology (IT) Security Program

In the fourth quarter of FY 2014, OPTR worked with the Office of the Chief Information Officer to include a full list of privacy controls in an updated version of TD 85-01, *Department of the Treasury Information Technology (IT) Security Program*. This revision was prompted by an update of NIST SP 800-53, which provides recommended security controls for federal information systems. The revised publication will add a new set of privacy controls in a new appendix, titled Appendix J. The current version of TD 85-01 is available at: http://www.treasury.gov/about/role-of-treasury/orders-directives/Pages/td85-01.aspx.

Do Not Pay

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 the Improper Payment Elimination and Recovery Improvement Act of 2012 (IPERIA). Treasury published a SORN for the system in the *Federal Register* on January 2, 2014. It is available at: https://www.federalregister.gov/articles/2014/01/02/2013-31356/privacy-act-of-1974-systems-of-records. As the Privacy Act requires,[22] notice of the new system of records was provided to the House of Representatives Committee on Oversight and Government Reform, the Senate Committee on Homeland Security and Governmental Affairs, and OMB.

In accordance with 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.

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

[22] 5 U.S.C. § 552a(r).

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. The consolidated report is available here: http://www.dhs.gov/sites/default/files/publications/2014-privacy-and-civil-liberties-assessment-report.pdf.

Treasury's Privacy Website

In FY 2014, OPTR completed substantial improvements to the privacy 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 2014, 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] This included developing and implementing a first-of- its- kind procedure for foreign nationals to request access to, and correction of, information maintained by Treasury under the Agreement. The TFTP Agreement allows Treasury to use EU-stored data to track terrorism financing while protecting the privacy interests of EU citizens. During the reporting period, OPTR also supported the Third Joint Review of the TFTP Agreement, by which EU and U.S. officials review the TFTP program to ensure compliance.

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

In FY 2014, 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

[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

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

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 2014, the Treasury DIB reviewed and approved one new computer matching program, one extension of a computer matching program, and renewed one of the Department's ongoing computer matching programs. A new matching agreement may exist for eighteen months; this agreement can be extended for an additional 12-months. After and extension expires, an agreement may be renewed for an additional eighteen months.

Published notices for all of Treasury ongoing computer matching programs are available online through the Treasury Privacy Act page at: http://www.treasury.gov/privacy/computer-matching-programs. The DIB's actions included:

Agreement Title	Agencies Involved	Action	Date of Action
Project 241	IRS-Department of Health and Human Services Centers for Medicare & Medicaid Services (CMS)-Social Security Administration (SSA)	Approved	December 13, 2013
Medicare Secondary Payer Program	IRS-SSA-CMS	Renewal	December 4, 2013
Taxpayer Address Request Program	IRS-Department of Education	Extension	January 24, 2014

SECTION TWO:
DEPARTMENT OF THE TREASURY 2014 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 803 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). These IRS, FinCEN, and TTB systems were discussed in previous Treasury data mining reports.

FINCEN 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's mission is to safeguard the financial system from illicit use, combat money laundering, and promote national security through the collection, analysis, and dissemination of financial intelligence and the strategic use of financial authorities. To meet its mission, FinCEN provides 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 provides "computer and data support and data analysis to the Secretary of the Treasury for tracking and controlling foreign assets."[26]

[25] 42 U.S.C § 2000ee-3(b)(1). "[T]elephone 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" are not "databases" under the Act. § 2000ee-3(b)(2).

[26] 31 U.S.C. § 310(b)(2)(E), (G).

FinCEN's analysts use various data searching and querying techniques and trend-spotting algorithms for "lead generation," that is, locating groups of subjects or institutions whose activities, as revealed by the algorithms, warrant outreach, investigation, or other statutorily mandated activities. For its data analysis activities, the analysts leverage FinCEN's Advanced Analytics system. This system is comprised of Commercial Off-the-Shelf (COTS) and custom-developed analytical and search tools with capabilities including social network analysis, modelling and visualization, link analysis, and geospatial analysis of Bank Secrecy Act (BSA) data.

FinCEN successfully developed algorithms designed to identify activity associated with specific types of financial crimes, such as check cashing activity associated with health insurance fraud. FinCEN also uses Advanced Analytics text mining capabilities in conjunction with structured field searches to examine filing patterns in financial sectors, which can reveal a spectrum of data from broad trends in criminal activity to flagging specific institutions for regulatory attention.

FinCEN recently began to develop the use of automated daily business rules to rapidly surface potential high value illicit financial activity. BSA filings that are flagged as a product of the business rules are then reviewed internally at FinCEN and distributed to law enforcement partners if applicable. These rules are vital when timely information is of the essence, such as in identifying terrorist financing. FinCEN plans to expand these activities in FY 2015.

(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 leverages two principal methods for deriving information relevant to illicit activity from the BSA data. The first is the content driven method. That is, searching for specific entity names, or term combinations used in reporting that are associated with various types of illicit activity. The second is the pattern driven method, which may take various forms. Patterns may be derived by searching for a particular type of subject and then identifying other subjects that fit that pattern and also may have certain filing profiles. Matching filing patterns across different types of statutorily required BSA reports highlights anomalous behavior and leads to the identification of new investigation subjects.

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

BSA reports administered by FinCEN, e.g., a report by a financial institution of a suspicious transaction relevant to a possible violation of law or regulation,[27] 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

[27] 31 U.S.C. § 5318(g).

information, and third party data sources, such as the Census Bureau, Federal Reserve, Social Security Administration ,[28] and Office of Foreign Assets Control data. This analysis can 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 four distinct audiences: law enforcement, including FinCEN's foreign Financial Intelligence Unit (FIU) partners; financial regulators; the financial industry; and the general public. Each of these sets of consumers has different restrictions or guidelines under which FinCEN can provide BSA-data or BSA-derived analysis.

For example, in FY 2014, FinCEN produced a total of 178 intelligence products for law enforcement partners and responded to 878 requests for BSA information from foreign FIU partners. For domestic and foreign law enforcement partners, it is FinCEN's assessment that the value of its data analytics is high. FinCEN annually receives the results of surveys of its foreign Egmont member counterparts and domestic law enforcement with respect to the usefulness of its analytic products. For FY 2014, these survey results are consistent with prior years, averaging in the mid-80th percentile. FinCEN also receives feedback on individual reports from law enforcement. Examples of some of these feedback reports are condensed below. Additionally, FinCEN has recently begun to distribute an analytical report based on output from its daily business rules and provides that to the requisite law enforcement agencies. Preliminary feedback suggests that these products are valued for their timeliness and are providing relevant information not previously uncovered.

More specifically, FinCEN receives 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 support of U.S. law enforcement counter-narcotics initiatives on the Southwest Border by employing advanced technology to build processes to more 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.

[28] 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.

- FinCEN queried Suspicious Activity Reports (SARs) related to significant tax evasion cases under investigation by another federal law enforcement agency and was able to provide data relevant to over 105 institutions in jurisdictions of interest—over 35,000 discrete BSA reports—in a sortable spreadsheet, within a week of the request and in time to meet a key litigation deadline.
- Since 2009, FinCEN has also been performing searches of SAR data that reflect suspicious use of proceeds from government financial support programs, such as the Trouble Asset Relief Program (TARP). In FY 2014, FinCEN provided the Special Inspector General for TARP with approximately17,388 SARs referencing 22,390 subjects of possible interest that reported more than $117 billion in suspicious financial activity.

Finally, FinCEN also provides, usually annually, aggregated statistics regarding SAR data by sector to the public in a publication currently titled "SAR Stats." The most recent version of SAR Stats was published on FinCEN's website in June 2014, and the latest data indicate that over 175,000 visits to SAR Stats have been made by the public readers since June 2014—an indication of the high public interest in the information.

(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."[29] Moreover, the Right to Financial Privacy Act of 1978[30] 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.[31]

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.

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

[30] 12 U.S.C. § 3401, *et seq.*

[31] 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.")

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

FinCEN publishes Privacy Impact Assessments (PIA) which informs the public of FinCEN activities and practices related to the collection, processing, retention, and distribution of the personally-identifiable information (PII) that FinCEN handles in its systems. The PII handled by these systems is necessary to properly assist regulators and law enforcement in identifying and monitoring the financial activities of individuals who are potentially committing financial crimes. FinCEN PIAs can be found at http://www.fincen.gov/foia/pia.html.

(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

[32] 5 U.S.C. § 552a.

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."[33] 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."[34] 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.

[33] 5 U.S.C. § 552a(a)(3) (defining a "record" to mean 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 and a "system of records" to mean 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);
[34] 5 U.S.C. § 552a(b)(3).

FinCEN has three SORNs that cover the information it collects under the BSA: (1) Treasury/FinCEN .001, *FinCEN Investigations and Examinations System*[35], (2) Treasury/FinCEN .002, *Suspicious Activity Report (SAR) System*[36], and (3) Treasury/FinCEN .003, *Bank Secrecy Act (BSA) Reports System.*[37]

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

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.[38] 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.[39] FinCEN is required to "analyze and disseminate" the data for a broad range of purposes consistent with the law.[40] 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."[41] 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.[42] The activities and procedures described in this report adhere to the requirements of this statute.

[35] 79 Fed. Reg. 20969 (April 14, 2014).
[36] *Id.* at 20972.
[37] *Id.* at 20974.
[38] Treasury Order 180-01, *Financial Crimes Enforcement Network* (July 1, 2014) (delegating to the Director of FinCEN various duties and responsibilities, including the authority to administer, implement, and enforce the BSA).
[39] 31 U.S.C.§ 310(b)(2)(B)
[40] *Id*. at § 310(b)(2)(C)(i)-(vii).
[41] *Id.* at § 310(b)(2)(E), (G).
[42] *Id*. at § 310(c)(1) and (c)(2).

(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. Because user agencies only use BSA information in conjunction with other evidence, a BSA 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.

FinCEN has mandatory training for its data users that includes secure handling and safeguarding of the information. FinCEN provides on-line training for all external users as a requirement for system access. Biennially, at a minimum, users must complete training as a requirement of continued system access. In addition to this online training, FinCEN hosts webinars as requested. All FinCEN staff are required to annually complete Privacy Awareness training which includes the staff's civil liberties responsibilities. Materials for each of these courses prominently articulate accountability for the security and confidentiality of the BSA data and its handling.

IRS 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.

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 indicators of potential criminal activity:

- Reveal;
- Investigative Data Analytics (IDA);

- Electronic Fraud Detection System (EFDS); and
- FinCEN Query

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),[43] Scheme Development Centers (SDC),[44] and field offices all use the system to identify and develop leads for refund frauds, 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.

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.

IRS-CI and now SB/SE users access the FinCEN Query system (see FinCEN report) as the system of record for BSA data. Until April 30, 2014, SB/SE used Web-CBRS for BSA data.

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

[44] 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.

Web-CBRS is a web-based application that accesses a database containing BSA forms and information. Currently, this system has very limited use with fewer than 50 active SB/SE users viewing Excise documents only.

(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

Reveal and IDA 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[45] and Questionable Refund Program.[46] 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) A thorough description of the data sources that are being 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 Modernized e-File (MeF)) or a paper filed tax return.

[45] 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.

[46] 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.

- **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.

(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 data uncovered during the query searches are only leads, and additional investigative steps are required for quality verification. There are 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 detection. 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 1:1.4 for FY 2014. This means that the data mining program accurately predicts fraudulent returns in 10 of 14 cases.

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

(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 or individuals 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 Internal Revenue Manual security and privacy 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.

(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.

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 leads and additional investigative steps are required 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. FinCEN' s 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 are used for investigative analysis purposes under Internal Revenue Manual standards and guidelines. The data uncovered during query searches are only used as leads, and additional investigative steps are required to verify the quality of the information. Therefore, IRS-CI uses these 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) A thorough description of the data mining activity, its goals, and, where appropriate, the target dates for the deployment of the data mining activity.

TTB's analytics program performs three types of activities that, together, qualify as data mining as defined by the Federal Agency Data Mining Reporting Act of 2007:

1. Queries of commercial transactions recorded by tax and trade databases maintained by TTB and other federal agencies;
2. Searches of public records and law enforcement databases for indications of illicit dealings;
3. Link analysis of connections between businesses and individuals.

The activities are conducted primarily for the purpose of discovering or locating patterns or anomalies indicative of activity that may violate federal regulations administered by TTB. The data used in these activities are, for the most part, gathered with queries of registered individuals or businesses. However, subsequent analysis of the data is primarily pattern-based, seeking anomalies in compiled records. The data mining activities also include some queries and searches that are solely pattern-based, e.g., queries of all tobacco product imports over a given time period.

The goals of TTB's data mining activity are to 1) automate routine oversight processes, and 2) improve detection of compliance violations. The activity supports predictive models and business intelligence that identifies compliance risks and potential fraudulent or criminal activity, that may be subject to further field review and action. TTB has predictive models in place that score the risk of tax diversion in the tobacco industry, and evaluate businesses seeking a TTB permit. TTB also compiles business intelligence that highlights patterns in tax and trade data. These models and business intelligence are in regular use today with improvements and expansions planned for fiscal year 2015.

(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.

TTB uses commercially available data mining technologies to access and analyze information. The experience of intelligence analysts and investigators provides the basis for determining whether a particular pattern or anomaly is indicative of compliance violations. The ability to identify patterns and anomalies is supplemented with statistical analysis and machine learning techniques.

Most data mining is conducted with a combination of SAS statistical analysis software and Oracle relational database systems. Data are retrieved with SAS data step programming and/or Structured Query Language (SQL) queries. Data fields are transformed with procedures that aggregate, correlate, cluster, and otherwise simplify available variables. The procedures include parsing unstructured text for entity extraction and topic modeling to find similarities between text documents.

Once data are collected and transformed, predictive models use the data to estimate the expected violation risk of a particular individual, business, or incident. The estimates are based primarily on business rules and templates defined by experienced analysts (and implemented in the SAS programming language). Analysts also are discovering new patterns based on analysis of historical violations using machine learning classification algorithms such as logistic regression, decision trees, and neural networks. Patterns identified through these methods are vetted with experienced analysts and evaluated against randomized test cases.

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

TTB uses data from its own databases, the databases of other federal agencies, and commercial data providers. The data sources include:

Internal Data:
1) Integrated Revenue Information System (IRIS) – tax data submitted by TTB industry members;
2) Permits Online (PONL) –application data from businesses requesting a TTB permit;

3) AutoAudit – data from TTB's audits and investigations;

External Data:
4) Automated Commercial Environment (ACE) / Automated Commercial System (ACS) / Automated Export System (AES) – data regarding imports and exports of products regulated by TTB;
5) Financial Crimes Enforcement Network Query (FinCEN Query) – data submitted in compliance with the Bank Secrecy Act transcripts such as Suspicious Activity Reports (SARs), Currency Transaction Reports (CTRs), etc.;
6) LexisNexis Accurint – public records data of court proceedings (including some criminal cases), property holdings, licenses, and registrations.

These databases are in use today and further integration of the sources is ongoing, as is identification of new data sources.

(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.

TTB's data mining activity is valuable for 1) automating certain routine oversight processes, and 2) improving detection of compliance violations. Initial evaluations indicate that data mining enables more regular oversight and produces indicators for further field review, including investigation and audit. This evaluation is continuing and generating new improvements as the data mining activity matures.

The data mining activity, models and business intelligence supported by the activity have been effective at helping to automate oversight processes. Predictive models automatically screen approximately 1,200 registered tobacco businesses and 800 new original permit applications every month. The models verify information and monitor patterns in operations, tax payments, and international trade activity. The models also automatically monitor financial and trade databases for indications of activity by unregistered businesses. Automating basic screens enables TTB to provide oversight to a wider section of its regulated industries.

The ability of predictive models to detect compliance violations depends greatly on the accuracy of the source data available for the models. Data quality and mining techniques continue to improve with increased use and scrutiny of data. In the meantime, predictive models that rely on data mining activity are showing promise in detecting compliance violations. The Tobacco Importer Risk Model has been in use for two years and demonstrated precision of approximately 0.60, i.e., 6 out of 10 cases recommended by the model result in the detection of previously undisclosed tax liability that also provided a positive return on investment. Initial evaluations of the Tobacco Diversion Model show precision measures above 0.80, i.e., 8 out of 10 leads uncover a compliance issue though not necessarily an undisclosed tax liability. Evaluation of these and other models will continue as part of TTB's ongoing effort to improve model accuracy.

(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 or individuals as a result of implementing the data mining activity.

TTB's data mining activity has little impact on the privacy and civil liberties of individuals. Insights gained from the activity only result in actions against property, or the privilege to operate in regulated industries, after thorough review by experienced specialists with oversight authorities mandated by federal laws and regulations. The data sources mined are also limited to include only tax records, commercial records, and law enforcement records authorized for use in oversight and enforcement.

Any data concerning individuals or businesses are vigorously protected against unauthorized use and disclosure. Policies and procedures prohibit the search of any database for reasons other than providing authorized oversight or enforcement. In cases when patterns in data are thought to be indicative of compliance issues the data and circumstances are carefully reviewed by experienced staff before any adverse action is taken. TTB also continues to protect data against any unauthorized disclosure through all investigation and enforcement actions.

Data gathered in data mining activities are considered private and confidential and protected from disclosure by 26 U.S.C. § 6103. TTB governs its use of these data consistent with the authorities described in that statute. Privacy protections are further assured by additional laws that provide for civil and criminal penalties for any unauthorized disclosure of taxpayer data. There are criminal penalties including: (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.

(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.

TTB administers the provisions of the Internal Revenue Code (IRC) relating to distilled spirits, wine, and beer (26 U.S.C. ch. 51), firearms and ammunition excise taxes (26 U.S.C. §§ 4181, 4182, and related portions of ch. 32), and the general rules of tax procedure with respect to these commodities (including related criminal provisions at 26 U.S.C. ch. 68 and 75). In addition, TTB administers the Federal Alcohol Administration Act (27 U.S.C. ch. 8, subchapter I), which covers basic permits, unfair trade practices, and labeling and advertising of alcohol beverages; the Alcoholic Beverage Labeling Act of 1988 (27 U.S.C. ch. 8, subchapter II), which requires a specific "Government Warning" statement on alcohol beverage labels; and the Webb-Kenyon Act (27 U.S.C. §§ 122-122b), which prohibits the shipment of liquor into a state in violation of state law.

The IRC establishes qualification criteria to engage in the businesses relating to manufacturing and importing or exporting tobacco products, and manufacturing or importing processed tobacco, and require that persons obtain permits to engage in these activities. 26 U.S.C. § 5713. A permit qualification requirement also applies to the production of distilled spirits and wine, as well as to the wholesaling and importation of all beverage alcohol products. 26 U.S.C. §§ 5171(c) and (d), 5271; *see also* 27 U.S.C. § 201, *et seq*.

Through an agreement with FinCEN, dated May 3, 2005, TTB is granted direct electronic access to data collected pursuant to provisions of the Bank Secrecy Act, 31 U.S.C. § 5311, *et seq*. The direct access is granted for tax or regulatory purposes relevant to the mission of TTB.

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. *See* 6 U.S.C. §§ 212 and 215. 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. "Customs revenue functions" are defined by the Homeland Security Act of 2002 as "[a]ssessing and collecting customs duties (including antidumping and countervailing duties and duties imposed under safeguard provisions), excise taxes, fees, and penalties due on imported merchandise, including classifying and valuing merchandise for purposes of such assessment." 6 U.S.C. § 215(a) (1).

TTB is authorized pursuant to the Homeland Security Act of 2002, Pub. L. 107-296; Executive Order 13439, July 18, 2007; the Internal Revenue Code of 1986 (IRC); and the Federal Alcohol Administration Act, 27 U.S.C. chapter 8 (FAA Act) to access data within Customs and Border Protection data systems necessary to fulfill its statutory mission. 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. Cooperative efforts across federal agency lines 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.

Insofar as the data analyzed by the models 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) of Section 6103 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) 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.

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

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

Information generated and accessed by the data mining activities is protected by internal controls that limit 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 returns and return information and the Trade Secrets Act, 18 U.S.C. § 1905, which protects confidential commercial, financial, or trades secrets information collected by the federal government.

System operators are notified of the requirements and legal consequences of accessing predictive models in production. 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.

Users of predictive models in production receive training in the proper handling of information. Users receive system demonstrations of the model and have access to a user guide. The same process will be followed for future models when successful testing and evaluation has been completed. Field Operations staff receive 26 U.S.C. § 6103 and disclosure training. In addition, all TTB employees complete the annual Privacy Awareness and Cyber Security 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 data mining activities rely on information collected through systems that have their own accuracy related checks and balances. TTB does not rely solely on information gathered through predictive models to take any adverse action against any individual or entity. Rather, the models are the first step in gathering data and this information is verified through subsequent research and audits of companies and importers before any adverse action is taken.

All data sets associated with TTB's systems are documented and managed using the TTB Systems Development Life Cycle (SDLC). Checks and balances are inherent to the data correction process ensuring different teams handle different steps of the effort and include oversight by the Office of the Chief Information Officer Quality Assurance (OCIO QA) Team. When the system owner identifies inconsistencies with data, a Data Correction process managed by TTB's OCIO QA Team may be initiated. All changes are documented via the Request for Change process managed by the Configuration Management Team and work orders track the correction through its lifecycle (from request to development and through implementation), which includes confirmation of successful completion by the system owner. The process includes specific identification of the data to be corrected along with rationale for the change. SDLC artifacts (e.g., database scripts) supporting data corrections conform to Data Management (DM) standards. 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 Tobacco Importer Risk Model were documented and tested using TTB's SDLC.

CONCLUSION

The Department of the Treasury is pleased to provide to Congress its Annual Privacy and Data Mining Reports for Fiscal Year 2014. OPTR has reviewed the activities and programs described in these reports and continues to work closely with all Treasury bureaus and offices to protect individual privacy and civil liberties in all Treasury activities.