OFFICE OF THE DIRECTOR OF NATIONAL INTELLIGENCE

2021–2023 Data Mining Report

1 January 2021 through 31 December 2023

Table of Contents

INTRODUCTION 3
SCOPE 3
REPORTING REQUIREMENT 3
REPORT CONTENT 3
Protection of Privacy and Civil Liberties 4
New Activities 5
Previously Reported Activities 5
National Counterterrorism Center (NCTC) 5
National Counterintelligence and Security Center (NCSC) 7
Intelligence Advanced Research Projects Activity (IARPA) 7
CONCLUDING STATEMENT 8

INTRODUCTION

The Office of the Director of National Intelligence (ODNI) provides this report pursuant to the *Federal Agency Data Mining Reporting Act of 2007*, 42 U.S.C. § 2000ee-3 (the "Act").

SCOPE

This report covers the activities of all ODNI components from 1 January 2021 through 31 December 2023.[1] Consistent with the Act, constituent elements of the Intelligence Community (IC) will report their activities to Congress through their own departments or agencies.

REPORTING REQUIREMENT

The Act requires that "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."[2] The Act defines "data mining" as "a program involving pattern-based queries, searches, or other analyses of 1 or more electronic databases, where—

(A) a department or agency of the Federal Government, or a non-Federal entity acting on behalf of the Federal Government, is conducting the queries, searches, or other analyses to discover or locate a predictive pattern or anomaly indicative of terrorist or criminal activity on the part of any individual or individuals;

(B) the queries, searches, or other analyses are not subject-based and do not use personal identifiers of a specific individual, or inputs associated with a specific individual or group of individuals, to retrieve information from the database or databases; and

(C) the purpose of the queries, searches, or other analyses is not solely —

(i) the detection of fraud, waste, or abuse in a Government agency or program; or

(ii) the security of a Government computer system."[3]

REPORT CONTENT

This report begins with an overview of ODNI's privacy and civil liberties infrastructure and describes how the infrastructure addresses potential civil liberties and privacy concerns in reportable activities. The report then provides descriptions of the activities that meet the definition of "data mining" under the Act. While this section of the report would provide updates on those activities, during this reporting period, ODNI did not undertake any new reportable activities. As in previous years, this report also provides an update on programs that meet some, but not all, of the criteria defining "data mining." ODNI discretionarily reports activities in this category in the interest of transparency.

[1] The annual reports for 2021, 2022, and 2023 are consolidated with this submission.

[2] 42 U.S.C. § 2000ee-3(c)(1).

[3] 42 U.S.C. § 2000ee-3(b)(1).

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 (OGC), 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 the National Security Act of 1947, and include: "ensur[ing] that the protection of civil liberties and privacy is appropriately incorporated in the policies [of ODNI and the IC]; oversee[ing] compliance by the [ODNI] with [legal] requirements . . . relating to civil liberties and privacy; review[ing] and assess[ing] complaints [about] potential abuses of civil liberties and privacy in the administration of [ODNI programs and activities]; and ensur[ing] that the use of 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 ODNI's Chief Transparency Officer, leading and coordinating the IC's efforts to enhance transparency.[5]

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 Executive Order (EO) 12333, the Privacy Act of 1974, and other applicable requirements to determine how these tools could be used consistent with protecting U.S. person (USP) information.

Like all IC elements, ODNI has 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. Consistent with section 2.3 of the EO, these procedures permit ODNI to collect, retain, and disseminate certain types of USP information if done in the course of ODNI's duly authorized intelligence activities and in fulfillment of ODNI's national security responsibilities. Each IC element's Attorney General-approved procedures are interpreted, applied, and overseen by that element's OGC, 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 relating to protections for all personal information contained in signals intelligence (SIGINT) activities under (a) Presidential Policy Directive 28, during the reporting period through October 2022, and then under (b) EO 14086, *Enhancing Safeguards for United States Signals Intelligence Activities*, which was signed on 2 October 2022.

[4] Section 103D(b) of the National Security Act of 1947, as amended by the Intelligence Reform and Terrorism Prevention Act of 2004, 50 U.S.C. § 3029.

[5] Office of the Director of National Intelligence, *Intelligence Community Directive 107: Civil Liberties, Privacy, and Transparency* (2018), https://www.dni.gov/files/documents/ICD/ICD-107.pdf.

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

Finally, the IC has developed and established three sets of principles that have been adopted as foundational to the IC mission: The *Principles of Professional Ethics for the Intelligence Community*, the *Principles of Intelligence Transparency for the Intelligence Community*, and the *Principles of Artificial Intelligence Ethics for the Intelligence Community*.[6] These sets of principles inform the IC's approaches to applying appropriate protections for the types of activities described in this report.

New Activities

ODNI did not undertake any new reportable activities during the reporting period.

Previously Reported Activities

As indicated in prior reports, ODNI does not always engage in activity meeting the statutory definition of "data mining." Nonetheless, in the interests of transparency, ODNI has included descriptions of certain data processing activities in prior reports. During this reporting period, ODNI continued to conduct these activities, which are described below.

National Counterterrorism Center (NCTC)

During the reporting period, NCTC continued to conduct "threat analyses," as described in prior reports:

> When responding to generalized threat reporting [...] NCTC narrows the data to be correlated with NCTC's terrorism information holdings in order to generate appropriately focused results. NCTC does this by deriving limiting parameters, based in part on analytic assumptions derived from experience and knowledge about the characteristics of the group or individuals historically involved in such threats, and about general terrorist tradecraft (e.g., communications, travel, and counterintelligence). NCTC then applies those parameters to the data at hand.[7]

This technique correlates information that NCTC receives with NCTC datasets to narrow the pool of information. However, such correlation does not meet the statutorily defined criteria for data mining. Specifically, as previously noted in prior reports:

> [C]ertain 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

[6] The principles, respectively, are available at: https://www.intelligence.gov/mission/our-values/336-ethics; https://www.intelligence.gov/index.php/mission/our-values/341-transparency; and https://www.intelligence.gov/images/AI/Principles_of_AI_Ethics_for_the_Intelligence_Community.pdf.

[7] Office of the Director of National Intelligence, *Data Mining Report for Calendar Year 2013*, at 4.

> 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.[8]

Nonetheless, this report provides information about this NCTC "threat analyses" in the interest of transparency and consistent with the *Principles of Intelligence Transparency for the Intelligence Community*.

To facilitate its counterterrorism mission, NCTC receives terrorism threat reports that may include details about threat actors, relevant dates and locations, modes and routes of travel, and other terrorism related information. NCTC uses such details to identify specific information about the potential terrorist threats. NCTC correlates that terrorism information with other pertinent datasets to which NCTC has access. NCTC then examines the results to identify, analyze, and provide leads, consistent with applicable laws and procedures, to its partners.

Only specially trained, authorized personnel are permitted to access the information involved in the NCTC threat analyses, and their analytic activities on NCTC systems are monitored, recorded, and audited. If erroneous or outdated data is identified, it must be corrected, updated, or removed from NCTC systems as appropriate, and the data provider must be notified of the error.

1. Narrowing the Data to Be Correlated for the Threat Analyses

When responding to generalized terrorism threat reporting rather than specific threats, NCTC narrows the data to be correlated within NCTC's terrorism information holdings to generate appropriately focused results. NCTC does this by deriving limiting parameters, based in part on analytic assumptions derived from experience and knowledge about the characteristics of the group or individuals historically involved in such threats, and about general terrorist tradecraft (e.g., communications, travel, and counterintelligence), and applying those parameters to the available information.

2. Procedures for Protecting Privacy and Civil Liberties in the NCTC Threat Analyses

To perform its mission, NCTC receives unevaluated information consistent with applicable law and procedures. NCTC then applies analytic parameters to that unevaluated information to identify connections and derive relevant terrorism information. NCTC analysts make analytic determinations based on any resulting "matches" with terrorism information accessible to NCTC. Authorized and trained analysts analyze the results to identify further leads, including information about the identity of individuals with an apparent nexus to international terrorism, to report to counterterrorism partners in response to the threat reporting. NCTC does not otherwise make use of the information that is "narrowed down" through the use of these parameters.

[8] Office of the Director of National Intelligence, *Data Mining Report for Calendar Year 2013*, at n.1.

If this technique is applied to U.S. Government datasets obtained or accessed by NCTC, NCTC must protect the privacy and civil liberties of USPs whose personal information is contained within this data. NCTC assessment of information in these datasets is designed to identify information that constitutes terrorism information, and to minimize the review of USP information that does not constitute terrorism information. Pattern-based assessment is permitted but must comply with the safeguards provided by the ODNI Attorney General Guidelines. NCTC's techniques, as described above, are designed to minimize the review of non-terrorism information by analysts.

NCTC may retain, use, and disseminate USP information that constitutes terrorism information. Disseminations must satisfy the dissemination requirements of the ODNI Attorney General-approved procedures, any requirements established by the agency that originally provided the data to NCTC, as well as the Privacy Act.

Once information has been disseminated by NCTC to its counterterrorism partners, the information is protected by applicable laws and policies, including the Privacy Act, which applies to all federal agencies, and EO 12333, as applied to IC elements. These measures are subject to compliance and oversight measures at NCTC, as implemented by the NCTC Civil Liberties Protection Officer in consultation with the ODNI Civil Liberty Protection Officer, NCTC legal counsel, and NCTC management.

National Counterintelligence and Security Center (NCSC)

During the reporting period, NCSC continued to correlate and review data as previously reported and consistent with NCSC's mission, applicable laws, and procedures. While NCSC's correlation of that data does not meet the statutorily defined criteria for data mining, it has been detailed in a classified annex that is provided to Congress in the interest of transparency to keep applicable overseers fully informed.

Intelligence Advanced Research Projects Activity (IARPA)

During the reporting period, IARPA continued to invest, consistent with its mission and applicable laws, in high-risk, high-payoff research programs with the potential to provide the United States with an overwhelming intelligence advantage over future adversaries. 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. However, one program, the Deep Intermodal Video Analytics (DIVA) Research Program, correlated data through pattern-based queries. Although the DIVA Research Program's pattern-based queries were not subject-based and did not use personal identifiers of a specific individual to retrieve information, the DIVA Research Program met the statutory definition of data mining and, thus, is again being reported here, though it has been concluded.

The DIVA Research Program began in 2017 and concluded in May 2022, during the reporting period. As previously reported, the program developed 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 did not research methods for

identifying individuals, such as face recognition. Instead, the program focused 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, IARPA's Test and Evaluation team and IARPA 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 would not facilitate personally identifiable information by collecting at sufficiently low resolution and/or obfuscating or redacting any faces of non-consenting subjects using technical means before use by the program.

The National Institute of Standards and Technology (NIST) performed the official evaluations under the program, and the program demonstrated the ability to recognize and localize activities in video that surpassed previous levels. Results were posted in real time on a public leaderboard hosted by NIST at https://actev.nist.gov/sdl.

CONCLUDING STATEMENT

The activities reported here are subject to compliance and oversight measures to ensure they remain consistent with applicable law and procedures and protect privacy and civil liberties.