OFFICE OF THE DIRECTOR OF NATIONAL INTELLIGENCE

2014 Data Mining Report

LEADING INTELLIGENCE INTEGRATION

Table of Contents

PART I: INTRODUCTION

The Office of the Director of National Intelligence (ODNI) provides this report pursuant to Section 804 of the *Implementing Recommendations of the 9/11 Commission Act of 2007*, entitled *The Federal Agency Data Mining Reporting Act of 2007* (Public Law 110–53).

Scope

This report covers the activities of all ODNI components from January 1, 2014 through December 31, 2014. Constituent elements of the Intelligence Community (IC) will report their activities to Congress through their own departments or agencies.

Reporting Requirement

The *Federal Agency Data Mining Reporting Act of 2007* requires that, each year "the head of each department or agency of the Federal Government that is engaged in an activity to use or develop data mining shall submit a report to Congress on all such activities of the department or agency."[1]

The Act defines "data mining" as "*a program involving pattern-based queries, searches or other analyses of one of 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;[2] *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."*[3]

Report Content

In recent years, we have followed a format that we believe enhances clarity and readability. Specifically, Part II of the report describes activities that meet the definition of "data mining" under the Act as well as programs that meet some, but not all, of the criteria defining "data mining." We report the latter category of activities in the interest of transparency. Part III provides updates on programs included in the prior year's report. Part IV of this report provides an overview of the Privacy and Civil Liberties infrastructure within which ODNI conducts its activities.

[1] 42 U.S.C. § 2000ee-3(c)(1).

[2] As stated in prior reports, certain analytic tools and techniques, such as link-analysis tools, rely on "personal identifiers of a specific individual, or inputs associated with a specific individual or group of individuals," such as a known or suspected terrorist, or other subject of foreign intelligence interest, and use various methods to uncover links or relationships between the known subject and potential associates or other persons with whom that subject has a "link" (a contact or relationship). Such tools and techniques are not considered to meet the "data mining" definition of the Act.

[3] 42 U.S.C. § 2000ee-3(b)(1).

PART II: NEW ACTIVITIES

The ODNI has undertaken no new reportable activities in the current report period.

PART III: UPDATES ON PREVIOUSLY REPORTED ACTIVITIES

As previously discussed, this section provides updates on programs that were described in last year's report.

National Counterterrorism Center (NCTC) Threat Analysis

The NCTC continues to conduct "threat analyses" as described in the 2013 ODNI Data Mining Report. As noted in the 2013 report, this is an analytic technique to narrow the pool of information within NCTC databases that analysts will assess in response to specific threat reports. This technique does not meet all of the statutorily defined criteria for data mining, but has been reported in the interest of transparency.

Intelligence Advanced Research Project Activity (IARPA)

IARPA continues to invest in high risk/high payoff research programs that have the potential to provide the United States with overwhelming intelligence advantage over future adversaries. As an organization that only funds scientific research, IARPA does not use, nor does it expect to make use of, data mining technology for its own purposes. IARPA programs are by nature experimental and designed to produce new capabilities, such as those described here.

Knowledge Discovery and Dissemination (KDD) Program

The KDD program began in 2009 and completed its fourth and final period in December 2014. KDD tackled two significant technical areas: (1) how to quickly understand novel data sets so that the contents can be correctly integrated with data sets that are already in use (this is termed "alignment"); and (2) how to construct automatic analysis tools that are able to work effectively across multiple aligned data sets. The KDD program created and evaluated a variety of technologies, a number of which have been determined useful in settings across the IC.

Automated Low–level Analysis and Description of Diverse Intelligence Video (ALADDIN Video) Program

The ALADDIN Video scientific research program completed its fourth round of testing in the Fall 2014. The objective of the ALADDIN Video program is to enable analysts to query large video data sets to quickly and reliably locate those video clips that show a specific type of event, thus automating a triage process that is currently performed manually for the most part. Although not "data mining," technologies that result from ALADDIN Video research could potentially be applied by operational organizations to support capabilities that involve pattern recognition.

ALADDIN Video research addressed three significant technical areas: (1) high-speed processing of large numbers of video clips to extract information needed to support queries about each clip's contents; (2) generation of effective queries from small sets of example video clips and a textual description; and (3) robust query processing that identifies the clips of interest and summarizes the rationale for their selection. ALADDIN Video research results will be evaluated by IARPA and the National Institute for Standards and Technology.

Security and Privacy Assurance (SPAR) Program

The SPAR program completed its second and final phase of research in June 2014, building on prior successes in the area of distributed private information retrieval (PIR). As discussed in the 2009 and 2010 ODNI Data Mining Reports, PIR protocols permit an entity to query a cooperating data provider and retrieve only the records that match the query without the provider learning what query was posed or what results were returned.

The final phase of SPAR involved research in three technical areas: (1) development of protocols ensuring privacy and security compliance even in the context of a "private" query; (2) implementation of fully homomorphic encryption in the absence of any third parties; and (3) application of PIR to publish/subscribe systems.

In September 2014, IARPA began a SPAR pilot project to demonstrate the utility of the protocols in real use cases within the IC. The pilot project will implement SPAR database query protocols and SPAR publish-subscribe protocols in a classified network with real data. The pilot project is expected to be complete by Fall 2015.

SPAR protocols have the potential to enable the IC to access specific records without having to disclose classified data and without accessing, learning, ingesting, or retaining any private information about non-relevant persons. By satisfying security and privacy concerns, the technology could enable enhanced cooperative information sharing across the IC and with other parts of the Federal Government and the private sector.

(U) PART IV: PROTECTION OF PRIVACY AND CIVIL LIBERITES

The ODNI Civil Liberties and Privacy Office (CLPO) works closely with the ODNI Office of General Counsel, ODNI components, and the IC elements to ensure appropriate legal, privacy, and civil liberties safeguards are incorporated into policies, processes, and procedures that support the intelligence mission. The CLPO is led by the Civil Liberties Protection Officer, a position established by the *Intelligence Reform and Terrorist Prevention Act of 2004* (Public Law 108-458). The duties of this Officer are set forth in that Act, and include: ensuring that the protection of civil liberties and privacy is appropriately incorporated in the policies of the ODNI and the IC; overseeing compliance by the ODNI with legal requirements relating to civil liberties and privacy; reviewing complaints about potential abuses of privacy and civil liberties in ODNI programs and activities; and ensuring that technologies sustain, and do not erode, privacy protections relating to the use, collection, and disclosure of personal information. Before using any tool or technology in an operational setting, the use of the tool or technology would need to be examined pursuant to Executive Order (EO) 12333, the *Privacy Act of 1974*, and other applicable requirements to determine how the tool could be used consistent with the framework for protecting United States Person (USP) information.

The IC has in place a protective infrastructure built in principal part on a core set of USP rules derived from EO 12333. This EO requires each IC element to maintain procedures, approved by the Attorney General, governing the collection, retention, and dissemination of USP information. These procedures limit the type of information that may be collected, retained, or disseminated to the categories listed in part 2.3 of the EO. Each IC element's Attorney General-approved USP guidance is interpreted, applied, and overseen by that element's Office of General Counsel, Office of Inspector General, and other compliance offices as appropriate. Violations are reported to the Intelligence Oversight Board of the President's Intelligence Advisory Board. In addition to EO 12333, IC elements are subject to the requirements of the *Privacy Act of 1974*, which protects information about U.S. citizens and permanent resident aliens that a government agency maintains and retrieves by name or unique identifier. Going

forward, the IC will also conform to policies and procedures under Presidential Policy Directive 28, relating to protections for all personal information contained in signals intelligence.

The IC's privacy and civil liberties protective infrastructure is bolstered also by guidance and directives issued by the Office of Management and Budget, including memoranda regarding the reporting of and response to incidents involving personally identifiable information and the minimization of social security numbers.

Finally, the IC has developed and established two sets of principles that have been adopted as "foundational" to the IC mission: The Principles of Professional Ethics for the IC, and The Principles of Intelligence Transparency for the IC. These two sets of principles inform the IC's approaches to applying appropriate protections for the types of activities described in this report.