OFFICE OF THE DIRECTOR OF NATIONAL INTELLIGENCE

2019 Data Mining Report

For the Period of 01 January 2019 through 31 December 2019

TABLE OF CONTENTS

INTRODUCTION

The Office of the Director of National Intelligence (ODNI) provides this report pursuant to the *Federal Agency Data Mining Reporting Act of 2007*, section 804 of Public Law 110-53 (codified at Title 42 United States Code section 2000ee-3) (the "Data Mining Reporting Act" or the "Act").

A. Scope

This report covers the activities of all ODNI components from 01 January 2019 through 31 December 2019. Consistent with the Act, constituent elements of the Intelligence Community (IC) will report their activities to Congress through their own departments or agencies.

B. Reporting Requirement

The Act 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 or more electronic databases, where —

1) 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;

2) 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

3) 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]"

C. Report Content

In recent years, the ODNI has followed a format believed to enhance 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" The ODNI reports the latter category of activities in the interest of transparency. Part III provides updates on programs included in the prior

[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).

year's report. Part IV of this report provides an overview of the Privacy and Civil Liberties infrastructure within which ODNI conducts its activities.

NEW ACTIVITIES

A classified annex provides summary information on a new effort administered by the National Counterintelligence and Security Center (NCSC) that does not meet the criteria for full reporting in this report. The ODNI did not undertake any other new reportable activities during the reporting period.

UPDATES ON PREVIOUSLY REPORTED ACTIVITIES

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

A. National Counterterrorism Center

The National Counterterrorism Center (NCTC) continues to conduct "threat analyses" as described in the 2013 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.

B. Intelligence Advanced Research Projects Activity

Intelligence Advanced Research Projects Activity (IARPA) continues to invest in high-risk, high-payoff research programs that have the potential to provide the U.S. with an overwhelming intelligence advantage over future adversaries. As a scientific research funding organization, IARPA does not use, nor does it expect to make use of, data mining technology. IARPA programs are experimental by nature, and are designed to produce new capabilities, such as those described in previous reports and summarized below.

i. Cyber-attack Automated Unconventional Sensor Environment Research Program.

The Cyber-attack Automated Unconventional Sensor Environment (CAUSE) Research Program began in 2016 and ended in 2019. The program investigated the use of automated systems to forecast and provide early warnings of cyber-attack events (e.g., denial of service, spearphishing, malware installation, and accessing malicious websites) with enough lead time to effectively protect information systems. Research identified and validated leading signals from large volumes of both traditional internal sources, such as host and technical sensor data,

and unconventional data sources, including help desk ticketing, social media, and the dark web. Research focused on external sources of publicly available data obtained legally through agreements, free access, subscription, or purchase. CAUSE Research Teams analyzed cyber-attacks and related activities against a particular U.S. business sector/industry (e.g., financial services, energy, defense) to develop forecasting models and cyber entity and event extraction methods that identify and detect patterns of activities that precede such cyber events of interest.

Such activities may contain non-personally identifiable information about the "attackers" (i.e., the sources of a cyber-attack), as well as the victim or target of a cyber-attack, which may include an organization or a person. While this information is publicly available, the CAUSE Research Program is not interested in further identifying individuals responsible for the cyber-attack. The CAUSE Research Program does not generate individual cyber attacker's identities and therefore does not constitute data mining as defined in the statute, but has been reported in the interest of transparency. The CAUSE Research Program to date has demonstrated the utility of various types of data features for forecasting cyber events, particularly from the deep and dark web. The CAUSE program concluded in 2019 and will not be reported in 2020.

ii. Deep Intermodal Video Analytics Research Program.

The Deep Intermodal Video Analytics (DIVA) Research Program began in 2017 and is expected to continue into 2021. The program develops methods for automated activity detection based on watching streaming video from multiple cameras and automatically detecting one or more defined activities of potential interest. The program intends to develop tools for forensic analysis, as well as real-time alerting for user-defined threat scenarios, such as an attack on a government facility or potential coordinated attacks planned at public events. The program will produce a common framework and software prototype for activity detection, person/object detection, and recognition across a multi-camera network. The program does not research methods for identifying individuals, such as face recognition. Instead, the program focuses on activity detection, meaning activities of one or more people performing a specified movement or interacting with an object or group of objects.

For this program, the Test and Evaluation team and researchers collected video data for the defined activities under institutional review board approval to ensure compliance with all relevant privacy concerns. Some collections were in a fully-controlled setting using hired actors, where the actors were provided notice that they were being recorded and consented to the collection of their video data. Other collections were performed under conditions that avoid collecting personally identifiable information

by collecting at sufficiently low resolution and/or obfuscating or redacting the faces of non-consenting subjects using technical means before use by the program.

The National Institute of Standards and Technology (NIST) is performing the official evaluations under the program, and the program has demonstrated the ability to recognize and localize activities in video at levels that surpass previous State-of-the-Art. The program continues to research novel approaches to the activity detection problem to reach false-alarm rates low enough for operational use. All results are posted in real time on a public leaderboard hosted by NIST at https://actev.nist.gov/sdl.

As the program involves pattern-based queries of one or more electronic databases and those pattern-based queries are not subject-based and do not use personal identifiers of a specific individual to retrieve information, the DIVA Research Program meets the statutory definition of data mining.

PROTECTION OF PRIVACY AND CIVIL LIBERTIES

The ODNI Office of Civil Liberties, Privacy, and Transparency (CLPT) works closely with the ODNI Office of General Counsel, ODNI components, and IC elements to ensure that appropriate legal, privacy, and civil liberties safeguards are incorporated into policies, processes and procedures that support the intelligence mission. CLPT is led by the Civil Liberties Protection Officer (CLPO), a position established by the Intelligence Reform and Terrorism Prevention Act of 2004. The duties of the CLPO 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."[4]

In addition, the CLPO also serves as the ODNI's Chief Transparency Officer and chairs the Intelligence Transparency Council, and leads and coordinates the IC's efforts to enhance transparency. Before any innovative capabilities or technology may be used in an operational setting, the use of the capability or technology must be examined for compliance with to Executive Order (EO) 12333, the Privacy Act, and other applicable requirements to determine how these tools could be used consistent with the framework for protecting U.S. persons (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,

[4] National Security Act of 1947, as amended by the Intelligence Reform and Terrorism Prevention Act of 2004 (IRTPA), 50 U.S.C. § 403-3d.

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, which protects information about U.S. citizens and permanent resident aliens that a government agency maintains and retrieves by name or unique identifier. The IC also conforms to policies and procedures under Presidential Policy Directive 28, relating to protections for all personal information contained in SIGINT.

The IC's privacy and civil liberties protective infrastructure is bolstered further by guidance and directives issued by the Office of Management and Budget pertaining to the protection of personally identifiable information and the development, procurement and operation of information technology systems that administer personally identifiable information.

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 Intelligence Community, and the Principles of Intelligence Transparency for the Intelligence Community. These two sets of principles inform the IC's approaches to applying appropriate protections for the types of activities described in this report.