CLOSING THE DATA MINES!

REPAIRING OVERSIGHT, PRESERVING RIGHTS

epic.org
ELECTRONIC PRIVACY INFORMATION CENTER

ABOUT EPIC

The Electronic Privacy Information Center (EPIC) was established in 1994 to protect privacy, freedom of expression, and democratic values in the information age. Our mission is to secure the fundamental right to privacy in the digital age for all people through advocacy, research, and litigation. We are an independent 501(c)(3) non-profit research and advocacy center, meaning that we do not accept funding from the government nor from corporate foundations.

AUTHOR

Abigail Kunkler, Law Fellow

ACKNOWLEDGEMENTS

The author thanks EPIC colleagues Senior Counsel Jeramie D. Scott, Counsel Chris Frascella, and Law Fellow Mayu Tobin-Miyaji for their hard questions and valuable edits and Counsel Kara Williams, Senior Counsel Enid Zhou, and Communications Fellow Aminah Koshul for their expertise in designing and launching this whitepaper.

EPIC extends special thanks to our generous donors, who make it possible for us to fight back against the unchecked expansion of surveillance and exploitation of our data as an independently funded organization.

TABLE OF CONTENTS

I. Executive Summary

In 2003, America was stunned by the reveal of Total Information Awareness, the Department of Defense's covert attempt to bring together every available piece of information on every person held by government and commercial sources. Acting under the unsubtle slogan "Knowledge Is Power," the government intended to mine its "centralized grand database" of personal information in a vain attempt to predict terrorism. Congress, also kept in the dark about the program, swiftly struck it down.

The Executive branch has not lost its fascination with data mining to predict illegal behavior. Just one problem: it does not work. Data issues abound and human biases and judgments are inescapable. What's more, government data mining is a constitutional minefield, rife with privacy disasters and standing invitations for government overstep and abuse. The Federal Agency Data Mining Reporting Act of 2007 was created to force government data mining into the light. The Act requires agencies to create detailed public reports of their data mining activities that seek patterns indicative of criminal or terrorist activity.

But while the Act is an important safeguard, the 18 years since its creation have exposed its weaknesses. The Act's blunt divisions between the types of data mining the law applies to, coupled with its lack of enforcement mechanisms, impair its ability to keep up with the field's development and instill effective and consistent accountability. The resulting Act is one that agencies treat as optional. When they do purport to fulfill their obligations, the reports are bare and unhelpful, at times relying on incorrect or confused definitions of data mining.

We desperately need the oversight promised by the Act. Armed with AI, data mining capabilities have escalated data collection, retention, and analysis at an unbelievable pace. And alarmingly, the ghost of Total Information Awareness has been revived in the federal government's reported plans to construct a massive and centralized repository of personal data, which it intends to mine as part of the Administration's ruthless anti-immigrant and antidemocratic campaign. The Act must be updated if the public is to have actionable protections for its rights and insight into the government's data activities.

Congress should move now to restore the Act to its intended purpose. Amending the Act to include technical background, set expectations for what satisfies reporting requirements, and require agencies to alert the public when it gives a confidential report

to Congress would make it easier for the public to understand and assess agency activities. The Act's definition of data mining also must be broadened in order to tackle developments in AI that endanger privacy and rights. More than anything, an accountable Act needs teeth. Congress should amend the Act to include an enforcement mechanism that ensures agency participation.

II. Data Mining Has Inherent Failure Points—Dubious Data Sources, Unreliable Outputs, and Lack of Cohesive Standards

Data mining is all about patterns and relationships.[1] By combining computer science with statistics, the system analyzes oceans of data to identify and surface patterns and relationships. The goal is to derive meaningful and useful patterns that human decisionmakers use to inform their next move. That goal can be realized in a number of ways. One of the first data mining tools was simple computer matching—the process of comparing two or more datasets to find matching data points. In fact, one of the earliest government forays into data mining was 1977's Project Match program that compared welfare recipient and federal employment records to determine whether any federal employees were fraudulently collecting welfare while employed.[2] Data mining techniques have become more complicated and sophisticated as computing power, data collection, and indefinite data storage have become far more affordable. Researchers have more recently used data mining to sift through the vast number of Google searches (searching "medicine for cough and fever," for example) and Twitter interactions in an attempt to predict flu outbreaks and disease spread.[3]

AI has significantly changed the data mining landscape. Until recently, data mining was primarily used to identify pre-existing patterns in data in order to describe the dataset. Once the pattern was plucked from a sea of data, a human decisionmaker took over. This workflow changed with the introduction of predictive data mining. Predictive data mining uses advanced statistical models based on machine learning techniques like neural networks or deep learning (techniques which are all, in turn, encompassed by the term AI).[4] Rather than simply identifying pre-existing patterns, the data mining system is now the pattern-finder and pattern-interpreter, attempting to predict future events, outcomes, risks, or opportunities based on the patterns found in historical data and route resources

---

[1] Within the law enforcement and intelligence communities, data mining approaches often are included in the definitions of buzzwords such as "evidence-based," "data-driven," "intelligence-led," "bottom-up," and "problem-oriented" policing.
[2] James P. Nehf, Recognizing the Societal Value in Information Privacy, 78 Wash. L. Rev. 1, 43 (2003).
[3] Elizabeth Joh, *Policing by Numbers: Big Data and the Fourth Amendment*, 89 Wash. L. Rev. 35, 41, n. 46 (2014). *See also* Letter from EPIC to Dr. Eric Schmidt, CEO, Google, Inc. (Nov. 2008), https://epic.org/wp-content/uploads/privacy/flutrends/EPIC_ltr_FluTrends_11-08.pdf.
[4] *What Is Predictive Analytics?*, IBM.com (Aug. 8, 2022), https://www.ibm.com/think/topics/predictive-analytics.

accordingly. Increasingly, predictive data mining is used to automate decisions in response to real-time information. Taking the flu outbreak example one step further, a predictive data mining system would not only surface connections between searches indicative of an ill person and disease spread, but may also allocate funds, vaccines, and manpower toward regions identified as highly infected.

The real innovation of AI in data mining is its video and image processing and natural language abilities. As a field, natural language processing teaches computers to utilize linguistic rules to better process and understand written and spoken words. The field attempts to not only correctly capture the content of the text but also endeavors to create systems that interpret and respond to the intentions of the speaker, including context, sarcasm, or sentiment. (It should be noted that these goals are unproven and, perhaps, impossible). Similarly, AI is rapidly improving at detecting specified objects and people in images and video. Like other jumps in computing power, AI operates at a speed unattainable by whole teams of analysts working mechanically to review images, videos, audio, or written text. Armed with these capabilities, AI-based data mining systems are able to process written data more effectively and to use visual data in ways it could not previously.

Despite its capabilities, data mining has a few functional problems, especially when done by government agencies. From the beginning, data mining is impeded by data sources and quality. The results that data mining turns out may be meaningless or based on faulty or useless patterns. The complete lack of cohesive standards for data mining exacerbates each of these problems. Any data mining project needs to address these points of failure before it ever begins, and certainly before it is ever trusted.

Dubious data sources undermine programs from the start.

The sources from which government data mining projects typically derive their data pose significant issues for obtaining usable results while respecting privacy and civil rights. Data mining relies on a broadscale collection and retention of all kinds of information. While agencies are already under the false impression that more data is the answer, their reliance on so-called data-driven solutions, including data mining, further encourages agencies to vacuum up data indiscriminately, sometimes in violation of the E-

Government Act[5] and Paperwork Reduction Act.[6] Under current privacy and data collection regimes, it is left up to the data mining system developers to accurately and responsibly vet, clean, and use data, creating significant roadblocks in assuring that responsible practices are deployed consistently across agencies.

This problem is compounded by federal agencies' use of commercially available information bought from data brokers.[7] Data brokers are commercial entities that collect, aggregate, and sell massive amounts of information on individuals from a wide variety of sources. The data accumulated is highly personal and often sensitive, including records on a person's internet browsing, purchases, geolocation, employment, finances, health, and more. The federal government is one of the data broker industry's biggest customers.[8] Not only does buying commercially available information pose serious problems under the Fourth and First Amendments, but data brokers rarely verify or correct the information they sell.

Wrapped up in the instinct to gather as much information as possible is the assumption that data is neutral or objective. This is plainly false. Data is not created in a vacuum but offers pieces of context that color and add nuance to a person's life. Plucking a piece of information out of that nuance strips it of its necessary context, including any biases that may influence the information. After collection, bias attaches in a few different ways. Bias may attach because the piece of data collected is deprived of explanatory context. But bias also attaches to the piece of data through the judgment of the collector. It happens like this: through its lifespan, data may be collected, sold, purposed, resold,

---

[5] *See, e.g.*, *EPIC v. Presidential Election Commission*, No. 1:17-cv-01320-CKK (D.D.C. July 3, 2017), https://epic.org/documents/epic-v-presidential-election-commission/. *See also* Ryan Singel, *DHS Data Mining System Shut Down After Privacy Slip Ups*, WIRED (Sept. 5, 2007) (DHS's $42 million ADVISE program was shut down after it was found to have tested on personal data without completing a privacy impact assessment required by law).

[6] *See* EPIC et al., *Joint Comments to the Department of State on Notice of Proposed Information Collection: U.S. Passport Renewal Application for Eligible Individuals* (Mar. 20, 2025), https://epic.org/documents/epic-comment-to-department-of-state-on-notice-of-proposed-information-collection-u-s-passport-application-renewal-application-and-limited-passport-replacement-for-eligible-individuals/.

[7] For further reading on law enforcement's use of commercially available information and exploitation of the data broker loophole, see Brennan Center for Justice et al., *Joint Comment Regarding OMB's Request for Information on Executive Branch Agency Handling of Commercially Available Information* (Dec. 17, 2024), https://epic.org/documents/join-comment-regarding-ombs-request-for-information-on-executive-branch-agency-handling-of-commercially-available-information/.

[8] *Id*. at 15. *See also* Chris Baumohl, *ODNI Report on Intelligence Agencies' Data Purchases Underscores Urgency of Reform*, EPIC (July 7, 2023), https://epic.org/odni-report-on-intelligence-agencies-data-purchases-underscores-urgency-of-reform/.

repurposed, and so on. Each of these instances alters the data by adding subjectivity to it. The decision to gather certain pieces of information over others, to gather more or different information from different groups, to add that information to or take it from a databank, to associate it with this data or that outcome, to share it with this company but not that one. In this way, the same data could point in different directions.

The false incentive to amass as much data as possible, coupled with the federal government's reliance on the inherently unreliable data broker industry, leads to the collection of false, inaccurate, duplicated, or outdated data. This includes information on U.S. persons. All of this means that agencies may base their decisions—which program to fund, which system to build out, or which people to surveil, for example—on information that is of dubious quality at best.

Data mining does not always yield useful results, particularly for law enforcement.

Even if we had perfectly curated and cleaned data, data mining is also particularly prone to meaningless results. The point of data mining is to pull patterns and insights from vast amounts of information. But that does not mean that the discovered pattern is meaningful or useful for making accurate predictions. The sheer volume of data fed into the data mining systems has exponentially expanded the number of possible patterns the algorithm could discover, leading to results that some researchers have labeled "fool's gold."[9] As an example, insurance provider Admiral Insurance announced in 2016 that it would determine rates by analyzing the social media posts of first-time car owners to look for personality traits indicative of safe driving.[10] From press statements, it appears that the algorithm performing the analysis was determining which personality traits were supposedly indicated based on patterns it gleaned, such as a post's organization tending to indicate conscientiousness or over-confidence.[11] Of course, a post's characteristics could indicate any number of things: education level, formality level, stylistic trends, or any number of other context-stripped factors that are unrelated to a person's ability to drive

[9] *See* Gary Smith, *Data Mining Fool's Gold*, 35 J. Info. Tech. 182, 182 (May 11, 2020), https://journals.sagepub.com/doi/epub/10.1177/0268396220915600.
[10] Graham Ruddick, *Admiral to Price Car Insurance Based on Facebook Posts*, Guardian (Nov. 1, 2016), https://www.theguardian.com/technology/2016/nov/02/admiral-to-price-car-insurance-based-on-facebook-posts.
[11] Natasha Lomas, *Facebook Slaps Down Admiral's Plan to Use Social Media Posts to Price Car Insurance Premiums*, TechCrunch (Nov. 2, 2016), https://techcrunch.com/2016/11/02/uk-car-insurance-firm-wants-to-scan-social-media-posts-to-price-premiums/.

safely.[12] Admiral's social media analysis also raises a more pernicious form of uselessness: proxy discrimination. Proxy discrimination occurs when seemingly neutral characteristics or factors—word choice, for example—intentionally or unintentionally functions as a stand in for a protected characteristic. This program was quickly shut down by Facebook.

Again, the problem of useless results is compounded by indiscriminate data practices and reliance on commercially available information supplied by data brokers. Just as data sources must be carefully considered and the data itself must be cleaned and verified, the output of a data mining model must be vetted to avoid following meaningless, illogical, or discriminatory recommendations. As machine learning and other forms of AI increasingly are used to power and refine data mining projects, human vetting and scrutiny only become more important.

The usefulness (or uselessness) of government data mining programs is particularly questionable in the context of criminal or terrorist activity. For decades, scholars and advocates from all sides have argued that data mining is ineffective to combat illegal activity simply because the data is not there to mine.[13] Successful data mining in this context requires a high number of known instances of a particular behavior—in the millions at least—before a pattern can emerge.[14] Particularly where the data mining program is aimlessly sifting through data and not focused on any one individual, the likelihood of generating false positives—where the system falsely identifies someone as a criminal or terrorist—is high.[15] In fact, Paul Rosenzweig, while in his role as Deputy Assistant Secretary for Policy at DHS, once said that false positives are "the only certainty"

---

[12] Read more about big data, AI, and proxy discrimination in Anya E.R. Prince & Daniel Schwarcz, *Proxy Discrimination in the Age of Artificial Intelligence and Big Data*, 105 Iowa L. Rev. 1257, 1257 (2020).

[13] For discussions of data mining efficacy, see Christopher Slobogin, *Government Data Mining and the Fourth Amendment*, 75 U. Chi. L. Rev. 317, 323–27 (2008); Fred H. Cate, *Government Data Mining: The Need for a Legal* Framework, 43 Harv. Civ. Rights-Civ. Lib. L. Rev. 435, 468–77 (2008); Jeff Jonas & Jim Harper, *Effective Counterterrorism and the Limited Role of Predictive Data Mining*, Cato Institute 7–8 (2006), https://www.cato.org/sites/cato.org/files/pubs/pdf/pa584.pdf; Bruce Schneier, *Why Data Mining Won't Stop Terror*, WIRED (Mar. 9, 2006), https://www.wired.com/2006/03/why-data-mining-wont-stop-terror-2/.

[14] Jeffrey S. Seifert, *Data Mining and Homeland Security: An Overview*, CRS Report to Congress at 3–4 (Jan. 18, 2007).

[15] *See* Steven R. Morrison, *The System of Domestic Counterterrorism Law Enforcement*, 25 Stanford L. & Pol. Rev. 341, 341 (2018). Even the Supreme Court has acknowledged the harm that flows from data brokers, such as credit reporting agencies, selling products which recklessly labels individuals as threats to American national security based on only computer matching first and last names. *Transunion, L.L.C. v. Ramirez*, 594 U.S. 413, 429–30 (2021).

in data mining.[16] This is still the case today, nearly two decades later. It is hard to justify the privacy costs of data mining in the face of these results.

[16] Cate, *supra* note 13, at 475.

III. Government Data Mining Poses Inherent Risks that the Federal Agency Data Mining Reporting Act Is Intended to Curb

Governments play with fire when they mine their citizens' data. Whatever the aim is, government data mining threatens to install a nasty status quo of constant surveillance while violating a person's right to speech, association, due process, and freedom from unreasonable searches. As data mining increasingly takes advantage of AI advances and the cost of collecting and retaining information continues to dwindle, data mining can enable a form of timeless and massive surveillance that the legal framework is ill-suited to address. The potential to sift and resift through pools of data or construct individual digital dossiers has proven all too tempting for agencies, but the resulting ubiquitous surveillance will have lasting chilling effects that must be grappled with. The Federal Agency Data Mining Reporting Act (the Act)[17] was created in response to these risks.

Government data mining risks violating essential constitutional protections.

Agencies across the federal government have utilized data mining for decades, from the Department of the Interior's Office of National Resources Revenue[18] to the Central Intelligence Agency (CIA).[19] But while data mining can help agencies achieve their goals more efficiently, its use to predict criminal or terrorist activity also poses significant risks to privacy and civil liberties that must be accounted for.

The discriminatory effects of so-called "predictive" or "data-driven" policing are well-known.[20] As has happened with other attempts to automate services, such as loan

---

[17] 42 U.S.C. § 2000ee-3 (2007) (hereinafter "the Act").
[18] *Interior Establishes Office of Natural Resources Revenue*, U.S. Dep't Interior (Oct. 1, 2010), https://www.doi.gov/pressreleases/news/pressreleases/Interior-Establishes-Office-of-Natural-Resources-Revenue.
[19] *See, e.g.*, Office of Privacy and Civil Liberties, *Data Mining Report*, CIA (2022), https://www.cia.gov/static/2ef313ef6e08a3051becf17361b36bb3/CIA_Data_Mining_Report_UnclassReport_2022.pdf.
[20] *See generally* Sarah Brayne, *Predict and Surveil: Data, Discretion, and the Future of Policing* (2021). *See also* Rashida Richardson et al., *Dirty Data, Bad Predictions: How Civil Rights Violations Impact*

and tenant approvals, the discriminatory practices and effects of policing have become digitized and automated rather than neutralized. The historic data on which the systems are built are riddled with prejudiced decisions, events, and practices that render even facially innocuous pieces of information discriminatory proxies.[21] In applying the patterns gleaned from this data, the system perpetuates its discriminatory effects.[22] It is doubtful that there is a solution to this, but it is certain that any solution will not be clean or quick.

The list of problems attendant to government data mining goes on. Data mining to investigate criminal or terrorist activity is fundamentally contrary to the American justice system and creates major risks for violating at least the First, Fourth and Fifth Amendments. Even if an agency is engaged in legitimate data mining activities with a clear and current link to its mission, unfettered data collection and analysis will enable mission creep over time, leading to unacceptable breaches of privacy as agencies impermissibly expand their practices.

Government data mining for evidence of illegal behavior can upend Fourth Amendment protections. That Amendment preserves the right to privacy by prohibiting unreasonable searches and seizures by the government.[23] However, this right is regularly threatened by agencies' unchecked accumulation and use of peoples' data, including personal and sensitive information that agencies acquire by sidestepping the Fourth Amendment's warrant requirement and purchasing from commercial sources.[24] Without articulating any suspicion, describing the information searched or sought, or demonstrating the reasonableness of its logic or methods—all requirements to obtain a warrant under the Fourth Amendment—agencies automatically thumb through the personal and sensitive information of millions of people. However, as many have noted, the Supreme Court's constraints on the Fourth Amendment—including the Court's overly narrow focus on the government's collection of information, rather than its use, as well as its broad exemption of information shared with third parties from the definition of a

---

Police Data, Predictive Policing Systems, and Justice, 94 N.Y.U. L. Rev. 192, 192 (2019). "Predictive policing" has several names, including "data-informed community-focused policing," "precision policing," and "intelligence-led policing." Ángel Díaz, *Data-driven Policing's Threat to Our Constitutional Rights*, Brookings Inst. (Sept. 13, 2021), https://www.brookings.edu/articles/data-driven-policings-threat-to-our-constitutional-rights/.

[21] Joh, *supra* note 3, at 58.

[22] Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 Cal. L. Rev. 671, 677–93 (2016).

[23] U.S. Const. amend. IV.

[24] For further discussion of the use of data brokers and information sharing agreements to circumvent the Fourth Amendment, see Baumohl, *supra* note 8.

search—have weakened the Amendment and most likely pushed data mining outside of its ambit.[25] Intervention is required to ensure that agencies are only acquiring and using information in ways that respect the Fourth Amendment.

Alongside concerns of invasive searches, government data mining raises due process concerns.[26] The Fifth Amendment requires government to follow certain procedures before depriving someone of a constitutionally protected liberty interest.[27] Law enforcement using data mining to predict illegality poses obvious risks under this requirement. Prosecution and incarceration are the ultimate deprivation of life and liberty. Without mandated transparency measures, however, an individual accused of a crime or terrorist activity based on data mining may never learn that data mining was involved nor have the opportunity to examine the technology and challenge its use, data, or pattern identification.[28]

Examining these concerns, it is clear that predictive data mining practices directly contradict the presumption of innocence that is the basis of the American criminal legal system.[29] The justice system requires law enforcement to have at least reasonable suspicion before searching a person. Similarly, arrests must be based on probable cause related to a specific crime. In contrast, predictive practices use automated systems to sift through massive amounts of information—much of it obtained without any suspicion, let alone a warrant—in order to find contextless anomalies that the system flags as suspicious and uses to both retroactively justify the original surveillance and lay the groundwork for any future law enforcement action.[30] The presumption of innocence is illusory when the government bases its decisions off so-called predictive systems.

---

[25] *See, e.g.*, *Data Mining, Dog Sniffs, and the Fourth Amendment*, 128 Harv. L. Rev. 691, 696–701 (2014). *See also* Slobogin, *supra* note 13, 328–36.
[26] Daniel J. Solove, *Data Mining and the Security-Liberty Debate*, 75 U. Chi. L. Rev. 343, 359 (2008).
[27] U.S. Const. amend V.
[28] In this case, prosecutors will have a *Brady* problem on their hands. *Brady v. Maryland*, 373 U.S. 83, 83 (1963) (mandating prosecutors to disclose material that may be exculpatory or material to guilt or punishment under the Due Process Clause).
[29] *See Coffin v. United States*, 156 U.S. 432, 432 (1895); *In re Winship*, 397 U.S. 358, 373 (1970) (Harlan, J., concurring) ("I view the requirement of proof beyond a reasonable doubt in a criminal case as bottomed on a fundamental value determination of our society that it is far worse to convict an innocent man than to let a guilty man go free."). For comprehensive histories of the presumption of innocence, see Hon. J. Harvie Wilkinson III, *The Presumption of Civil Innocence*, 104 Va. L. Rev. 589, 597–611 (2018) and Alexander Volokh, *N Guilty Men*, 146 U. Penn. L. Rev. 173, 173 (1997).
[30] For further discussion of the ways that data mining does away with individualized suspicion, see Joh, *supra* note 3, at 40–41.

With the amount of granular information available for purchase or housed within the federal government, the opportunities to misuse and abuse information grow. As with data mining generally, it is difficult to assure that agencies and their employees[31] act consistently and responsibly. Natural checks on the government's surveillance power are eroded by innovations in data analysis.[32] Data storage and computing power cost next to nothing,[33] allowing an algorithm to swiftly complete a review that many teams of officers could not have done mere decades ago. Implementing AI in data mining kicks up analysis speeds and sophistication by several notches, making troves of accumulated data searchable and understandable at scale. Combined, these advancements allow law enforcement to piece together a detailed dossier of a person's every moment for any reason or no reason at all.[34] This is different in kind from the surveillance and spying of the past.[35] The allure of these abilities will lead (and has led) to mission creep—the tendency of an agency to self-enlarge its mission over time and collect and use information with only a tenuous relationship to the agency's original purpose. As the dragnet of surveillance grows ever larger, it will have (and has had)[36] chilling effects.[37]

Congress understood the risks. On January 10, 2007, Wisconsin Senator Russ Feingold introduced the Federal Agency Data Mining Reporting Act. In his remarks to Congress, Senator Feingold questioned the reliability and efficacy of data mining for combatting terrorism and emphasized that government data mining opened an enormous opportunity for unchecked government overreach and abuse.[38] In short, he feared that

---

[31] In fact, agency personnel misusing their data access for personal surveillance is already a problem. *See, e.g.*, EPIC et al., *Joint Comments to the FCC on Notice of Proposed Rulemaking: In the Matter of Lifeline and Link Up Reform and Modernization; Affordable Connectivity Program; Supporting Survivors of Domestic and Sexual Violence*, app. 2 (Apr. 12, 2023), https://epic.org/documents/in-the-matter-of-supporting-survivors-of-domestic-and-sexual-violence-nprm.

[32] The encryption community offers valuable insight on this argument. *See* Hal Abelson et al., *Bugs in Our Pockets: The Risks of Client-Side Scanning*, 10 J. Cybersec. 1, 16 (2024) ("One way that democratic societies protect their citizens against the ever-present danger of government intrusion is by making search expensive.").

[33] Professor Orin Kerr argued that the plummeting costs of storage have made it ubiquitous and a significant threat to privacy in his piece arguing for a new privacy statute to replace the Electronic Communications Privacy Act. *See* Orin S. Kerr, *The Next Generation Communications Privacy Act*, 162 U. Penn. L. Rev. 373, 376, 391–95 (2014).

[34] Slobogin, *supra* note 13, at 327.

[35] *See, e.g.*, Bruce Schneier, *The Internet Enabled Mass Surveillance. A.I. Will Enable Mass Spying.*, Slate (Dec. 4, 2023), https://slate.com/technology/2023/12/ai-mass-spying-internet-surveillance.html.

[36] Nina Wang et al., *American Dragnet: Data-Driven Deportation in the 21st Century*, Geo. L. Ctr. Priv. & Tech. (2022) (updated with a new forward in 2025), https://americandragnet.org/.

[37] Slobogin, *supra* note 13, at 327.

[38] 153 Cong. Rec. S359–61 (daily ed. Jan. 10, 2007) (statement of Sen. Russell Feingold).

federal agencies would go "fishing for patterns of criminal or terrorist activity" among the "vast quantities of digital data" the government holds.[39] Such overreach had already occurred four years prior, when Congress and the public learned that the Department of Defense was running the massive, expensive, and invasive data mining program, known as Total Information Awareness, without any oversight or accountability. The bill quickly gained bipartisan support, garnering six cosponsors, before its adoption without amendment into the Implementing Recommendations of the 9/11 Commission Act of 2007. That Act was signed into law on August 3, 2007.

Data mining may have rebranded, but it has not slowed down. At the same time, the worries expressed by Congress in 2007 have proven unfortunately prescient. Since President Trump retook office on January 20, 2025, his executive branch has ransacked databases maintained by federal and state agencies,[40] consolidated access to the data of hundreds of millions of individuals,[41] and fed troves of sensitive and personal information into a centralized repository.[42] This information has been used to inflict catastrophic damage to the integrity of federal databases and the lives of individuals included in them.[43] Federal agencies, particularly the Department of Homeland Security (DHS) and its subcomponent, Immigrations and Customs Enforcement (ICE), have abused their access to surveil and arrest based on political dissent[44] and have not been shy about their interest in

---

[39] *Id*.
[40] Class Action Complaint, *League of Women Voters v. DHS*, Case No. 1:25-cv-03501 (Sept. 30, 2025), https://democracyforward.org/wp-content/uploads/2025/09/LWV-V.-DHS-Complaint-as-filed-9-30-25.pdf. See *also Defeating the DOGE: EPIC's Campaign to Repel the Administration's Attack on Our Privacy*, EPIC, https://epic.org/issues/surveillance-oversight/defeating-the-doge-epics-campaign-to-repel-the-administrations-attack-on-our-privacy/.
[41] Emily Badger & Sheera Frenkel, *Trump Wants to Merge Government Data. Here Are 314 Things It Might Know About You,* N.Y. Times (Apr. 9, 2025), https://www.nytimes.com/2025/04/09/us/politics/trump-musk-data-access.html.
[42] Makena Kelly & Vittoria Elliott, *DOGE Is Building a Master Database to Surveil and Track Immigrants*, WIRED (Apr. 18, 2025), https://www.wired.com/story/doge-collecting-immigrant-data-surveil-track/.
[43] For example, over 4 million still-living people were declared dead after the DOGE accessed systems maintained by the Social Security Administration. James Liddell, *How Social Security Claimants Are Being 'Resurrected' After DOGE Falsely Declares Them Dead*, Independent (Apr. 24, 2025), https://www.independent.co.uk/news/world/americas/us-politics/social-security-dead-doge-claims-musk-b2738662.html. These databases are also being used to carry out mass "deportations." Makena Kelly & Vittoria Elliott, *DOGE Is Building a Master Database to Surveil and Track Immigrants*, WIRED (Apr. 18, 2025), https://www.wired.com/story/doge-collecting-immigrant-data-surveil-track/.
[44] *See Designating Antifa as a Domestic Terrorist Organization*, 90 Fed. Reg. 46317, 46317 (Sept. 25, 2025) (enabling broad, opinion-based surveillance and prosecution based on membership in a political

targeting minority groups. To these disastrous ends, the federal government is blowing past safeguards left and right to sweep up as much information as possible, including from government and commercial sources, into a single massive, searchable database that it will use in its violent anti-immigrant and anti-democratic campaign.[45] The combination of data mining, now powered by AI, with that much information results in an alarmingly swift, powerful, and massive surveillance machine.

The Federal Agency Data Mining Reporting Act is not keeping up.

To establish the oversight and accountability necessary to protect privacy and civil liberties, and ensure that the nation's funds are allocated toward programs which respect these rights, the Federal Agency Data Mining Reporting Act creates a reporting requirement for federal agencies that use data mining to identify patterns of criminal or terrorist activity.[46] The Act requires that the agency engaged in activities to use or develop data mining to submit a publicly available report to Congress.[47] The report should include:

- A thorough description of the data mining activity, its goals, and, where appropriate, the target dates for the deployment of the data mining activity.
- 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.
- A thorough description of the data sources that are being or will be used.
- 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.
- 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

---

group which does not exist). For real world examples, see Ella Lee, *Pritzker, US Lawmakers Condemn 'Broadview 6' Indictment Over ICE Protest*, Hill (Oct. 30, 2025), https://thehill.com/homenews/state-watch/5582331-pritzker-johnson-defend-broadview-six/ and Zolan Kanno-Youngs & Hamed Aleaziz, *Inside Trump's Crackdown on Dissent: Obscure Laws, ICE Agents and Fear*, N.Y. Times (Mar. 12, 2025), https://www.nytimes.com/2025/03/12/us/politics/trump-crackdown-dissent.html.

[45] Several Executive Orders are directed toward consolidating information. *See, e.g., Executive Order 14243*, 90 Fed. Reg. 13681, 13681 (Mar. 20, 2025), https://www.whitehouse.gov/presidential-actions/2025/03/stopping-waste-fraud-and-abuse-by-eliminating-information-silos/.

[46] 42 U.S.C. § 2000ee-3 (2007).

[47] *Id.* at § 2000ee-3(b)(1).

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.
- 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.
- 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—
    - protect the privacy and due process rights of individuals, such as redress procedures; and
    - ensure that only accurate and complete information is collected, reviewed, gathered, analyzed, or used, and guard against any harmful consequences of potential inaccuracies.[48]

If an agency's report discusses sensitive law enforcement systems, proprietary business information, trade secrets, or confidential information, the agency instead may send these confidentially to Congress.[49]

The Act has some limitations. Notably, it includes no enforcement mechanism. Agencies face no penalty when they do not report, fail to make the report public, or submit obviously insufficient reports. As the next section discusses, this is a common occurrence. In the same vein, the Act does not require agencies that transmit their reports confidentially to make any statement to the public at all, making it impossible to know the extent of government data mining.

Perhaps most troubling, though, is that agencies are only required to report pattern-based data mining by defining data mining as a program involving "pattern-based queries, searches, or other analyses" looking for "predictive pattern[s] or anomal[ies]."[50] Subject-based data mining, which "uses an initiating individual" to find associations,[51] is explicitly excluded. It makes sense that Congress would have made this distinction; including every subject-based data mining activity would increase the burden of reporting due to the sheer volume of individual investigations and shed light on a large swath of ongoing

[48] *Id.* at § 2000ee-3(c)(2)(A)-(G).
[49] The Act at § 2000ee-3(c)(3).
[50] *Id.* at § 2000ee-3(b)(1), (b)(1)(A).
[51] *See* Nat'l Research Counsel of the Nat'l Academies, *Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment* (2008), https://nap.nationalacademies.org/catalog/12452/protecting-individual-privacy-in-the-struggle-against-terrorists-a-framework.

investigations and techniques. But excluding all subject-based data mining is equally dangerous, if not more so. The distinction allows for invasive searching without particularized suspicion. The combination of AI-powered data mining and shrunken costs associated with data collection supercharges the government's ability to use the "surveillance time machine"[52] and assemble digital dossiers on any given person at any time in their lives. In a time where this is reality and not speculation, some amount of information on subject-based data mining must be made available for public and Congressional scrutiny.

[52] John Villasenor, *Recording Everything: Digital Storage as an Enabler of Authoritarian Governments*, Brookings Inst. Ctr. Tech. Innovation at 1 (Dec. 14, 2011), https://www.brookings.edu/wp-content/uploads/2016/06/1214_digital_storage_villasenor.pdf.

IV. Putting the Act into Practice: Lessons Learned from Agency Reports

EPIC reviewed 29 data mining reports made available by federal intelligence and law enforcement agencies. DHS[53] and the Office of the Director of National Intelligence[54] (ODNI) publish the bulk of publicly available data mining reports, which include overviews of projects led by numerous subcomponent agencies. Our review revealed several ways in which agencies fail to clear the bar set by the Act. Generally, agency reports do not contain enough information to evaluate an agency's compliance with the Act's reporting requirements. Agencies treat the reports as an optional activity rather than a requirement. Similarly, agencies do not need to inform the public when they transmit confidential reports, and there is no enforcement mechanism when agencies fail to produce reports. This leaves the public unable to distinguish delinquent agencies from compliant agencies with confidential reports and without transparency and accountability in either case. Further, important sections are left un- or under-addressed. Finally, agencies vary widely on the robustness of their discussions of a program's efficacy, impact on privacy and civil liberties, and installed safeguards.

Vague language and misapplied definitions undermine the purpose of reporting.

When agencies publish reports at all, they provide insufficient information to determine whether they are accurately understanding and using the data-mining concepts in the Act. Our review suggests that ODNI repeatedly relied on inconsistent and inaccurate definitions of pattern-based data mining when deciding what to report, suggesting the agency may be underreporting. Further, while some variety existed between them, both ODNI and DHS offered vague descriptions of their programs that obscured the uses and breadth of the programs. The brevity and vagueness of the reports along with the questionable implementation of the Act's requirements all raise the question of whether agencies are engaged in self-selected and misleading reporting.

---

[53] DHS began publishing data mining reports in 2006 and continued to publish until 2021. Its reports are available at *DHS Data Mining Reports*, DHS.gov, https://www.dhs.gov/publication/dhs-data-mining-reports (last accessed Oct. 21, 2025).

[54] ODNI began publishing data mining reports in 2008 and continued to publish until 2023, when it published a single report covering its activities 2021-2023. Unlike DHS, ODNI's reports are not available in a single location. However, each is on file with EPIC.

First, DHS and ODNI repeatedly apply inconsistent (and at times inaccurate) definitions of pattern-based and subject-based data mining when making its decisions on what would be included in the agency's final report. This adds considerable confusion and makes the already-brief explanations of how the programs work difficult to understand. ODNI's National Counterterrorism Center's (NCTC) DataSphere program, which is meant to surface "unknown terrorism relationships" and identify "previously undetected terrorist and terrorism information," provides one example.[55] In 2010, ODNI stated that DataSphere is not reportable data mining because "the data set loaded into the tool is comprised entirely of terrorism information" and thus queries and searches of the program "begin with known identifiers…."[56] It is true DataSphere is exempt because it starts with queries based on known identities.[57] However, ODNI's statement of exemption muddies the water of subject-based data mining by stating that the data mining is subject-based because of the information included in the data set rather than the user's query or search.[58] Following ODNI's logic would incorrectly exempt large swaths of data mining just because personal identifiers are included in data sets that are mined.

On some occasions, ODNI has subtly altered the definitions of the Acts to exclude its programs from the Act's oversight. In one example, it stated that the Catalyst program (developed to analyze information held across agencies) was exempt because it does not have "pattern-matching functionality," incorrectly limiting the world of pattern-based data mining to this one technique.[59] ODNI again relied on an incorrect definition of subject-based data mining to exempt its Mercury Research Program (MRP) in 2016 and 2017. That program attempts to detect patterns of changes in communications before events of interest such as terrorist activity,[60] a technique that certainly seems to fit the Act's definitions. ODNI excluded MRP from reportable data mining programs because it "does not generate individuals' identities and therefore does not constitute data mining as

---

[55] ODNI, 2010 Data Mining Report at 7.
[56] *Id*. The report states that DataSphere will be able to detect patterns in data and that "[d]etails regarding such functionality will be provided in a future report, once such functionality is designed and in development." *Id*. at 7. To its credit, ODNI did follow up in 2011 to state that "no such functionality was developed" and the program was discontinued. ODNI, 2011 Data Mining Report at 5. DHS makes a similar blunder in its 2007 description of the ADVISE program. DHS, 2007 Data Mining Report at 2.
[57] ODNI, 2010 Data Mining Report at 6. Additionally, it is arguable that DataSphere was not exempt from the Act, as ODNI admits that it "contemplate[s] potential pattern-based functionality in future stages." *Id*. If ODNI means that there are plans to develop pattern-based data mining as part of the DataSphere Program, it would have to report that program.
[58] *See* The Act at § 2000ee-3(b)(1).
[59] ODNI, 2010 Data Mining Report at 5–6.
[60] ODNI, 2016 Data Mining Report at 3. *See also* ODNI, 2017 Data Mining Report at 4–5.

defined in the statute."[61] This is simply an incorrect understanding of the distinctions between the Act's definitions of subject-based and pattern-based data mining. The Act applies if the program applies pattern-based analyses to discover patterns indicative of criminal activity, not necessarily individuals. Through its reasoning, ODNI has invented an extra criterion to the Act's definition, that the program generate individuals' identities, which it used to claim its program was exempt.

It is unlikely that ODNI's choices are motivated by its interpretation of the Act's definition of pattern-based data mining, which searches for patterns or anomalies "indicative of terrorist or criminal activity on the part of any individual or individuals."[62] In 2017 and 2018, the Deep Intermodal Video Analytics program was exempted under the same logic as the MRP.[63] The next year, that program was the only one formally reported despite few changes to its description.[64]

Notably, ODNI has only twice determined that it was actually required to report active data mining programs.[65] In all other instances, it "voluntarily" reported programs either because of their future potential for data mining or just "in the interest of transparency." While agencies are encouraged to publish program information and strive for ever-increased transparency, its self-determination and inconsistent application of data mining concepts raise questions of what programs they leave unreported. And because agencies are not required by the Act to clarify why an activity is exempted from the report, this uncertainty and potential manipulation go unchecked.

For its part, DHS does not provide enough information on the functionality of any one program to evaluate its grasp on the distinctions between pattern-based and subject-based data mining. In fact, rather than the "thorough" discussions prescribed by the Act, both DHS and ODNI employ incredibly vague and general descriptions of their programs, making it difficult to get much more than a cursory understanding of what a program is and how it is used. This is explored more deeply in the next section.

The numerous examples of misunderstood or misapplied data mining principles, coupled with a level of generality that significantly reduces oversight potential, raises a significant concern: that these agencies engage in "opt-in" reporting, determining for

---

[61] *Id*.
[62] The Act at § 2000ee-3(b)(1)(A).
[63] ODNI, 2017 Data Mining Report at 3; *see also* ODNI, 2018 Data Mining Report at 4.
[64] ODNI, 2019 Data Mining Report at 5–6.
[65] *See id*. *See also* ODNI, 2020 Data Mining Report at 3–4.

themselves which programs will be included or excluded based on preference or faulty information.

Chronic under-discussion of programs, technologies, and data sources leave critical questions unaddressed.

Across reports, the agencies do not consistently and rigorously discuss key sections that the Act requires. The Act lays out several topics that a data mining report must cover, including "thorough" discussions of the program's data sources, project goals, and the data mining technology used, including the "basis for determining" when patterns or anomalies are indicative of criminal or terrorist activity.[66] Not only does insufficient discussion make evaluation of compliance with the Act difficult, it undercuts the ability of the Act to install any accountability. Conducting any kind of robust cost-benefit analysis of a program requires that these sections be detailed and regularly updated. In particular, the intrusiveness of a program on a person's privacy and rights cannot be well understood without a complete discussion of indices of criminal or terrorist activity. However, where agencies only dedicate anywhere from one paragraph to a few pages to its entire discussion of a program, there can never be adequate information to evaluate the program.

First, some general observations on the "thorough" nature of the reports. While neither ODNI nor DHS provide an adequate level of detail, DHS has historically provided far more detail when describing their programs. This is evident even at a glance—where a typical report from ODNI would clock in around ten pages, DHS reports regularly total over 60 pages.[67] Even so, individual reports are internally inconsistent, giving detailed descriptions of privacy guardrails to some programs but skating over the topic in others. The agencies are united, however, in the brevity of each report section. ODNI is by far the worst offender; while some early reports devote a page or more on one program (including each required topic of discussion), many others devote only one or two paragraphs to the descriptions.[68] Most importantly, the agencies often fail to address

[66] The Act at § 2000ee-3(b)(2)(A)-(C).
[67] Of course, 60 pages is still lean given the number of subcomponents covered in its report (including Immigration and Customs Enforcement, Customs and Border Protection, and the Transportation Security Authority).
[68] *Compare, e.g.*, ODNI's 2010 and 2013 Data Mining Reports with its 2020 and 2021-2023 Data Mining Reports.

required topics entirely. In 2010 and 2011, for example, ODNI failed to discuss data sources at all.

When agencies do cover each topic the Act requires, the discussions are generally lacking. Consider discussions of the data sources used by each data mining program. ODNI will often only describe the category of data source—for example, video, help desk ticketing, social media, and so on. A few will name the origin of the data set, with ODNI's 2008, 2011, and 2018 reports identifying the National Institute of Standards and Technology as a data source.[69] DHS, though varied from program to program, provides far more information. In some program discussions, the agency identifies information systems from which the program gathers data or particular pieces of information relied upon, such as arrival and departure times or certain types of personally identifiable information. Two observations emerge from this. First, many disclosures are too incomplete to be helpful in determining a program's intrusiveness. Without more, we cannot decide whether the data source is something the agency should even be able to access (let alone be allowed to store or process). And, in general, there is no consistency in the discussions.

Similarly, the agencies fail to deliver significant discussions of the programs and the data mining technology they use. Indeed, discussions of the technology betray an utter disinterest in participating in the reporting requirement. In ODNI's breathtakingly brief discussions of technology, it often copies descriptions over from year to year with little or no updates made. In 2018, it directly copied its discussion of the CAUSE program down to the typo.[70] DHS similarly copies information from one year to the next. However, it refuses to copy information from other sources to fulfill its obligation. DHS frequently outsources this discussion by pointing the reader to Privacy Threshold and Privacy Impact Analyses or previous data mining reports rather than completing a description tailored to the data mining aspects of the program.

Agencies return similarly vague and unhelpful discussions of their criteria for establishing indices of criminal or terrorist activity. ODNI reports rarely include criteria, and at least seven reports do not discuss this section at all.[71] DHS is similarly spotty in its

[69] ODNI, 2008 Data Mining Report at 3; ODNI, 2011 Data Mining Report at 6; ODNI, 2018 Data Mining Report at 4.
[70] *Compare* ODNI, 2017 Data Mining Report at 3 *with* ODNI, 2018 Data Mining Report at 4. In 2018, ODNI added a single sentence to the end describing the program's abilities to date. *Id*.
[71] *See generally* ODNI's Data Mining Reports from 2010 to 2016, 2018.

inclusion of this discussion.[72] Most discussions of this point are vague to the point of uselessness. In another egregious example, ODNI merely stated in 2013 that "analysts are trained to narrowly tailor their queries to 'identify information that is reasonably believed to constitute terrorism information'…"[73] DHS has moments of increased specificity,[74] but generally employs similarly vague language, focusing on repeating that it has "risk-based rules" and "targeting rules" that are based on "intelligence." Indeed, DHS more than once states flat-out that it uses data mining to determine what the rules should be.[75] The most helpful discussions do not go beyond single examples or references to trends identified from aggregated data.

The agencies clearly view the data mining reports as a box-checking exercise that it can use to self-justify their actions and waive away with subpar and vague descriptions rather than a meaningful way to gauge the risks and benefits of its own programs or enable Congress and the public to do the same. However, discussions of a program's data sources, data mining techniques, and indices of illegal activity are necessary predecessors to understanding the risks a program poses.

Feeble discussions of program efficacy and impact offer little basis for determining whether additional protections are needed.

The Act requires agencies to go beyond technical discussions and consider the program's effectiveness in light of its goals, its potential or actual impacts on privacy and civil liberties, and the guardrails put in place to prevent or lessen those impacts. Like the topics surveyed above, the reports reflect a general lack of care in their repetitive nature, outsourcing, and brief, surface-level discussion.

Both DHS and ODNI decline to discuss program efficacy with any specificity or real analysis. In fact, ODNI often skips over discussing efficacy or likely efficacy altogether.[76] When it does discuss efficacy in a report, it may not do so for each program,

---

[72] For example, DHS did not discuss criteria for establishing indices of illegal activity for its AFI discussion in 2019, 2020 and 2021.
[73] ODNI, 2013 Data Mining Report at 5.
[74] See DHS's discussion of TSA's Silent Partner and Quiet Skies, which indicates that a traveler may be identified "based on travel patterns matching intelligence regarding terrorist travel[,] upon submitting passenger information matching a partially identified terrorist," or upon submitting passenger information matching a known or suspected terrorist. DHS, 2020-21 Data Mining Report at 29.
[75] *See* DHS, 2020-21 Data Mining Report at 21, 26–27.
[76] At least 5 ODNI reports do not discuss efficacy, including 2010, 2011, 2013, and 2016-17.

as happened in 2019. In both cases, discussions of efficacy boil down to the agency saying, "the program works." In 2008, ODNI stated that "researchers … have articulated sound reasons why they believe their technological approaches could ultimately be successful…." Similarly, DHS offered two sentences to discuss its FALCON-DARTTS program, stating that it "is helpful as an investigative tool in numerous Homeland Security criminal investigations."[77] At other times, DHS offers singular anecdotes (which, coincidentally, tend to differ little across the years) that do not provide any real insight into the real success of the program.

Similarly, neither agency discusses impacts on civil liberties or privacy in any detail, offering little actual insight. ODNI rarely discusses impacts at all, routinely including a copy-paste catch-all section at the end and stating that it recognizes that "data mining techniques… could, potentially, impact the privacy or civil liberties of individuals" if used "without careful consideration."[78] Of course, the remaining discussion (a few paragraphs meant to cover each program reported) falls short of a "thorough" description undertaken with "careful consideration." DHS provides more robust discussion and analyzes impacts for each discrete program that it includes in the report. However, as it does with the technology discussions, DHS outsources much of this by linking readers out to its Privacy Impact Assessments (PIAs). Of course, PIAs have their own drawbacks. When implemented properly, agencies use PIAs to analyze their data practices and related programs before they can be implemented, carefully considering the dangers and benefits. This analysis is later disclosed to the public if the agency proceeds with the program. In practice, however, PIAs are more often done post-hoc (if they are done at all)[79] and without the level of detail, context, and transparency necessary.[80] And unlike data mining reports, PIAs do not have to be updated annually, and predictably most are not. When

[77] DHS, 2020-21 Data Mining Report at 46.
[78] *See, e.g.*, ODNI, 2008 Data Mining Report at 4.
[79] In one egregious example, Immigration and Customs Enforcement, a DHS subcomponent, ran an invasive surveillance program for almost 20 years without conducting a PIA. EPIC, *Comments to OMB on Request for Information: Privacy Impact Assessments* (Apr. 1, 2024), https://epic.org/documents/comments-of-epic-to-omb-on-privacy-impact-assessments/.
[80] For more detail on PIAs, see EPIC, *Comments to OMB on Request for Information: Privacy Impact Assessments* (Apr. 1, 2024), https://epic.org/documents/comments-of-epic-to-omb-on-privacy-impact-assessments/.

DHS points to a PIA in its data mining report, it is rejecting its duty to provide regular and robust discussion.[81]

Unsurprisingly, discussion of protective measures is similarly lacking in detail. ODNI's discussion, which it largely cuts and pastes from year to year with minor deviations, reads more like a restatement of the prompt with an appended laundry list of involved parties. Similarly, DHS has stated that it implements programs and makes decisions "in accordance with applicable law, executive orders, and policy," without further detail.[82] At times, the agencies provide helpful discussion of methods of redress for affected individuals. However, this is still limited in its usefulness. It is good to know that "only specially trained, authorized personnel" whose work is "monitored, recorded, and audited" can access the information. What would be better is to know how these audits are happening and how access controls are implemented. The same can be said for NCTC's process for correcting, updating, or removing erroneous or outdated data. The lack of detail precludes the agencies' programs and activities from scrutiny and purported safeguards amount to a pinky promise.

Discussions of program efficacy and impacts create necessary space for both agencies and Congress to consider whether the programs should continue to exist and how they can be implemented responsibly. Reporting habits like the ones discussed above do not offer that space. The Act cannot be the oversight and privacy protective tool it was intended to be without better reporting from the agencies on these points.

A toothless act yields a haphazard compliance culture.

The Act does not contain any consequences for agencies that have an "opt-in" interpretation of the Act. In fact, there are no consequences if an agency fails to report at all. Further, it contains no enforcement mechanisms that Congress or the public can use to enforce the Act's reporting requirement. The Act is predicated on the idea that Congress may review agency activities and alter agency budgets accordingly. However, with no consequences for failing to report, and no threat of enforcement, the compliance culture surrounding the Act is haphazard at best, and willful avoidance at worst. Public records

[81] This is not to say that pointing to a PIA could never be a sufficient reference for an agency's data mining report. However, federal agencies, including DHS, have shown that they cannot or will not create PIAs that are accurate, timely, and detailed.
[82] DHS, 2019 Data Mining Report at n. 58.

requests only sometimes yield data mining reports, leaving Congress and the public in the dark.

Take DHS as an example. Since 2008, the agency had regularly published data mining reports describing programs across DHS and numerous subcomponents. This stopped in 2020, when DHS failed to publish for the first time. It also did not publish in 2021 or 2022. It was not until EPIC filed a Freedom of Information Act (FOIA) request in June 2023 that DHS complied with its obligation and published a joint report covering 2020 and 2021.[83] Dated August 2022, DHS' joint report states that its "2020 Data Mining Report to Congress was not submitted as planned."[84] While DHS stated then that it was actively working on reports covering 2022 and 2023, none have been published. EPIC submitted another FOIA request seeking all unpublished reports earlier this year.

In another case, the Drug Enforcement Administration (DEA) wholly denied EPIC's FOIA request for data mining reports.[85] In its determination letter, the DEA refused to acknowledge the existence of any data mining reports and stated that to even acknowledge whether such reports exist would risk circumvention of the law. Given the vagueness with which other agencies describe programs in their reports, which are still made publicly available, this is just implausible. Regardless, agencies are required to release portions of responsive documents that do not fall under any exemption to the FOIA's disclosure requirement.[86] And while it is possible the agency only conducts data mining activities that would be confidential, there is no way for the public to know whether these agencies transmit a confidential data mining report to Congress. In light of reports made public from ODNI and DHS—both handling more sensitive programs—it is quite difficult to imagine that there are no portions of the data mining reports that could be released without any threat to the DEA's mission.[87] (Indeed, the Department of Defense has done so.).[88]

---

[83] *See DHS Releases Data Mining Report After EPIC FOIA Request*, EPIC.org (Jan. 2, 2024), https://epic.org/dhs-releases-data-mining-report-after-epic-foia-request/.
[84] DHS, 2020-2021 Data Mining Report at 7–8.
[85] *See* Appendix 1.
[86] 5 U.S.C. § 552(b).
[87] EPIC successfully appealed the DEA's decision. At the time of publication, the remanded FOIA request is still pending.
[88] *See, e.g.*, FOIA Response Obtained by Federation of American Scientists from Department of Defense, (Oct. 14, 2015), https://irp.fas.org/agency/dod/datamine2014.pdf (containing the DOD's final response letter and portions of its 2014 data mining report).

The DEA is not the only agency to dodge EPIC's requests for data mining reports. For this whitepaper, EPIC sent similar requests to the FBI and Department of Justice, both highly likely to be engaged in data mining. Like the DEA, neither agency had ever published a report. These requests have been either wholly denied or effectively ignored and tied up in extensive administrative delays. In its eventual response, which was affirmed on appeal, the FBI claimed no responsive records were found.[89] But we know that the FBI has mined data for decades,[90] including tax data,[91] social media, and location data.[92] Its programs certainly *look* like data mining—its partnership with Palantir, for example, is likely to visualize relationships between the information.[93] But since the FBI never created data mining reports, we cannot know whether the agency thinks its activities are exempt or whether it is just flat-out neglecting its responsibility to report, even to Congress. The reality is that an untold number of data mining activities are unreported. Agencies obscure their activities with meaningless and ambiguous statements and treat the Act as an opt-in activity instead of an obligation that is necessary to provide the oversight democracy needs. While FOIA is an important tool for holding federal agencies accountable, it cannot be the only means of enforcing the Act.

Agencies do not take the data mining reports seriously. Agency treatment of the reporting requirement is reminiscent of other oversight mechanisms, including the Privacy Impact Assessments agencies must complete prior to creating or altering information systems. Rather than taking the opportunity to truly evaluate the need for or efficacy of a program and weigh its costs against its benefits, agencies treat them as a post-hoc box-check. Treated this way, the reports cannot provide oversight and reflection or stop harms before they occur. In short, they cannot serve their intended purposes.

---

[89] *See* Appendix 2.
[90] *See, e.g.*, GAO, *Data Mining: Agencies Have Taken Key Steps to Protect Privacy in Selected Efforts, but Significant Compliance Issues Remain*, Report to the U.S. Senate Subcomm. on Oversight of Gov't Mgmt., Comm. on Homeland Sec. & Gov't Affairs (Aug. 2005), https://www.gao.gov/assets/gao-05-866.pdf.
[91] Dalia Naamani-Goldman, *Anti-terrorism Program Mines IRS' Records*, L.A. Times (Jan. 15, 2007), https://www.latimes.com/archives/la-xpm-2007-jan-15-fi-reveal15-story.html.
[92] Lee Fang, *FBI Expands Ability to Collect Cellphone Location Data, Monitor Social Media, Recent Contracts Show*, Intercept (June 24, 2020), https://theintercept.com/2020/06/24/fbi-surveillance-social-media-cellphone-dataminr-venntel/.
[93] *Id*.

V. Moving Toward True Accountability: EPIC's Recommended Amendments

The Federal Agency Data Mining Reporting Act is an important tool for protecting privacy and civil liberties and preventing agency overreach and shadowy behavior. However, its focus is too narrow and its bite too weak to create meaningful oversight for modern agency activities. As a result, agencies make self-guided decisions, apparently based on ease and interest rather than statutory compulsion. The intentions behind the Act—installing oversight, responsibly allocating Congressional funds, and respecting the fundamental rights of individuals—are undermined by this regime. To restore the Act, EPIC makes the following recommendations.

1. **Amend the Act's definitions of data mining to be more protective and inclusive of modern developments in data analysis**. The Act's focus on pattern-based data mining ignores the myriad ways that subject-based data mining may baselessly violate a person's privacy and other rights. This is increasingly true as data mining practices speed up (and wholly invent) law enforcement tactics and make compiling digital dossiers cheaper and easier. The Act's intentions are best served by broadening its protections to include subject-based data mining activities that may intrude on a person's privacy and rights without evidence sufficient to obtain a warrant. Broadening the definition also mitigates the likelihood of agencies interpreting certain data mining activities as exempt to their own convenience.
2. **Amend the Act's definitions of data mining to include examples, technical background, additional defined terms, or required elements of discussion**. As the previous section discussed, agencies have inconsistently applied the definitions of "pattern-based" and "subject-based" data mining when determining whether they are required to report a program, likely resulting in broad under-reporting of activities that have significant impacts on privacy. Further, agencies inconsistently interpret topics of discussion, omitting important details from their scope. Additional technical background, an included glossary of technical terms, or similar amendments would address some oversight challenges by allowing non-technical members of Congress and the public to evaluate agency programs.
3. **Require agencies that transmit confidential annexes to Congress disclose a statement to the public that (1) certifies that such an annex was sent, and (2)**

provides a high-level description of the included data mining programs. Currently, it is impossible to know whether agencies are completing reports when they are required to do so. Instituting a statement requirement would alert the public that the agency is required to report and is attempting to fulfill that requirement and provides some insight into the efficacy of that attempt. This requirement breathes new life into the Act's oversight goals without burdening agencies. Indeed, some agencies, including DHS and the CIA, produce such a statement each year.[94] Further, the vagueness with which DHS and ODNI discuss current data mining activities suggests that the goals of the program would not be lost by reporting the existence of a confidential annex.

4. **Include an enforcement mechanism to curtail agency's 'opt-in' treatment of the reporting requirement**. Agencies that are clearly engaged in data mining under the Act have never published reports. After years of regular reporting, DHS has suddenly halted its reports and only releases them when the public reminds them of their duty through public records requests and the threat of potential FOIA litigation. The Act includes no consequences for an agency's failure to report and no mechanisms for the public to enforce the Act, such as restricting funds for data mining activities when agencies fail to produce adequate reports. Without these, there is no incentive for agencies to comply with the Act, diminishing oversight and limiting the potential for continued improvement of civil liberties protections.

[94] *See, e.g.*, DHS, 2020-2021 Data Mining Report; CIA, 2022 Data Mining Report, https://www.cia.gov/static/2ef313ef6e08a3051becf17361b36bb3/CIA_Data_Mining_Report_UnclassReport_2022.pdf.

VI. Appendices

Appendix 1: DEA FOIA Chain for Data Mining Reports

epic.org

Electronic Privacy Information Center
1718 Connecticut Avenue NW, Suite 200
Washington, DC 20009, USA

+1 202 483 1140
+1 202 483 1248
@EPICPrivacy
https://epic.org

October 4, 2024

VIA EMAIL

Drug Enforcement Administration
Freedom of Information and Privacy Act Unit
Attn: Intake Sub-Unit
8701 Morrissette Drive
Springfield, Virginia 22152
DEA.FOIA@dea.gov

Dear FOIA Officer:

This letter constitutes a request under the Freedom of Information Act ("FOIA"), 5 U.S.C. § 552(a)(3) and is submitted on behalf of the Electronic Privacy Information Center ("EPIC") to the Drug Enforcement Administration ("DEA").

EPIC requests the public release of any and all reports on the DEA's data mining activities generated by DEA in response to the Federal Agency Data Mining Reporting Act of 2007.

Background

The Federal Agency Data Mining Reporting Act of 2007 ("the Act") requires any agency engaged in data mining activities to submit yearly written reports to Congress on such activities, including: the goals, technological details, data sources, efficacy, potential impacts on civil liberties and privacy interests, and guardrails put in place to protect these interests.[1] Such report must also be made available to the public, excepting any annex containing classified information, law enforcement sensitive information, proprietary business information, or trade secrets.[2]

According to the Act, an agency is engaged in data mining if it conducts queries, searches, or other analyses to "discover a predictive pattern or anomaly indicative of terrorist or criminal activity,"[3] so long as the analyses "are not subject-based and do not use personal identifiers" of a specific individual or group of individuals.[4] Further, the agency does not need to run such analyses itself; it must also report if any non-Federal entity is acting on behalf of the Federal Government.[5]

[1] Federal Agency Data Mining Reporting Act, 42 U.S.C. § 2000ee-3(c) (2007).
[2] *Id*. at § 2000ee-3(C)(3).
[3] *Id*. at § 2000ee-3(b)(1)(A).
[4] *Id*. at § 2000ee-3(b)(1)(B).
[5] *Id*. at § 2000ee-3(b)(1)(A).

The DEA is likely engaged in data mining activities. The Administration is commanded to use all means at its disposal to carry out its mission.[6] Data analysis is one such tool.[7] As early as the 1990s, the DEA worked with the CIA and NSA to amass data and identify links between individuals.[8] Data mining is particularly useful for identifying links among large datasets. Given its utility and the DEA's history, current DEA programs likely use data mining. Operation Overdrive, for example, "uses a data-driven, intelligence-led approach" to identify criminal drug networks within the United States.[9] Another example is the Prescription Drug Monitoring Program Analytics System (PDMPAS), which supports the DEA's Diversion Control Division by analyzing data and identifying patterns.[10]

Data mining presents serious implications for Americans' privacy and civil liberties. Further, it is impossible for citizens to know who has access to their data, what information is taken, and how it is used. For these reasons, Congressional and public oversight are key safeguards. Indeed, Congress established the Act to inject transparency into a system that is "prone to produce inaccurate results and [is] ripe for abuse, error, and unintended consequences."[11] The reporting requirement allows Congress to review the costs and benefits of data mining on a program-by-program basis and evaluate whether new rules are needed to protect Americans' privacy.[12]

In light of the Act's public reporting requirement and Congress' intent, the public has a right to transparency concerning the DEA's use of data mining under the Act.

Request for "News Media" Fee Status and Fee Waiver

EPIC is a "representative of the news media" for fee classification purposes.[13] Based on EPIC's status as a "news media" requester, EPIC is entitled to receive the requested record with only duplications fees assessed.[14]

---

[6] Exec. Order No. 12333 §§ 1.1, 1.14, 46 Fed. Reg. 59941 (Dec. 4, 1981).
[7] *See, e.g.*, U.S. Dep't of Justice Office of the Inspector General, *A Review of the Drug Enforcement Administration's Use of Administrative Subpoenas to Collect or Exploit Bulk Data* (Mar. 2019); David J. Mudd, *Privacy Impact Assessment for the Prescription Drug Monitoring Program Analytics System (PDMPAS)*, at (Aug. 29, 2022).
[8] *See* Bridge Initiative Team, *Factsheet: War on Drugs: Surveillance*, Bridge (Jul. 31, 2023), https://bridge.georgetown.edu/research/factsheet-war-on-drugs-surveillance/ (discussing the DEA's use of subject-based data mining in Project CrissCross and its partnership with the NSA's ICREACH to "sift through billions of metadata records….").
[9] *Statement of Anne Milgram: Hearings on Drug Enforcement Administration Oversight Before the Committee on the Judiciary Subcommittee on Crime and Federal Government Surveillance* (Jul. 27, 2023).
[10] David J. Mudd, *Privacy Impact Assessment for the Prescription Drug Monitoring Program Analytics System (PDMPAS)*, at 2 (Aug. 29, 2022).
[11] *See e.g.*, 153 Cong. Rec. S5 (daily ed. Jan. 10, 2007) (statement of Sen. Russell Feingold).
[12] *Id*.
[13] *EPIC v. DOD*, 241 F. Supp. 2d 5, 15 (D.D.C. 2003).
[14] 5 U.S.C. § 552(a)(4)(A)(ii)(II); 28 C.F.R. § 16.10(d)(1).

In addition, any duplication fees should be waived because EPIC's request satisfies the standards in 28 C.F.R. § 16.10(k) for granting a fee waiver.[15] EPIC satisfies § 16.10(k) because disclosure is "in the public interest because it is likely to contribute significantly to public understanding of the operations or activities of the government and is not primarily in the commercial interests of" EPIC, the requester.[16]

First, disclosure "would shed light on the operations or activities of the government" because the request pertains to the operations and procedures of the DEA.[17]

Second, disclosure would be "likely to contribute significantly to public understanding" of how the DEA uses US citizen data. Pursuant to DOJ's FOIA regulations, this factor is satisfied where disclosure is "meaningfully informative" about the government operations or activities in question, and where disclosure "contribute[s] to the understanding of a reasonably broad audience of persons interested in the subject, as opposed to the individual understanding of the requester."[18]

In addition to the DEA's obligation under the Act, disclosure of data mining reports would be "meaningfully informative" to the public because individuals have no other way of knowing how their personal data may be collected. These reports apprise the public on how the DEA uses their data as well as how it safeguards their data and their rights. These reports also keep Congress meaningfully informed so that it can make legislative and budgetary decisions.

Second, EPIC's request "contribute[s] to the understanding of a reasonably broad audience" because it is a news media representative.[19]

Third, disclosure of the requested information is "not primarily in the commercial interest" of EPIC.[20] EPIC is a non-profit organization committed to privacy, open government, and civil liberties.[21] As demonstrated above, Further, the DOJ "components will presume that when a news media requester has satisfied the requirements of paragraphs (k)(2)(i) and (ii) of this section, the request is not primarily in the commercial interest of the requester."[22] As a non-profit research organization, EPIC has no commercial interest in the requested information. Therefore, as demonstrated above, EPIC is a news media requester and satisfies the public interest standard under (k)(2)(i) and (ii).

For these reasons, a fee waiver should be granted.

---

[15] 28 C.F.R. § 16.10(k); 5 U.S.C. 552(a)(4)(A)(iii).
[16] 5 U.S.C. § 552(a)(4)(A)(iii); 28 C.F.R. § 16.10(k)(1).
[17] 28 C.F.R. § 16.10(k)(2)(i).
[18] 28 C.F.R. § 16.10(k)(2)(ii).
[19] 28 C.F.R. § 16.10(k)(2)(ii)(B) (requiring that components "presume that a representative of the news media will satisfy this consideration.").
[20] 28 C.F.R. § 16(10)(k)(2)(iii).
[21] *See* EPIC, *About EPIC*, https://epic.org/epic/about.html.
[22] 28 C.F.R. § 16(10)(k)(2)(iii)(B).

Conclusion

Thank you for your consideration of this request. EPIC anticipates your determination on its request within 20 calendar days. Please send any responsive documents in searchable PDF form via email to FOIA@epic.org and CC jscott@epic.org. For questions regarding this request contact Jeramie Scott at 202-483-1140 x108 or FOIA@epic.org, cc: jscott@epic.org.

Respectfully submitted,

/s Jeramie Scott
Jeramie Scott
Senior Counsel
Director, Project on Surveillance Oversight

/s Abigail Kunkler
Abigail Kunkler
EPIC Law Fellow

U.S. Department of Justice
Drug Enforcement Administration
FOIA and Privacy Act Unit
8701 Morrissette Drive
Springfield, VA 22152

May 29, 2025

Case Number: 25-00132-F

Abigail Kunkler
Sent via e-mail: kunkler@epic.org

Dear Abigail Kunkler:

This letter responds to your enclosed Freedom of Information Act/Privacy Act (FOIA/PA) request dated October 4, 2024, addressed to the Drug Enforcement Administration (DEA), FOIA/PA Unit.

After a thorough review of your request, we neither confirm nor deny the existence of such records pursuant to Exemption 7(E) of the FOIA. *See* 5 U.S.C. § 552(b)(7)(E). Even to acknowledge the existence of such records in and of itself would disclose techniques, procedures, and/or guidelines that could reasonably be expected to risk circumvention of the law. This is our standard response to such requests and should not be taken to mean that records do, or do not, exist. Please be advised that for the exemption cited, it is reasonably foreseeable that disclosure of the information withheld would harm the interests protected by this exemption.

Finally, FOIA exemption (b)(7)(E) protects "records or information complied for law enforcement purposes when disclosure would disclose techniques and procedures for law enforcement investigations or prosecutions, or would disclose guidelines for law enforcement investigations or prosecutions if such disclosure could reasonably be expected to risk circumvention of the law." How the DEA applies its investigative resources against a particular allegation, report of criminal activity, or perceived threat is, itself a law enforcement technique or procedure that the DEA protects pursuant to exemption (b)(7)(E) of 5 U.S.C. § 552. Accordingly, a confirmation by the DEA that it has or does not have responsive records would be tantamount to acknowledging where the DEA is or is not applying investigative resources thus disclosing the scope of law enforcement techniques and procedures. This is our standard response to such requests and should not be taken to mean that records do, or do not, exist. Please be advised that for each of the exemptions cited, it is reasonably foreseeable that disclosure of the information withheld would harm the interests protected by these exemptions.

Because this office is not assessing any fees in connection with the processing of your request, there is no need for us to consider your request for a waiver of fees.

For your information, Congress excluded three discrete categories of law enforcement and national security records from the requirements of the FOIA. *See* 5 U.S.C. § 552(c). This response is limited to those records that are subject to the requirements of the FOIA. This is a standard notification that is given to all our requesters and should not be taken as an indication that excluded records do, or do not, exist.

You may contact our FOIA Public Liaison at (571) 776-2300 for any further assistance and to discuss any aspect of your request. Additionally, you may contact the Office of Government Information Services (OGIS) at the National Archives and Records Administration to inquire about the FOIA mediation services they offer. The contact information for OGIS is as follows: Office of Government Information Services, National Archives and Records Administration, Room 2510, 8601 Adelphi Road, College Park, Maryland 20740-6001; e-mail at ogis@nara.gov; telephone at (202) 741-5770; toll free at 1-877-684-6448; or facsimile at (202) 741-5769.

If you are not satisfied with DEA's determination in response to this request, you may administratively appeal by writing to the Director, Office of Information Policy (OIP), United States Department of Justice, 441 G Street, NW, 6th Floor, Washington, D.C. 20530, or you may submit an appeal through OIP's FOIA STAR portal by creating an account following the instructions on OIP's website: https://www.justice.gov/oip/submit-and-track-request-or-appeal. Your appeal must be postmarked or electronically transmitted within 90 days of the date of my response to your request. If you submit your appeal by mail, both the letter and the envelope should be clearly marked "Freedom of Information Act Appeal." If possible, please provide a copy of your original request and this response letter with your appeal.

If you have any questions regarding this letter, you may contact our Requester Service Center at (571) 776-2300.

Sincerely,

ANGELA DAVIS

Digitally signed by ANGELA DAVIS

Angela C. Davis, Unit Chief
FOIA and Information Law Section
Freedom of Information and Privacy Act Unit
Office of Chief Counsel
Drug Enforcement Administration

Enclosure

FOIA EXEMPTIONS
SUBSECTIONS OF TITLE 5, UNITED STATES CODE, SECTION 552

(b)(1): Information that is classified to protect national security.
(b)(2): Information related solely to the internal personnel rules and practices of an agency.
(b)(3): Information that is prohibited from disclosure by another federal law.
(b)(4): Trade secrets or commercial or financial information that is confidential or privileged.
(b)(5): Privileged communications within or between agencies, including those protected by the: (1) Deliberative Process Privilege (provided the records were created less than 25 years before the date on which they were requested); (2) Attorney-Work Product Privilege; or (3) Attorney-Client Privilege.
(b)(6): Information that, if disclosed, would invade another individual's personal privacy.
(b)(7): Information compiled for law enforcement purposes that: 7(A) Could reasonably be expected to interfere with enforcement proceedings; 7(B) Would deprive a person of a right to a fair trial or an impartial adjudication; 7(C) Could reasonably be expected to constitute an unwarranted invasion of personal privacy; 7(D) Could reasonably be expected to disclose the identity of a confidential source; 7(E) Would disclose techniques and procedures for law enforcement investigations or prosecutions, or would disclose guidelines for law enforcement investigations or prosecutions if such disclosure could reasonably be expected to risk circumvention of the law; or 7(F) Could reasonably be expected to endanger the life or physical safety of any individual.
(b)(8): Information that concerns the supervision of financial institutions.
(b)(9): Geological information on wells.

PRIVACY ACT EXEMPTIONS
SUBSECTIONS OF TITLE 5, UNITED STATES CODE, SECTION 552a

(d)(5): Information compiled in reasonable anticipation of a civil action proceeding.
(j)(2): Material reporting investigative efforts pertaining to the enforcement of criminal law including efforts to prevent, control, or reduce crime or to apprehend criminals.
(k)(1): Information that is currently and properly classified pursuant to an executive order in the interest of the national defense or foreign policy, for example, information involving intelligence sources or methods.
(k)(2): Investigatory material compiled for law enforcement purposes, other than criminal, which did not result in loss of a right, benefit or privilege under Federal programs, or which would identify a source who furnished information pursuant to a promise that his/her identity would be held in confidence.
(k)(3): Material maintained in connection with providing protective services to the President of the United States or any other individual pursuant to the authority of Title 18, United States Code, Section 3056.
(k)(4): Required by statute to be maintained and used solely as statistical records.
(k)(5): Investigatory material compiled solely for the purpose of determining suitability, eligibility, or qualifications for federal civilian employment or for access to classified information, the disclosure of which would reveal the identity of the person who furnished information pursuant to a promise that his/her identity would be held in confidence.
(k)(6): Testing or examination material used to determine individual qualifications for appointment or promotion in federal government service, the release of which would compromise the testing or examination process.
(k)(7): Material used to determine potential for promotion in the armed services, the disclosure of which would reveal the identity of the person who furnished the material pursuant to a promise that his/her identity would be held in confidence.

epic.org

Electronic Privacy Information Center
1718 Connecticut Avenue NW, Suite 200
Washington, DC 20009, USA

+1 202 483 1140
+1 202 483 1248
@EPICPrivacy
https://epic.org

August 26, 2025

VIA FOIASTAR

Office of Information Policy (OIP)
United States Department of Justice
441 G St., NW
6th Fl.
Washington, DC 20530

RE: Freedom of Information Act Appeal, DEA Request No. 25-00132-F

Dear Director:

This letter constitutes an appeal of the U.S. Drug Enforcement Administration's ("DEA") withholding of records under the Freedom of Information Act ("FOIA"), 5 U.S.C. 552(a)(6)(A). The FOIA request was submitted by the Electronic Privacy Information Center (EPIC) on October 4, 2024 ("EPIC's FOIA Request").

EPIC's FOIA Request sought any data mining reports created by the United States Drug Enforcement Administration (DEA) pursuant to the Federal Agency Data Mining Report Act of 2007, 42 U.S.C. § 2000ee-3. See Appendix A.

The DEA issued a final response to EPIC on May 29, 2025. *See* Appendix B. The agency's final response denied EPIC's request pursuant to Exemption 7(E) to the FOIA. *See* 5 U.S.C. § 552(b)(7)(E). In its muddled final response, the agency argues that it could neither confirm nor deny the existence of the data mining reports EPIC sought because "[e]ven to acknowledge the existence" of the reports "would be tantamount to" disclosing the scope of law enforcement techniques or procedures and could be "reasonably be expected to risk circumvention of the law." Appendix B at 1.

EPIC has reviewed the DEA's final response and has determined that the agency has not established that the *Glomar* response under Exemption 7(E) applies to reports that are not compiled for law enforcement purposes and that do not disclose specific procedures or techniques used in law enforcement investigations. The agency provided bare and conclusory reasoning on why disclosure of the requested records could reasonably be expected to risk

circumvention of the law and applied the *Glomar* response to categorically exempt all material under Exemption 7(E).

This appeal challenges the DEA's final determination regarding the application of Exemption 7(E) as well as the agency's decision to close EPIC's FOIA Request. EPIC also challenges the DEA's failure to release reasonably segregable material. The determination should be withdrawn, and responsive records should be disclosed to EPIC.

Background on the Federal Agency Data Mining Reporting Act of 2007

Data mining combines computer science with statistics to identify patterns in and extract them from massive data sets. Mining techniques vary widely, from setting rules of association to predictive analytics built from machine learning. Various terms are attached to data mining, including "knowledge mining from data," "knowledge extraction," "data analysis," and "data dredging." Within the intelligence and law enforcement communities, data mining approaches are often included in the definitions of a myriad of buzzwords, including "evidence-based,"[1] "data-driven,"[2] "intelligence-led,"[3] "bottom-up,"[4] and "problem-oriented"[5] policing.

The Federal Agency Data Mining Reporting Act of 2007 ("the Act") requires any agency engaged in data mining activities to submit yearly written reports to Congress on such activities, including: a description of the technology and its goals, data sources, efficacy, potential impacts on civil liberties and privacy interests, and guardrails put in place to protect those interests.[6] Such report must also be made available to the public, excepting any annex containing classified information, law enforcement sensitive information, proprietary business information, or trade secrets.[7]

An agency is engaged in data mining if it conducts queries, searches, or other analyses to "discover a predictive pattern or anomaly indicative of terrorist or criminal activity."[8] Importantly, an agency's activities are not considered reportable if they are "subject-based" or use personal identifiers as the starting place for the search or query.[9] This means that any search based on an individual does not qualify, including if it uses their name, birthday, address, and so

[1] Sarah Brayne, *The Criminal Law & Law Enforcement Implications of Big Data*, 14 Annu. Rev. Law Soc. Sci. 293, 293 (Oct. 2018), https://pmc.ncbi.nlm.nih.gov/articles/PMC10817721/.
[2] Muhammad Afzal & Panos Panagiotopoulos, *Data in Policing: An Integrative Review*, 48 Int'l J. Pub. Admin. 411, 417-421 (2025).
[3] Jerry H. Ratcliffe, *Integrated Intelligence and Crime Analysis: Enhanced Information Management for Law Enforcement Leaders*, 1, 2 (2d Ed. Aug. 2007).
[4] Colleen McCue & Andre Parker, *Connecting the Dots: Data Mining and Predictive Analytics in Law Enforcement and Intelligence Analysis*, 70 Police Chief 115, 115 (Oct. 2003).
[5] Muhammad Afzal & Panos Panagiotopoulos, *Data in Policing: An Integrative Review*, 48 Int'l J. Pub. Admin. 411, 417-421 (2025).
[6] Federal Agency Data Mining Reporting Act, 42 U.S.C. § 2000ee-3(c) (2007).
[7] *Id*. at § 2000ee-3(c)(3).
[8] *Id*. at § 2000ee-3(b)(1)(A).
[9] *Id.* at § 2000ee-3(b)(1)(B).

on to begin the search. Further, the agency does not need to conduct the data mining activity itself. Rather, it must report data mining activities conducted by any non-federal entity on the agency's behalf.

Responsive data mining reports should exist. The DEA is commanded to use all means at its disposal to carry out its mission.[10] Data analysis, including data mining, is one such tool. As early as 2008, in fact, the Government Accountability Office reported that the DEA conducted two data mining efforts: one for detecting criminal activities or patterns, the Statistical Management Analysis and Reporting Tool System, and another for the TOLLS database.[11] These would be reportable programs under the Act. The agency very likely engaged in pattern-based data mining activities since that GAO report was issued, including through its Drug Flow Attack Strategy (which employs "predictive intelligence"),[12] the "data-driven, intelligence-led" Operation Overdrive,[13] and the Prescription Drug Monitoring Program Analytics System.[14]

Data mining presents serious implications for Americans' privacy and civil liberties. It is impossible for citizens to know what information is held by the federal government, who has access to it, or how it is used. Congressional and public oversight are critical for injecting transparency into utterly opaque systems that are "prone to produce inaccurate results" and "ripe for abuse, error, and unintended consequences."[15]

Procedural Background

On October 4, 2024, EPIC submitted EPIC's FOIA Request to the DEA via email at DEA.FOIA@dea.gov. EPIC requested the public release of any and all reports on the DEA's data mining activities generated by the DEA pursuant to the Federal Agency Data Mining Reporting Act of 2007. The DEA did not acknowledge EPIC's FOIA Request.

On November 1, 2024, EPIC emailed the DEA's FOIA Office to request an update on EPIC's FOIA Request. This email was not answered. EPIC again requested a status update on November 12, 2024. The DEA acknowledged EPIC's request the following day, November 13, 2024. The agency advised EPIC that its Request, assigned the case number 25-00132-F, was under review.

Between November 13, 2024, and March 29, 2025, EPIC followed up with the DEA's FOIA office four additional times on December 12, 2024, January 23, 2025, and February 4,

[10] Exec. Order No. 12333 §§ 1.1, 1.14, 46 Fed. Reg. 59941 (Dec. 4, 1981).
[11] United States General Accounting Office, *Data Mining: Federal Efforts Cover a Wide Range of Uses*, GAO-04-548 at 47 (May 2004), https://www.gao.gov/assets/gao-04-548.pdf.
[12] *See* Bureau of International Narcotics & Law Enforcement Affairs, 2015 International Narcotics Control Strategy Report (2014), https://2009-2017.state.gov/j/inl/rls/nrcrpt/2014/vol1/223182.htm.
[13] Statement of Anne Milgram: Hearings on Drug Enforcement Administration Oversight Before the Committee on the Judiciary Subcommittee on Crime and Federal Government Surveillance (Jul. 27, 2023).
[14] David J. Mudd, *Privacy Impact Assessment for the Prescription Drug Monitoring Program Analytics System*, at 2 (Aug. 29, 2022).
[15] 153 Cong. Rec. S5 (daily ed. Jan. 10, 2007) (statement of Sen. Russell Feingold).

2025, and March 29, 2025. Three of the four follow-ups went unanswered. The DEA issued its final response on May 29, 2025.

EPIC Appeals the DEA's *Glomar* Response

EPIC appeals the DEA's *Glomar* response. Generally, agencies must acknowledge the existence of responsive information and provide specific, nonconclusory justifications for their withholdings. *ACLU v. CIA*, 710 F.3d 422, 426 (D.C. Cir. 2013). Agencies "may refuse to confirm or deny the existence of records where to answer the FOIA inquiry would cause harm cognizable under an FOIA exception." *Gardels v. CIA*, 689 F.2d 1100, 1103 (D.C. Cir. 1982). While such *Glomar* responses are an exception to the FOIA's presumption of disclosure, they are only available in limited circumstances. *Id*. The agency cannot use a *Glomar* response regarding information that it has acknowledged and that it is statutorily obligated to disseminate to the public. Thus, EPIC's appeal should be granted.

The DEA may not claim a *Glomar* exception over material that it is obligated to disseminate to the public and which the public knows should exist. The DEA, through its statements and publications, has acknowledged its use of data mining. The Act mandates that the DEA release parts of its data mining reports that do not qualify for the confidential annex transmitted to Congress.[16] The agency may not claim that it cannot confirm nor deny the existence of a record that it is statutorily required to create for and release to the public. While the Act recognizes that some material should not be made public,[17] the remaining portions of a report must be distributed. As detailed throughout this appeal, the DEA's data mining activities are a commonly known technique. It is required to report these activities to the public. The agency undermines the Act and rejects all accountability when it refuses to release the remaining portions of the report.

Clearly, the DEA is itself or through third parties utilizing data mining. Both are reportable. Because the DEA readily acknowledges these instances of data mining, and because it is compelled to release portions of the Act, it cannot claim any *Glomar* exceptions.

EPIC Appeals the DEA's *Glomar* Response Under Exemption 7(E)

EPIC also appeals the DEA's grounding of its *Glomar* response in Exemption 7(E). The agency bears the burden of demonstrating that the fact of the existence or non-existence of agency records fits a FOIA exemption. *Wolf v. CIA*, 473 F.3d 370, 374 (D.C. Cir. 2007). The review standards for non-*Glomar* cases are applied to determine whether a *Glomar* response was appropriate. *Gardels*, 689 F.2d at 1103-05. The DEA has failed to carry its burden to establish that the exemption applies and has failed to release any reasonably segregable records which would not be subject to the exemption. Thus, EPIC's appeal should be granted.

[16] 42 U.S.C. § 2000ee-(b).
[17] *Id.*

Exemption 7(E) to the FOIA permits the agency to withhold records that "would disclose techniques and procedures for law enforcement investigations or prosecutions, or would disclose guidelines for law enforcement investigations or prosecutions if such disclosure could reasonably be expected to risk circumvention of the law." 5 U.S.C. § 552(b)(7)(E). The phrase "techniques and procedures" refers to the means by which agencies conduct investigations. *See Hansten v. DEA*, 2022 WL 2904151 at *3 (D.D.C. Jul. 22, 2022) *(*quoting *Allard K. Lowenstein Int'l Hum. Rts. Project v. DHS*, 626 F.3d 678, 683 (2d Cir. 2010)).

As a threshold consideration, exemption 7(E) does not apply to the requested records because that exemption only protects "records or information compiled for law enforcement purposes." 5 U.S.C. § 552(b)(7)(E). EPIC's FOIA Request seeks any data mining reports created by the DEA for the purposes of oversight and accountability. Because these are not law enforcement purposes, the data mining reports EPIC seeks do not fall within exemption 7(E). *See Ibrahim v. U.S. State Dep't,* 311 F. Supp. 3d 134, 143 (D.D.C. 2018).

Further, the agency's *Glomar* response does not fall under Exemption 7(E) because the agency's use of data mining is common knowledge, and acknowledging the existence of the requested records would not provide any detailed insight into the agency's specific application of any data mining programs.

In its final letter, the agency stated that "[h]ow the DEA applies its investigative resources against a particular allegation, report of criminal activity, or perceived threat is, itself a law enforcement technique or procedure that the DEA protects" under the exemption. Appendix B at 1. The agency further stated that confirming the existence of the requested data mining reports would be "tantamount to acknowledging where the DEA is or is not applying investigative resources thus disclosing the scope of law enforcement techniques or procedures." *Id*.

The agency cannot properly claim that the very existence of data mining reports may not be disclosed when the techniques are well known and the agency itself has acknowledged their use. *See Billington v. DOJ*, 69 F. Supp. 2d 128, 140 (D.D.C. 1999). In several publications, spanning nearly two decades, the agency has acknowledged its use of data mining activities. In a 2010 statement, former Assistant Administrator for Intelligence Anthony Placido acknowledged that the International Drug Flow Attack Strategy utilizes "intelligence-driven enforcement" and "predictive intelligence."[18] As discussed above, both rely on pattern-based data mining capabilities. This is just one example. For another, the EPIC Inquiry System (formerly the EPIC Seizure System) "incorporates subsystems and applications that aid in analyzing trends and identifying patterns of criminal activities"—precisely the kind of reportable data mining activity

[18] Statement of Anthony P. Placido Before the House Oversight and Government Reform Subcommittee on National Security and Foreign Affairs (Mar. 3, 2010), https://www.dea.gov/sites/default/files/pr/speeches-testimony/2012-2009/ct030310.pdf.

contemplated by the Act.[19] Further, the agency has contracted for analytics support from external suppliers that use data mining techniques, including the Brite Group[20] and Deloitte.[21] Where the public is already aware, based on the agency's own communications, that the DEA is using data mining, the very fact that data mining reports exist cannot be hidden. A *Glomar* response was thus unwarranted.

Further, acknowledging the existence or nonexistence of the requested records would reveal nothing about the DEA's techniques or procedures. *See Hansten v. DEA*, No. 21-2043, 2022 WL 2904151, at *1, *4 (D.D.C. Jul. 22, 2022) (holding that DEA improperly categorically denied request under 7E because "such forms say nothing about how the DEA would '*go about* investigating' a diversion case."). Data mining reports do not, as the agency asserts, relate to any particular allegations, reports, or threats. Rather, the reports are meant to convey the agency's goals in using data mining and assess its strategies for limiting any privacy and civil liberties impacts. For the same reasons, acknowledging the existence of data mining reports would not "acknowledge where the DEA is or is not" applying its resources as no specific investigations are implicated. Appendix B at 1. The reports are more akin to a "list of tools" than "specific instructions." *Brennan Ctr. for Just. At N.Y. Univ. Sch. Of L. v. ICE*, 571 F. Supp. 3d 237, 248 (S.D.N.Y. 2021). As such, the existence of the records is not properly within Exemption 7(E). *Compare Id.* (holding that "a list of tools rather than specific instructions for how, when, and why to use such tools" does not qualify as a 'specialized, calculated technique or procedure' that would 'not be apparent to the public'") *with Kendrick*, 2022 WL 4534627, at * 8 (D.D.C. 2022) (withholding records which would detail non-public information such as targets, dates of use, types of devices, or installation information); *Vazquez v. DOJ*, 887 F. Supp. 2d 114, 117-119 (D.D.C. 2012) (withholding records which may specify individuals mentioned in databases); *Shapiro v. DOJ*, 239 F. Supp. 3d 100, 111-16 (D.D.C. 2017) (holding that FBI request search slips may be withheld because they may serve to confirm the existence of an investigation); *Sanders v. FBI*, No. 20-3672, 2022 WL 888191, at *2, *4-5 (D.D.C. Marc. 25, 2022) (upholding a *Glomar* response because confirming or denying whether the FBI coordinated with specific foreign agencies may allow bad actors to circumvent the law).

---

[19] James Robert Bryden, *Privacy Impact Assessment for the EPIC Inquiry System (EIS)*, DEA (Jul. 21, 2025), https://www.dea.gov/sites/default/files/2025-07/DEA%20PIA%20for%20EPIC%20Inquiry%20System.pdf.

[20] *DOJ DEA Data Analytics BPA*, Brite Group (last accessed Aug. 13, 2025), https://www.thebritegroup.com/doj-dea-data-analytics-bpa/. The Brite Group employs data mining as part of its analysis work, according to job postings from the company. *See* Brite Group, *Junior Data Scientist-DOD Secret Clearance Required*, Glassdoor (last accessed Aug. 13, 2025), https://www.glassdoor.com/job-listing/junior-data-scientist-dod-secret-clearance-required-the-brite-group-JV_KO0,51_KE52,67.htm?jl=1009789729691.

[21] *See* BPA Call on Specialized Data Analytics Support Services Awarded to Deloitte & Touche LLP, BPA 5DDHQ22A00000012, https://www.fpds.gov/common/jsp/LaunchWebPage.jsp?command=execute&requestid=316412487&version=1.5.

EPIC Challenges the DEA's Failure to Release Reasonably Segregable Material

EPIC also challenges the scope of the DEA's assertion of Exemption 7(E). As with the agency's assertion of the exemption, the DEA entirely failed to justify withholding potentially responsive records in its entirety and refusing to release any reasonably segregable portions of the requested records.

Even if the DEA has properly invoked Exemption 7(E), which EPIC does not concede, the agency must still release any "reasonably segregable portion" of the records. 5 U.S.C. § 552(b); *Stolt-Nielsen Transp. Group Ltd. V. United States*, 534 F.3d 728, 734 (D.C. Cir. 2008); *Oglesby v. United States Dep't of the Army*, 79 F.3d 1172, 1176 (D.C. Cir. 1996). The burden is on the agency to "provide a detailed justification for its non-segregability." *Johnson v. EOUSA*, 310 F.3d 771, 776 (D.C. Cir. 2002) (internal quotations marks omitted). This includes "a statement of [the government's] reasons" and a "descri[ption of] what proportion of the information in a document is non-exempt and how that material is dispersed throughout the document." *Mead Data Cent., Inc. v. Dep't of Air Force*, 566 F.2d 242, 261 (D.C. Cir. 1977).

The DEA did not explain its decision to withhold potentially responsive records. For example, its final letter did not state, let alone provide evidence, that any non-exempt material in the records is inextricably entwined with exempt material under Exemption 7(E) or any other exemption. To the extent that it is found on appeal that Exemption 7(E) does apply, the DEA must still release any reasonably segregable portion of pages. While the DEA may claim (and actually provide its "detailed justification") that no portion is reasonably segregable, this is very unlikely to be the case. When the same reports were sought from the Department of Defense, for example, the agency was able to segregate and release portions of the reports.[22]

Further, the Act mandates the public release of data mining reports.[23] Like the FOIA, the Act recognizes that some information is too sensitive to release to the public. Thus, the Act provides that agencies may create confidential annexes for Congress that discuss classified information, law enforcement sensitive information, proprietary business information, or trade secrets.[24] Importantly, these annexes are subparts of the report—the DEA is still obligated to release the remaining portions of its data mining reports.[25] The DEA's failure to release portions of its reports that are reasonably segregable from any confidential annexes is not only a violation of the FOIA, but of the Federal Agency Data Mining Reporting Requirement Act which mandates the creation and public release of the reports.

---

[22] *See, e.g.*, FOIA Response Obtained by Federation of American Scientists from Department of Defense, (Oct. 14, 2015), *available at* https://irp.fas.org/agency/dod/datamine2014.pdf (containing the DOD's final response letter and portions of its 2014 data mining report).
[23] 42 U.S.C. § 2000ee-3(c)(1).
[24] 42 U.S.C. § 20003-ee(c)(3)(A).
[25] *See Id*. at § 2000ee-3(c)(1) ("The report shall be made available to the public, *except for* an annex[.]") (emphasis added).

For the reasons set forth above, the DEA is required by the FOIA to release reasonably segregable portions of the requested documents. EPIC certifies this explanation is true and correct to the best of its knowledge and belief. 5 U.S.C. § 552(a)(6)(E)(vi). For the foregoing reasons, EPIC's appeal of the agency's withholding under Exemption 7(E) must be granted.

Conclusion

Thank you for your consideration of this appeal. EPIC anticipates your determination on our appeal within twenty business days. 5 U.S.C. Sec. 552(a)(6)(A)(ii). For questions regarding this appeal, contact Abigail Kunkler at FOIA@epic.org, cc: kunkler@epic.org.

Respectfully submitted,

/s Abigail Kunkler
Abigail Kunkler
EPIC Law Fellow

U.S. Department of Justice
Office of Information Policy
Sixth Floor
441 G Street, NW
Washington, DC 20530-0001

Telephone: (202) 514-3642

Abigail Kunkler

Re: Appeal No. A-2025-02518
Request No. 25-00132-F

foia@epic.org

VIA: Online Portal - 9/29/2025

Dear Abigail Kunkler:

You appealed from the action of the Drug Enforcement Administration (DEA) on your Freedom of Information Act (FOIA) request for access to reports on the DEAs data mining activities generated by DEA in response to the Federal Agency Data Mining Reporting Act of 2007. I note that your appeal concerns DEA's refusal to confirm or deny the existence of records pursuant to Exemption (b)(7)(E).

After carefully considering your appeal, and as a result of discussions between DEA personnel and this Office, I am remanding your request to DEA for a search for responsive records. If DEA locates releasable records, it will send them to you directly, subject to any applicable fees. You may appeal any future adverse determination made by DEA. If you would like to inquire about the status of this remanded request or to receive an estimated date of completion, please contact DEA directly at 571-776-2300.

If you are dissatisfied with my action on your appeal, the FOIA permits you to file a lawsuit in federal district court in accordance with 5 U.S.C. § 552(a)(4)(B).

Sincerely,

X Jillian Warzynski
Jillian Warzynski

Associate Chief, for Christina Troiani, Chief,
Administrative Appeals Staff

Appendix 2: FBI FOIA Chain for Data Mining Reports

epic.org

Electronic Privacy Information Center
1718 Connecticut Avenue NW, Suite 200
Washington, DC 20009, USA

+1 202 483 1140
+1 202 483 1248
@EPICPrivacy
https://epic.org

October 4, 2024

VIA WEB PORTAL

Federal Bureau of Investigation
Attn: Initial Processing Operations Unit
Record/Information Dissemination Section
200 Constitution Drive
Winchester, VA 22602

Dear FOIA Officer:

This letter constitutes a request under the Freedom of Information Act ("FOIA"), 5 U.S.C. § 552(a)(3) and is submitted on behalf of the Electronic Privacy Information Center ("EPIC") to the Federal Bureau of Investigation ("FBI").

EPIC requests the public release of any and all reports on the FBI's data mining activities generated by the FBI in response to the Federal Agency Data Mining Reporting Act of 2007.

Background

The Federal Agency Data Mining Reporting Act of 2007 ("the Act") requires any agency or department engaged in data mining activities to submit yearly written reports to Congress on such activities, including: the goals, technological details, data sources, efficacy, potential impacts on civil liberties and privacy interests, and guardrails put in place to protect these interests.[1] Such report must also be made available to the public, excepting any annex containing classified information, law enforcement sensitive information, proprietary business information, or trade secrets.[2]

According to the Act, an agency is engaged in data mining if it conducts queries, searches, or other analyses to "discover a predictive pattern or anomaly indicative of terrorist or criminal activity,"[3] so long as the analyses "are not subject-based and do not use personal identifiers" of a specific individual or group of individuals.[4] Further, the agency does not need to run such analyses itself; it must also report if any non-Federal entity is acting on behalf of the Federal Government.[5]

[1] Federal Agency Data Mining Reporting Act, 42 U.S.C. § 2000ee-3(c) (2007).
[2] *Id*. at § 2000ee-3(C)(3).
[3] *Id*. at § 2000ee-3(b)(1)(A).
[4] *Id*. at § 2000ee-3(b)(1)(B).
[5] *Id*. at § 2000ee-3(b)(1)(A).

The FBI has been engaged in data mining as defined by the Act. The FBI is commanded to use all means at its disposal to carry out its mission,[6] including data analysis.[7] At least as early as 2007, the FBI was conducting pattern-based data mining[8] as defined by an identical provision of the reauthorized PATRIOT Act.[9] In 2013, the Foreign Terrorist Tracking Task Force used data mining to identify and assess unknown terrorist threats or persons of interest, including U.S. citizens.[10]

It is likely the FBI continues to incorporate data mining into its work. In 2022 alone, the FBI produced around 3,000 domestic terrorism-related intelligence products which incorporated collection, domain, targeting, or threat analysis.[11] Examples of likely data mining include the Violent Criminal Apprehension Program (ViCAP) Web National Crime Database (which can locate patterns between otherwise-unrelated crimes),[12] The Uniform Crime Reporting (UCR)'s National Incident-Based Reporting System,[13] and the National Data Exchange (N-DEx) system that "enables users to 'connect the dots' between data on people, places, and things that may seem unrelated…."[14]

Data mining presents serious implications for Americans' privacy and civil liberties. Further, it is impossible for citizens to know who has access to their data, what information is taken, and how it is used. For these reasons, Congressional and public oversight are key safeguards. Indeed, Congress established the Act to inject transparency into a system that is "prone to produce inaccurate results and [is] ripe for abuse, error, and unintended consequences."[15] The reporting requirement allows Congress to review the costs and benefits of data mining on a program-by-program basis and evaluate whether new rules are needed to protect Americans' privacy.[16]

In light of the Act's public reporting requirement and Congress' intent, the public has a right to transparency concerning the FBI's use of data mining under the Act.

---

[6] Exec. Order No. 12333 §§ 1.1, 1.14, 46 Fed. Reg. 59941 (Dec. 4, 1981).
[7] *See Intelligence*, FBI.gov, https://www.fbi.gov/investigate/how-we-investigate/intelligence (last visited Oct. 1, 2024) ("Gathering intelligence has always been critical to fulfilling the FBI's mission. Some techniques we use to do this include[s] … data analysis.").
[8] *See* U.S. Dep't of Justice, Report on "Data-Mining" Activities Pursuant to Section 126 of the USA PATRIOT Improvement and Reauthorization Act of 2005 (Jul. 9, 2007), available at https://epic.org/wp-content/uploads/privacy/fusion/doj-dataming.pdf.
[9] USA Patriot Improvement and Reauthorization Act of 2005, Pub. L. No. 109-177, § 126, 120 Stat. 192, 227 (2006), available at https://www.congress.gov/109/plaws/publ177/PLAW-109publ177.pdf.
[10] *See* U.S. Dep't of Justice, The Federal Bureau of Investigation's Foreign Terrorist Tracking Task Force Audit Report 13-18 at 5-7 (Mar. 2013), available at https://oig.justice.gov/reports/2013/a1318r.pdf.
[11] Fed. Bureau of Investigation & Dep't of Homeland Sec., *Strategic Intelligence Assessment and Data on Domestic Terrorism* at 27-28 (Jun. 2023).
[12] *Behavioral Analysis*, FBI.gov, https://www.fbi.gov/investigate/how-we-investigate/behavioral-analysis (last visited Oct. 1, 2024).
[13] Uniform Crime Reporting Program, *Benefits of NIBRS Participation* at 3 (2016).
[14] National Data Exchange (N-DEx), FBI.gov, https://www.fbi.gov/investigate/how-we-investigate/behavioral-analysis (last visited Oct. 1, 2024).
[15] *See e.g.*, 153 Cong. Rec. S5 (daily ed. Jan. 10, 2007) (statement of Sen. Russell Feingold).
[16] *Id*.

Request for "News Media" Fee Status and Fee Waiver

EPIC is a "representative of the news media" for fee classification purposes.[17] Based on EPIC's status as a "news media" requester, EPIC is entitled to receive the requested record with only duplications fees assessed.[18]

In addition, any duplication fees should be waived because EPIC's request satisfies the standards in 28 C.F.R. § 16.10(k) for granting a fee waiver.[19] EPIC satisfies § 16.10(k) because disclosure is "in the public interest because it is likely to contribute significantly to public understanding of the operations or activities of the government and is not primarily in the commercial interests of" EPIC, the requester.[20]

First, disclosure "would shed light on the operations or activities of the government" because the request pertains to the operations and procedures of the FBI.[21]

Second, disclosure would be "likely to contribute significantly to public understanding" of how the FBI uses US citizen data. Pursuant to DOJ's FOIA regulations, this factor is satisfied where disclosure is "meaningfully informative" about the government operations or activities in question, and where disclosure "contribute[s] to the understanding of a reasonably broad audience of persons interested in the subject, as opposed to the individual understanding of the requester."[22]

In addition to the FBI's obligation under the Act, disclosure of data mining reports would be "meaningfully informative" to the public because individuals have no other way of knowing how their personal data may be collected. These reports apprise the public on how the FBI uses their data as well as how it safeguards their data and their rights. These reports also keep Congress meaningfully informed so that it can make legislative and budgetary decisions.

Second, EPIC's request "contribute[s] to the understanding of a reasonably broad audience" because it is a news media representative, as discussed above.[23]

Third, disclosure of the requested information is "not primarily in the commercial interest" of EPIC.[24] EPIC is a non-profit organization committed to privacy, open government, and civil liberties.[25] As demonstrated above, Further, the DOJ "components will presume that

[17] *EPIC v. DOD*, 241 F. Supp. 2d 5, 15 (D.D.C. 2003).
[18] 5 U.S.C. § 552(a)(4)(A)(ii)(II); 28 C.F.R. § 16.10(d)(1).
[19] 28 C.F.R. § 16.10(k); 5 U.S.C. 552(a)(4)(A)(iii).
[20] 5 U.S.C. § 552(a)(4)(A)(iii); 28 C.F.R. § 16.10(k)(1).
[21] 28 C.F.R. § 16.10(k)(2)(i).
[22] 28 C.F.R. § 16.10(k)(2)(ii).
[23] 28 C.F.R. § 16.10(k)(2)(ii)(B) (requiring that components "presume that a representative of the news media will satisfy this consideration.").
[24] 28 C.F.R. § 16(10)(k)(2)(iii).
[25] *See* EPIC, *About EPIC*, https://epic.org/epic/about.html.

when a news media requester has satisfied the requirements of paragraphs (k)(2)(i) and (ii) of this section, the request is not primarily in the commercial interest of the requester."[26] As a non-profit research organization, EPIC has no commercial interest in the requested information. Therefore, as demonstrated above, EPIC is a news media requester and satisfies the public interest standard under (k)(2)(i) and (ii).

For these reasons, a fee waiver should be granted.

Conclusion

Thank you for your consideration of this request. EPIC anticipates your determination on its request within 20 calendar days. Please send any responsive documents in searchable PDF form via email to FOIA@epic.org and CC jscott@epic.org. For questions regarding this request contact Jeramie Scott at 202-483-1140 x108 or FOIA@epic.org, cc: jscott@epic.org.

Respectfully submitted,

/s Jeramie Scott
Jeramie Scott
Senior Counsel
Director, Project on Surveillance Oversight

/s Abigail Kunkler
Abigail Kunkler
EPIC Law Fellow

[26] 28 C.F.R. § 16(10)(k)(2)(iii)(B).

U.S. Department of Justice

Federal Bureau of Investigation
Washington, D.C. 20535

November 6, 2024

MR. JERAMIE D. SCOTT
ELECTRONIC PRIVACY INFORMATION CENTER
1519 NEW HAMPSHIRE AVENUE NORTHWEST
WASHINGTON, DC 20036

Request No.: 1649273-000
Subject: FBI Data Mining Activities

Dear Mr. Scott:

This is in response to your Freedom of Information/Privacy Acts (FOIPA) request. Based on the information you provided, we conducted a main entity record search of the Central Records System (CRS) per our standard search policy. However, we were unable to identify records subject to the FOIPA that are responsive to your request. Therefore, your request is being closed. If you have additional information pertaining to the subject of your request, please submit a new request providing the details, and we will conduct an additional search. For more information about records searches and the standard search policy, see the enclosed FBI FOIPA Addendum General Information Section.

Please see the paragraphs below for relevant information that may be specific to your request. Only checked boxes contain corresponding paragraphs relevant to your request. If no boxes are checked, the corresponding information does not apply.

☐ Please be advised that your request was reopened based on the additional information you provided. A new search was conducted, and we were unable to identify records subject to the FOIPA that are responsive to your request.

☐ Records potentially responsive to your request were destroyed. Since this material could not be reviewed, it is not known if it was responsive to your request. Record retention and disposal is carried out under supervision of the National Archives and Records Administration (NARA) according to Title 44 United States Code Section 3301, Title 36 Code of Federal Regulations (CFR) Chapter 12 Sub-chapter B Part 1228, and 36 CFR 1229.10. Please be advised that the General Records Schedule (GRS) disposition authority for FOIPA records is DAA-GRS-2016-0002-0001 (GRS 4.2, Item 020).

☐ Records potentially responsive to your request were transferred to the National Archives and Records Administration (NARA). If you wish to review these records, file a FOIPA request with NARA at the following address:

National Archives and Records Administration
Special Access and FOIA
8601 Adelphi Road, Room 5500
College Park, MD 20740-6001

☐ Potentially responsive records were identified during the search. However, we were advised that they were not in their expected locations. An additional search for the missing records also met with unsuccessful results. Since we were unable to review the records, we were unable to determine if they were responsive to your request.

☐ The identification records requested are maintained by the FBI's Criminal Justice Information Services (CJIS) Division; therefore, we have forwarded a portion of your request to CJIS for processing. To check the status of this request, please contact CJIS directly at (304) 625-5590. For additional information, see the enclosed FBI FOIPA Addendum General Information Section.

☐ Requests for expedited processing are not applicable when a final response is issued within ten calendar days.

☐ Police departments should be aware that the search conducted was limited to FBI records. Requests for criminal history records or rap sheets should be directed to Criminal Justice Information Services (CJIS). Information regarding CJIS is listed in the enclosed FBI FOIPA Addendum General Information Section.

☐ Records potentially responsive to your request were transferred to the National Personnel Records Center - Civilian Personnel Records (NPRC-CPR). In order to obtain information on a file located at the NPRC, your request must be mailed to the following address:

National Archives and Records Administration
ATTN: Archival Programs
P.O. Box 38757
St. Louis, MO 63138

☐ You also requested information regarding one or more third parties. Please be advised the FBI will neither confirm nor deny the existence of such records pursuant to FOIA exemptions (b)(6) and (b)(7)(C), 5 U.S.C. §§ 552 (b)(6) and (b)(7)(C). The mere acknowledgement of the existence of FBI records on third party individuals could reasonably be expected to constitute an unwarranted invasion of personal privacy. This is our standard response to such requests and should not be taken to mean that records do, or do not, exist.

Please refer to the enclosed FBI FOIPA Addendum for additional standard responses applicable to your request. **"Part 1"** of the Addendum includes standard responses that apply to all requests. **"Part 2"** includes additional standard responses that apply to all requests for records about yourself or any third party individuals. **"Part 3"** includes general information about FBI records that you may find useful. Also enclosed is our Explanation of Exemptions.

Additional information about the FOIPA can be found at www.fbi.gov/foia. Should you have questions regarding your request, please feel free to contact foipaquestions@fbi.gov. Please reference the FOIPA Request number listed above in all correspondence concerning your request.

If you are not satisfied with the Federal Bureau of Investigation's determination in response to this request, you may administratively appeal by writing to the Director, Office of Information Policy (OIP), United States Department of Justice, 441 G Street, NW, 6th Floor, Washington, D.C. 20530, or you may submit an appeal through OIP's FOIA STAR portal by creating an account following the instructions on OIP's website: https://www.justice.gov/oip/submit-and-track-request-or-appeal. Your appeal must be postmarked or electronically transmitted within ninety (90) days of the date of this response to your request. If you submit your appeal by mail, both the letter and the envelope should be clearly marked "Freedom of Information Act Appeal." If possible, please provide a copy of your original request and this response letter with your appeal.

You may seek dispute resolution services by emailing the FBI's FOIA Public Liaison at foipaquestions@fbi.gov. The subject heading should clearly state "Dispute Resolution Services." Please also cite the FOIPA Request Number assigned to your request so it may be easily identified. You may also contact the Office of Government Information Services (OGIS). The contact information for OGIS is as follows: Office of Government Information Services, National Archives and Records Administration, 8601 Adelphi Road-OGIS, College Park, Maryland 20740-6001, e-mail at ogis@nara.gov; telephone at 202-741-5770; toll free at 1-877-684-6448; or facsimile at 202-741-5769.

Sincerely,

Michael G. Seidel
Section Chief
Record/Information Dissemination Section
Information Management Division

Enclosures

FBI FOIPA Addendum

As referenced in our letter responding to your Freedom of Information/Privacy Acts (FOIPA) request, the FBI FOIPA Addendum provides information applicable to your request. Part 1 of the Addendum includes standard responses that apply to all requests. Part 2 includes standard responses that apply to requests for records about individuals to the extent your request seeks the listed information. Part 3 includes general information about FBI records, searches, and programs.

Part 1: The standard responses below apply to all requests:

(i) **5 U.S.C. § 552(c).** Congress excluded three categories of law enforcement and national security records from the requirements of the FOIPA [5 U.S.C. § 552(c)]. FBI responses are limited to those records subject to the requirements of the FOIPA. Additional information about the FBI and the FOIPA can be found on the www.fbi.gov/foia website.

(ii) **Intelligence Records**. To the extent your request seeks records of intelligence sources, methods, or activities, the FBI can neither confirm nor deny the existence of records pursuant to FOIA exemptions (b)(1), (b)(3), and as applicable to requests for records about individuals, PA exemption (j)(2) [5 U.S.C. §§ 552/552a (b)(1), (b)(3), and (j)(2)]. The mere acknowledgment of the existence or nonexistence of such records is itself a classified fact protected by FOIA exemption (b)(1) and/or would reveal intelligence sources, methods, or activities protected by exemption (b)(3) [50 USC § 3024(i)(1)]. This is a standard response and should not be read to indicate that any such records do or do not exist.

Part 2: The standard responses below apply to all requests for records on individuals:

(i) **Requests for Records about any Individual—Watch Lists.** The FBI can neither confirm nor deny the existence of any individual's name on a watch list pursuant to FOIA exemption (b)(7)(E) and PA exemption (j)(2) [5 U.S.C. §§ 552/552a (b)(7)(E), (j)(2)]. This is a standard response and should not be read to indicate that watch list records do or do not exist.

(ii) **Requests for Records about any Individual—Witness Security Program Records.** The FBI can neither confirm nor deny the existence of records which could identify any participant in the Witness Security Program pursuant to FOIA exemption (b)(3) and PA exemption (j)(2) [5 U.S.C. §§ 552/552a (b)(3), 18 U.S.C. 3521, and (j)(2)]. This is a standard response and should not be read to indicate that such records do or do not exist.

(iii) **Requests for Confidential Informant Records.** The FBI can neither confirm nor deny the existence of confidential informant records pursuant to FOIA exemptions (b)(7)(D), (b)(7)(E), and (b)(7)(F) [5 U.S.C.§ § 552 (b)(7)(D), (b)(7)(E), and (b)(7)(F)] and Privacy Act exemption (j)(2) [5 U.S.C.§ 552a (j)(2)]. The mere acknowledgment of the existence or nonexistence of such records would reveal confidential informant identities and information, expose law enforcement techniques, and endanger the life or physical safety of individuals. This is a standard response and should not be read to indicate that such records do or do not exist.

Part 3: General Information:

(i) **Record Searches and Standard Search Policy.** The Record/Information Dissemination Section (RIDS) searches for reasonably described records by searching systems, such as the Central Records System (CRS), or locations where responsive records would reasonably be found. The CRS is an extensive system of records consisting of applicant, investigative, intelligence, personnel, administrative, and general files compiled by the FBI per its law enforcement, intelligence, and administrative functions. The CRS spans the entire FBI organization, comprising records of FBI Headquarters, FBI Field Offices, and FBI Legal Attaché Offices (Legats) worldwide; Electronic Surveillance (ELSUR) records are included in the CRS. The standard search policy is a search for main entity records in the CRS. Unless specifically requested, a standard search does not include a search for reference entity records or administrative records of previous FOIPA requests.

a. *Main Entity Records* – created for individuals or non-individuals who are the subjects or the focus of an investigation
b. *Reference Entity Records*- created for individuals or non-individuals who are associated with a case but are not known subjects or the focus of an investigation

(ii) **FBI Records.** Founded in 1908, the FBI carries out a dual law enforcement and national security mission. As part of this dual mission, the FBI creates and maintains records on various subjects; however, the FBI does not maintain records on every person, subject, or entity.

(iii) **Foreseeable Harm Standard.** As amended in 2016, the Freedom of Information Act provides that a federal agency may withhold responsive records only if: (1) the agency reasonably foresees that disclosure would harm an interest protected by one of the nine exemptions that FOIA enumerates, or (2) disclosure is prohibited by law (5 United States Code, Section 552(a)(8)(A)(i)). The FBI considers this foreseeable harm standard in the processing of its requests.

(iv) **Requests for Criminal History Records or Rap Sheets.** The Criminal Justice Information Services (CJIS) Division provides Identity History Summary Checks – often referred to as a criminal history record or rap sheet. These criminal history records are not the same as material in an investigative "FBI file." An Identity History Summary Check is a listing of information taken from fingerprint cards and documents submitted to the FBI in connection with arrests, federal employment, naturalization, or military service. For a fee, individuals can request a copy of their Identity History Summary Check. Forms and directions can be accessed at www.fbi.gov/about-us/cjis/identity-history-summary-checks. Additionally, requests can be submitted electronically at www.edo.cjis.gov. For additional information, please contact CJIS directly at (304) 625-5590.

EXPLANATION OF EXEMPTIONS

SUBSECTIONS OF TITLE 5, UNITED STATES CODE, SECTION 552

(b)(1) (A) specifically authorized under criteria established by an Executive order to be kept secret in the interest of national defense or foreign policy and (B) are in fact properly classified to such Executive order;

(b)(2) related solely to the internal personnel rules and practices of an agency;

(b)(3) specifically exempted from disclosure by statute (other than section 552b of this title), provided that such statute (A) requires that the matters be withheld from the public in such a manner as to leave no discretion on issue, or (B) establishes particular criteria for withholding or refers to particular types of matters to be withheld;

(b)(4) trade secrets and commercial or financial information obtained from a person and privileged or confidential;

(b)(5) inter-agency or intra-agency memorandums or letters which would not be available by law to a party other than an agency in litigation with the agency;

(b)(6) personnel and medical files and similar files the disclosure of which would constitute a clearly unwarranted invasion of personal privacy;

(b)(7) records or information compiled for law enforcement purposes, but only to the extent that the production of such law enforcement records or information (A) could reasonably be expected to interfere with enforcement proceedings, (B) would deprive a person of a right to a fair trial or an impartial adjudication, (C) could reasonably be expected to constitute an unwarranted invasion of personal privacy, (D) could reasonably be expected to disclose the identity of confidential source, including a State, local, or foreign agency or authority or any private institution which furnished information on a confidential basis, and, in the case of record or information compiled by a criminal law enforcement authority in the course of a criminal investigation, or by an agency conducting a lawful national security intelligence investigation, information furnished by a confidential source, (E) would disclose techniques and procedures for law enforcement investigations or prosecutions, or would disclose guidelines for law enforcement investigations or prosecutions if such disclosure could reasonably be expected to risk circumvention of the law, or (F) could reasonably be expected to endanger the life or physical safety of any individual;

(b)(8) contained in or related to examination, operating, or condition reports prepared by, on behalf of, or for the use of an agency responsible for the regulation or supervision of financial institutions; or

(b)(9) geological and geophysical information and data, including maps, concerning wells.

SUBSECTIONS OF TITLE 5, UNITED STATES CODE, SECTION 552a

(d)(5) information compiled in reasonable anticipation of a civil action proceeding;

(j)(2) material reporting investigative efforts pertaining to the enforcement of criminal law including efforts to prevent, control, or reduce crime or apprehend criminals;

(k)(1) information which is currently and properly classified pursuant to an Executive order in the interest of the national defense or foreign policy, for example, information involving intelligence sources or methods;

(k)(2) investigatory material compiled for law enforcement purposes, other than criminal, which did not result in loss of a right, benefit or privilege under Federal programs, or which would identify a source who furnished information pursuant to a promise that his/her identity would be held in confidence;

(k)(3) material maintained in connection with providing protective services to the President of the United States or any other individual pursuant to the authority of Title 18, United States Code, Section 3056;

(k)(4) required by statute to be maintained and used solely as statistical records;

(k)(5) investigatory material compiled solely for the purpose of determining suitability, eligibility, or qualifications for Federal civilian employment or for access to classified information, the disclosure of which would reveal the identity of the person who furnished information pursuant to a promise that his/her identity would be held in confidence;

(k)(6) testing or examination material used to determine individual qualifications for appointment or promotion in Federal Government service the release of which would compromise the testing or examination process;

(k)(7) material used to determine potential for promotion in the armed services, the disclosure of which would reveal the identity of the person who furnished the material pursuant to a promise that his/her identity would be held in confidence.

epic.org

Electronic Privacy Information Center
1519 New Hampshire Avenue NW
Washington, DC 20036, USA

+1 202 483 1140
+1 202 483 1248
@EPICPrivacy
https://epic.org

December 19, 2024

VIA FOIA STAR

Bobak Talebian
Director, Office of Information Policy (OIP)
United States Department of Justice
441 G Street, NW, 6th Fl.
Washington, DC 20001

RE: Freedom of Information Act Appeal, FBI FOIPA Request No. 1649273-00

Dear Director Talebian:

This letter constitutes an appeal under the Freedom of Information Act (FOIA), 5 U.S.C. § 552(a)(6)(A), and is submitted to the Director of the Office of Information Policy in the United States Department of Justice by the Electronic Privacy Information Center (EPIC).

This appeal arises from EPIC's October 4, 2024, request (EPIC's FOIA Request), Appendix A, seeking any data mining reports created by the Federal Bureau of Investigation (FBI) pursuant to the Federal Agency Data Mining Report Act of 2007, 42 U.S.C. § 2000ee-3. This appeal challenges the sufficiency of the FBI's search for responsive records as well as the FBI's decision to close EPIC's FOIA Request.

The FBI failed to meet its obligations to make a "good faith effort" to search for requested records because it did not search all relevant offices, search beyond the Central Record Service, disclose its search queries, or consult all relevant individuals most likely to hold relevant information, including the privacy officer specified by the Act.[1]

Procedural Background

On October 4, 2024, EPIC submitted a FOIA Request via the FBI's eFOIPA portal.[2] EPIC's FOIA Request included a request for a fee waiver.[3] EPIC's FOIA Request was assigned FOIA Request No. 1649273-00 by the eFOIPA system.

On October 16, 2024, the FBI emailed two letters to EPIC. The first letter acknowledged EPIC's FOIA Request and granted EPIC's requested fee waiver.[4] The second letter informed EPIC that our request was subject to "unusual circumstances" which would delay the agency's ability to

[1] *See* 42 U.S.C. § 2000ee-3(c)(1).
[2] *See* Appendix A (EPIC's FOIA Request).
[3] *Id*.
[4] *See* Appendix B (FBI's Confirmation Letter).

make a determination on EPIC's FOIA Request.[5] The letter did not clarify which unusual circumstances applied.

EPIC emailed the FBI's FOIA office on November 1, 2024, to request an update on EPIC's FOIA Request. This email asked for clarification on which unusual circumstance scenarios described in the FBI's October 16, 2024, letter applied to EPIC's FOIA Request and requested confirmation that the agency would be able to respond within the extended 30-day window.

On November 6, 2024, the FBI sent a letter ("FBI's Final Response Letter") to EPIC stating that, following a "main entity record search of the Central Records System," no responsive documents were identified.[6] It also included a FOIPA Addendum with "standard responses" including "general information about FBI records, searches, and programs."[7] Main entity records are those "created for individuals or non-individuals who are the subjects or the focus of an investigation."[8] The same letter informed EPIC that its FOIA Request was being closed.[9]

EPIC Appeals the Sufficiency of the FBI's Search for Responsive Records and the Improper Closure of EPIC's FOIA Request

EPIC challenges the sufficiency of the FBI's search for responsive records and the closure of EPIC's FOIA Request. An agency must "show that it 'conducted a search reasonably calculated to uncover all relevant documents" in order to fulfill its obligations under FOIA. *Freedom Watch, Inc. v. Nat'l Sec. Agency*, 49 F. Supp. 3d 1, 5 (D.D.C. 2014) (quoting *Weisberg v. Dep't of the Army*, 920 F.2d 1344, 1351 (D.C. Cir. 1983)), *aff'd and remanded*, 783 F.3d 1340 (D.C. Cir. 2015). It is the agency's burden to "show that it made a good faith effort to conduct a search for the requested records, using methods which can be reasonably expected to produce the information requested." *Oglesby v. U.S. Dep't of Army*, 920 F.2d 57, 68 (D.C. Cir. 1990). The FBI cannot meet its burden to show that it has made a good faith effort nor that it used methods which can be reasonably expected to produce documents that are responsive to EPIC's FOIA Request seeking data mining reports created by the agency.

The agency provides no details specific to EPIC's FOIA Request regarding the search methodology. In its November 6, 2024, letter, the agency stated that it conducted only a "main entity record search."[10] It does not clarify whether "main entity" records include all or merely a subset of all files collected in the CRS.[11] The agency also does not provide the specific search terms used to search the CRS.[12] EPIC's FOIA Request was highly particularized: it sought Data

[5] *See* Appendix C (FBI's Letter of Unusual Circumstances); 5 U.S.C. § 552(a)(6)(B)(iii).
[6] *See* Appendix D (FBI's Letter Closing FOIA Request).
[7] *Id*.
[8] *Id*.
[9] *Id*.
[10] *Id*.
[11] *Id*.
[12] *Id*.

Mining Reports created by the agency in response to a federal statute.[13] Responsive documents can easily be missed if improper, overly narrow, or overly broad search terms are used. However, the agency's letter does not provide enough information for EPIC to evaluate the likelihood of this possibility.

Further, it is unclear whether a main entity search of the CRS could produce responsive documents at all. The data mining reports required by the Act are not "created for individuals or non-individuals who are the subjects or the focus of an investigation."[14] Indeed, these reports are mandated to be created if the agency engages in data mining, whether or not this is connected to specific investigations. If main entity records are only connected to investigations, as the FBI's appendix to its letter states, a main entity records search may not yield responsive documents.

The FBI's search was also inadequate because the agency did not consult those most likely to possess responsive documents, despite the likelihood that such documents would be missing from main entity records. Namely, the FBI failed to consult its Privacy Officer when conducting its search. The Federal Agency Data Mining Report Act requires that the head of an agency work with the agency's privacy officer to produce its data mining reports.[15] Additionally, the FBI is required to designate a senior officer to serve as the agency's principal advisor on privacy and civil liberties issues.[16] The FBI's Privacy and Civil Liberties Officer (PCLO) fills this role for the agency as part of the Office of General Counsel and is supported by the Privacy and Civil Liberties Unit.[17] However, the FBI did not indicate that it consulted the PCLO or any of its staff in the Unit.

There is evidence that responsive records exist. EPIC's FOIA Request detailed several cases where the FBI was or was likely to be using data mining.[18] In addition to that evidence, Director Wray's Congressional testimony explained that the FBI Laboratory employs digital forensics,[19] an investigational method that employs several data mining techniques.[20] Additionally, the FBI's business solicitations show that it plans to employ pattern-based data mining techniques in its future work. For example, the agency is actively soliciting solutions that involve data mining,

---

[13] *See* Appendix A.
[14] *See* Appendix D.
[15] 42 U.S.C. § 2000-ee(c)(1).
[16] FISA Amendments Reauthorization Act of 2017, Pub. L. No. 115-118, § 109, 132 Stat. 3, 15 (2018).
[17] *See* FBI, Privacy and Civil Liberties Semi-Annual Report (2023) https://www.justice.gov/usdoj-media/opcl/media/1344026/dl?inline.
[18] *See* Appendix A.
[19] Christopher Wray, Statement for the Record on Oversight of the Before the House Judiciary Committee (Jul. 24, 2024) https://www.fbi.gov/news/testimony/oversight-of-the-federal-bureau-of-investigation-072424.
[20] Dipo Dunsin et al., *A Comprehensive Analysis of the Role of Artificial Intelligence and Machine Learning in Modern Digital Forensics and Incident Response*, 48 Forensic Science International: Digital Investigation (Mar. 2024) https://www.sciencedirect.com/science/article/pii/S2666281723001944?ref=pdf_download.

including AI tools for identifying data relationships, locating trends and patterns, and conducting threat analysis.[21]

Finally, the FBI improperly closed EPIC's FOIA Request. Responsive records exist, and the FBI failed to meet its obligation to make a good faith effort in conducting its search. By failing to disclose its search terms, consider the type of record sought when choosing to conduct a main entity search, or consult colleagues likely to possess responsive records, the FBI did not conduct a search reasonably calculated to return responsive documents.

For these reasons, the FBI did not conduct a sufficient search for responsive records. OIP should find in favor of EPIC and require the agency to reopen EPIC's FOIA request and conduct a sufficient search for responsive records.

Conclusion

Thank you for your consideration of this appeal. EPIC anticipates a determination on this appeal within 20 working days as required by 5 U.S.C. § 552(a)(6)(A)(ii). For questions regarding this appeal, please contact Abigail Kunkler at FOIA@epic.org, cc: kunkler@epic.org.

Respectfully submitted,

/s Abigail A. Kunkler
Abigail Kunkler
EPIC Law Fellow

*s/*Jeramie D. Scott
Jeramie D. Scott
EPIC Senior Counsel

Enclosures

[21] FBI, *FBI Enterprise BAA Problem Sets*, https://biz.fbi.gov, https://biz.fbi.gov/broad-agency-announcement-baa/fbi-enterprise-baa-problem-sets.

U.S. Department of Justice
Office of Information Policy
Sixth Floor
441 G Street, NW
Washington, DC 20530-0001

Telephone: (202) 514-3642

Jeramie Scott, Esq.

Re: Appeal No. A-2025-00568
Request No. 1649273-00

jscott@epic.org

VIA: Online Portal - 07/11/2025

Dear Jeramie Scott:

You appealed from the action of the Federal Bureau of Investigation (FBI) on your Freedom of Information Act (FOIA) request for access to "any and all reports on the FBI's data mining activities generated by the FBI in response to the Federal Data Mining Reporting Act of 2007." I note that your appeal concerns the adequacy of the FBI's search.

After carefully considering your appeal, I am affirming the FBI's action on your request. The FBI informed you that it could locate no responsive records subject to the FOIA in its files. Please be advised that the FBI inadvertently informed you that it searched its Central Records System (CRS) for responsive records; rather, the FBI searched multiple offices, including the Office of Congressional Affairs and Office of General Counsel. I have determined that the FBI's action was correct and that it conducted an adequate, reasonable search for such records.

Please be advised that this Office's decision was made only after a full review of this matter. Your appeal was assigned to an attorney with this Office who thoroughly reviewed and analyzed your appeal, your underlying request, and the action of the FBI in response to your request.

If you are dissatisfied with my action on your appeal, the FOIA permits you to file a lawsuit in federal district court in accordance with 5 U.S.C. § 552(a)(4)(B).

For your information, the Office of Government Information Services (OGIS) offers mediation services to resolve disputes between FOIA requesters and Federal agencies as a non-exclusive alternative to litigation. Using OGIS services does not affect your right to pursue litigation. The contact information for OGIS is as follows: Office of Government Information Services, National Archives and Records Administration, Room 2510, 8601 Adelphi Road,

College Park, Maryland 20740-6001; email at ogis@nara.gov; telephone at 202-741-5770; toll-free at 1-877-684-6448; or facsimile at 202-741-5769. If you have any questions regarding the action this Office has taken on your appeal, you may contact this Office and speak with the undersigned agency official by calling 202-514-3642.

Sincerely,

X _____

Christina Troiani
Chief, Administrative Appeals Staff