The Imbalance of Security & Privacy: What the Snowden Revelations Contribute to the Data Mining Debate

GEORGE GUTIERREZ*

INTRODUCTION

Data mining is the process of searching data for previously unknown patterns and using these patterns to predict future outcomes.[1] Recently, the United States has been embroiled in a fiery debate concerning the government's ability to use data mining as a tool to protect domestic security.[2] Critics of the government using data mining argue that the practice fails to adequately address terrorism, burns through financial capital, and ineffectually uses scarce resources, which collectively threaten civil liberties.[3] However, data-mining supporters argue that it provides useful counterterrorism intelligence that enhances domestic security and ensures the safety of Americans.[4]

Whistleblower, Edward Snowden's, recent revelations have given the public access to information regarding previously undisclosed government data mining programs.[5] This Article uses the work of two leading scholars to present the data mining debate as it stood before the Snowden leaks, and argues that the government programs exposed by the leaks undermine the previous arguments of data mining supporters.

Part I provides an overview of data mining and briefly discusses

---

*Mr. Gutierrez obtained his J.D. degree in 2014 from the University of San Francisco School of Law, and his B.A. in Political Science from California State University, Long Beach. The Author owes special thanks to Professor Susan Freiwald for her expertise and thoughtful guidance throughout the writing process. The Author would also like to thank his parents for their enduring love and support.

1. Jeff Jonas, *What is Data Mining? Depends Who You Ask . . .*, JEFF JONAS (Sept. 8, 2006), http://jeffjonas.typepad.com/jeff_jonas/2006/09/what_is_data_mi.html.

2. *See* Byron Acohido, *Is Government Data Mining Necessary to Keep Us Safe?*, USA TODAY (June 7, 2013), http://www.usatoday.com/story/cybertruth/2013/06/07/data-mining-prism-government-cyberspying/2399259/.

3. Jeff Jonas & Jim Harper, *Effective Counterterrorism and the Limited Role of Predictive Data Mining*, 584 POL'Y ANALYSIS 1 (2006), *available at* http://object.cato.org/sites/cato.org/files/pubs/pdf/pa584.pdf.

4. *See generally* Shane Ham & Robert Atkinson, *Using Technology to Detect and Prevent Terrorism*, PROGRESSIVE POL'Y INST. (Jan. 2002) (noting that the information technology revolution will make domestic defense easier, less expensive, and more effective making all Americans safer).

5. *See infra* Part IV.

the controversial characteristics of the government's data-mining programs. Part II presents the opposing arguments against data mining, promulgated by noted scholars, Jeff Jonas and Jim Harper. Part III addresses expert, Kim Taipale's, arguments in favor of data mining. Part IV evaluates the strength and relevance of each side's arguments in light of the new information provided by the Snowden revelations. This Article concludes that in order to have effective discourse about government data-mining, it is imperative to recognize the flaws in security that Snowden exposed.

I. DATA MINING—WHAT IS IT? WHY IS IT CONTROVERSIAL?

A. DATA MINING OVERVIEW

Data mining is a complex process that is not uniformly defined. For example, the Government Accountability Office defines data mining as "the application of database technology and techniques—such as statistical analysis and modeling—to uncover hidden patterns and subtle relationships in data and to infer rules that allow for the prediction of future results."[6] In comparison, the Congressional Research Service's definition of data mining is "the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. These tools can include statistical models, mathematical algorithms, and machine learning methods (algorithms that improve their performance automatically through experience, such as neural networks or decision trees)."[7] Scholars Jeff Jonas and Jim Harper describe data mining as "the process of searching data for previously unknown patterns and using those patterns to predict future outcomes."[8]

The lack of a precise definition makes it challenging to reach a consensus on whether the government's use of data mining is effective against terrorism and ultimately worth sacrificing civil liberties to continue.[9] Scholars seem to agree that the varying definitions sometimes prevent a productive dialogue between supporters and opposers because there is a lack of consensual understanding of what data mining is.[10] This Article assumes that data mining technologies use advanced computer software to predict patterns from large amounts of aggregated data using specific

---

6. U.S. GEN. ACCOUNTABILITY OFFICE, GAO-04-548, DATA MINING: FEDERAL EFFORTS COVER A WIDE RANGE OF USES 4 (2004), *available at* http://www.gao.gov/new.items/d04548.pdf.

7. JEFFREY W. SEIFERT, CONG. RES. SERV., RL31798, DATA MINING AND HOMELAND SECURITY: AN OVERVIEW 1 (2007), *available at* http://www.fas.org/sgp/crs/intel/RL31798.pdf.

8. Jonas & Harper, *supra* note 3, at 1.

9. *Id.* at 5 ("'Data mining' is a broad and fairly loaded term that means different things to different people. . . . [C]ollective failure to get to the root of the term 'data mining' may have preserved disagreements among people who may be in substantial agreement.").

10. *See* Ham & Atkinson, *supra* note 4.

computer algorithms.[11] Further, this Article accepts Jonas and Harper's definition of data mining as only a subset of the larger category of data analysis.[12]

Within this category of data mining are two subcategories—subject-based and pattern-based analysis.[13] Subject-based or "link analysis" data mining "seeks to trace links from known individuals or things to others."[14] This type of analysis begins with "information about specific suspects, combined with general knowledge."[15] The criticisms relating to data mining this Article presents exclude link analysis and focus specifically on predictive pattern-based data mining. In contrast to subject-based analysis, pattern-based predictive analysis "use[s] statistical probabilities to seek predicates in large data sets."[16] Jonas and Harper describe the process of pattern-based data mining as "seek[ing] to find new knowledge, not from the investigative and deductive process of following specific leads, but from statistical, inductive processes. Because it is more characterized by prediction than by the traditional notion of suspicion, we refer to it as 'predictive data mining.'"[17]

B. DATA MINING CONTROVERSY

The controversy surrounding data mining stems from a lack of consensus on how to address the privacy and civil liberties concerns[18] that it implicates. One of the most controversial aspects of data mining is its ability to indiscriminately collect massive amounts of aggregated data from millions of Americans.[19] Those opposing data mining argue that this bulk collection is ineffective in predicting and preventing acts of terrorism, and any minimal uses it has to address this problem are outweighed by the privacy and civil liberties concerns it creates.[20] Some scholars view the issue as coming up with the most effective way to balance the interests between enhancing

---

11. *What is Data Mining (Predictive Analytics, Big Data)*, DELL SOFTWARE, http://www.statsoft.com/Textbook/Data-Mining-Techniques (last visited Apr. 17, 2015).

12. Jonas & Harper, *supra* note 3, at 5.

13. *Id.* at 6.

14. *Id.*

15. *Id.*

16. *Id.*

17. *Id.*

18. *See* Susan Freiwald, *Nothing to Fear or Nowhere to Hide: Competing Visions of the NSA's 215 Program*, 12 COLO. TECH. L.J. 309 (2014).

19. Am. Civil Liberties Union v. Clapper, 959 F. Supp. 2d 724, 746 (S.D.N.Y. 2013) ("[W]ithout all the data points, the Government cannot be certain it connected the pertinent ones. As FISC Judge Eagan noted, the collection of virtually all telephony metadata is 'necessary' to permit the NSA, not the FBI, to do the algorithmic data analysis that allow the NSA to determine connections between known and unknown international terrorist operatives.") (internal quotation marks omitted). However, mining with smaller data sets can be and is conducted as well. *See* Razvan Andonie, *Extreme Data Mining: Inferences from Small Datasets*, 5 INT. J. COMP. COMM'NS & CONTROL 280 (2010).

20. *See supra* text accompanying note 3.

national security and respecting citizen's privacy.[21] In contrast, civil libertarians dispute the very existence of such a tradeoff because they maintain privacy can be upheld without resulting losses to security because data mining is ineffective and never enhanced security to begin with.[22] They further argue that the zealous pursuit of security will trample citizens' privacy rights, sidestepping the Fourth Amendment's requirements of "individualized suspicion" and "particularity" before the government begins a search.[23] In response, the government argues that preventing another terrorist attack should be the nation's top priority and any privacy loss is a necessary tradeoff.[24] Furthermore, government litigators stress that the government should be trusted to secure Americans' privacy rights when in pursuit of national security counterterrorism objectives.[25] As Part II discusses, there are numerous disagreements as to the appropriateness of government data mining practices.

II. JONAS AND HARPER'S ARGUMENTS AGAINST DATA MINING

Jeff Jonas, chief scientist of IBM's Entity Analytics group, has contributed significantly to the data mining debate.[26] His work has been featured in the Wall Street Journal, Washington Post, Fortune Magazine, MSNBC, and on NPR.[27] Jonas has repeatedly taken the position that data mining is not particularly useful for counterterrorism purposes and needlessly infringes on privacy and civil liberties.[28] Jim Harper is a CATO Institute senior fellow, whose areas of expertise include privacy, telecommunications, intellectual property, government transparency, and cybersecurity.[29] As a preliminary matter, Jonas and Harper argue that data mining is

---

21. *See* Daniel J. Solove, *Data Mining and the Security-Liberty Debate*, 75 U. CHI. L. REV. 343, 345 (2008).

22. Eugene Volokh, *NSA Data Mining and the False Trade-Off Between Privacy and Security*, VOLOKH CONSPIRACY (May 31, 2006), http://www.volokh.com/2006/05/31/nsa-data-mining-and-the-false-trade-off-between-privacy-and-security/.

23. Jim Harper, *Data Mining or the Fourth Amendment?*, CATO INST. (Aug. 22, 2006), http://www.cato.org/blog/data-mining-or-fourth-amendment.

24. *Obama Defends Surveillance Programs After Revelations US Authorities Collecting Phone Records, Mining Data*, ABC NEWS (July 12, 2013), http://www.abc.net.au/news/2013-06-08/obama-defends-us-surveillance-programs/4741664.

25. *See* Freiwald, *supra* note 18 (providing a greater analysis of the "trust problem").

26. Alex Woodie, *Jeff Jonas Explores the Nature of Data in COMMON Keynote*, IT JUNGLE (June 24, 2009), http://www.itjungle.com/tfh/tfh051809-story03.html.

27. *Jeff Jonas Biography*, JEFF JONAS, http://jeffjonas.typepad.com/about.html (last visited Apr. 17, 2015).

28. Jeff Jonas, *110th Congress Debates Data Mining*, JEFF JONAS (Jan. 21, 2007), http://jeffjonas.typepad.com/jeff_jonas/2007/01/110th_congress_.html. *See also* Barton Gellman, Dafna Linzer & Carol D. Leonnig, *Surveillance Net Yields Few Suspects*, WASH. POST (Feb. 5, 2006), http://www.washingtonpost.com/wp-dyn/content/article/2006/02/04/AR2006020401373.html; Jeff Jonas, *Sometimes a Big Picture is Worth a 1,000 False Positives*, JEFF JONAS (May 7, 2006), http://jeffjonas.typepad.com/jeff_jonas/2006/05/sometimes_a_big.html.

29. *Jim Harper*, CATO INST., http://www.cato.org/people/jim-harper (last visited Apr. 17, 2015).

ineffective at stopping terrorist attacks.

The notion that data mining is useful in preventing terrorism is the government's most frequently cited reason for continuing data-mining practices.[30] Jonas and Harper disagree with this assertion, and contrast data mining used for national security purposes with data mining used to discover identity theft and credit card fraud.[31] When data mining is used in the latter scenario, data analysts can examine thousands of instances of fraud that occur every year to create a profile of the type of behavior or transactions most frequently associated with suspicious credit activity.[32] Jonas and Harper argue that acts of terrorism do not occur as frequently, and consequently do not yield a similar set of incidents.[33] They explain, "With a relatively small number of attempts every year and only one or two major terrorist incidents every few years—each one distinct in terms of planning and execution—there are no meaningful patterns that show what behavior indicates planning or preparation for terrorism."[34] Terrorist attacks are infrequent, unlike the fraudulent uses of credit cards, thus it is not possible to create credible predictive models for terrorist attacks.[35] The absence of useful patterns renders any algorithm designed to identify suspicious behavior useless because such algorithm will produce a large number of "false positives."[36]

Jonas and Harper also call attention to the high costs of data mining resulting from the frequent number of false positives which it generates.[37] A false positive occurs when a test incorrectly evidences the existence of some pattern or circumstance such as a terrorist plot.[38] Citing the work of Jeffrey Rosen, a law professor at George Washington University, Jonas and Harper highlight the large rate of false positives that would occur in a system designed to prevent the September 11th attacks.[39] They argue, "Assuming a 99 percent accuracy rate, searching our population of nearly 300,000,000, some 3,000,000 people would be identified as potential terrorists."[40] The

---

30. *See supra* note 24.

31. Jim Harper, *Balancing Privacy and Security: The Privacy Implications of Government Data Mining Programs*, CATO INST. (Jan. 10, 2007), http://www.cato.org/publications/congressional-testimony/balancing-privacy-security-privacy-implications-government-data-mining-programs.

32. JIM HARPER, IDENTITY CRISIS: HOW IDENTIFICATION IS OVERUSED AND MISUNDERSTOOD 219 (2006).

33. Jim Harper, *Data Mining Can't Improve Our Security*, CATO INST. (Dec. 7, 2006), http://www.cato.org/publications/commentary/data-mining-cant-improve-our-security.

34. Jonas & Harper, *supra* note 3, at 7–8.

35. *Id.* at 8.

36. *See supra* text accompanying note 34.

37. Jonas & Harper, *supra* note 3, at 8.

38. Bruce Schneier, *Data Mining for Terrorists*, SCHNEIER ON SEC. (Mar. 9, 2006), https://www.schneier.com/blog/archives/2006/03/data_mining_for.html.

39. JEFFREY ROSEN, THE NAKED CROWD: RECLAIMING SECURITY AND FREEDOM IN AN ANXIOUS AGE 104–07 (2004).

40. Jonas & Harper, *supra* note 3, at 8.

high number of false positives data mining produces squanders our limited national resources and detracts from legitimate leads extracted from better intelligence methods.[41]

While the outcome of the cost-benefit analysis varies depending on one's worldview,[42] at some point it should be recognized that the large number of false positives data mining produces infringes too much on the civil liberties of Americans with no countervailing benefit. Jonas and Harper highlight the fact that information regarding the September 11th hijackers was available to the U.S. government prior to the attacks, and the terrorists were interconnected in various ways.[43] Nawaf al-Hazmi and Khalid al-Midhar, two of the terrorists responsible for the September 11th attacks, were allowed to enter and reside in the United States despite having been linked to the prior U.S.S. Cole bombing.[44] The lesson taken from this failure is the need to focus more acutely on the traditional intelligence methods already available to our intelligence agencies, rather than to aggregate greater amounts of data.

III. KIM TAIPALE'S ARGUMENTS IN FAVOR OF DATA MINING

Kim Taipale is a noted scholar who specializes in information, technology, and national security policy.[45] Taipale argues that data mining is a useful tool in preventing terrorism and supports its use in counterterrorism pursuits.[46] Taipale admits that while data mining techniques are not perfected, this fact should not warrant complete dismissal of the technology.[47] Taipale rejects arguments from those who view privacy and security as mutually exclusive.[48] Instead he perceives them as "dual obligations of a liberal democracy."[49] He notes that "[t]he current public debate [presenting] security and privacy as dichotomous rivals to be traded one for another in a zero-sum game is based on a general misunderstanding and apprehension

41. *Id.* at 9; *see also Balancing Privacy and Security: The Privacy Implications of Government Data Mining Programs: Hearing Before the S. Comm. on the Judiciary*, 110th Cong., 1st Sess. 8–10 (2007) (statement of Jim Harper, Dir. of Info. Pol'y Studs., CATO Inst.).

42. *See* Freiwald, *supra* note 18, at 319 n.62.

43. Jonas & Harper, *supra* note 3, at 2.

44. Jim Harper, *Seth Stodder's Weak Defense of ATS-P*, CATO INST. (Mar. 21, 2008), http://www.cato.org/blog/seth-stodders-weak-defense-ats-p; *see also USS Cole Bombing*, 9/11 MEM'L, http://www.911memorial.org/uss-cole-bombing (last visited Apr. 17, 2015).

45. *Kim Taipale Info.*, TAIPALE, http://www.taipale.org (last visited Apr. 17, 2015).

46. Kim Taipale, *Whispering Wires and Warrantless Wiretaps: Data Mining and Foreign Intelligence Surveillance*, N.Y.U. REV. L. & SEC., No. VII: *The NSA and the War on Terror* (Supp. Spring 2006), *available at* http://ssrn.com/abstract=889120.

47. Kim Taipale, *Data Mining and Domestic Security: Connecting the Dots to Make Sense of Data*, 5 COLUM. SCI. & TECH. L. REV. 1, 57–59 (2003).

48. *Balancing Privacy and Security: The Privacy Implications of Government Data Mining Programs: Hearing Before the S. Comm. on the Judiciary*, 110th Cong., 1st Sess. 12–15 (2007) (testimony of Kim Taipale).

49. Kim Taipale, *Technology, Security and Privacy: The Fear of Frankenstein, the Mythology of Privacy and the Lessons of King Ludd*, 7 YALE J. L. & TECH. 123, 127 (2005).

of technology on the one hand and a mythology of privacy that conflates secrecy with autonomy on the other."[50] Taipale further argues that increasing use of data aggregation is inevitable and relying exclusively on legal mechanisms and sanctions to address consequential privacy concerns is unrealistic.[51] Instead, he proposes that privacy protecting features be "built into the design" of these technologies.[52] Furthermore, he faults privacy lobbyists and their "fetish for absolute secrecy" for "delay[ing] the development of appropriate technologies to improve security while also protecting civil liberties, and leav[ing] us with little security and brittle privacy protection."[53] Taipale maintains that "'strangling nascent technology in its crib'" before giving it an opportunity to be properly developed is not fair and that technology and appropriate safeguards can adequately address the privacy concerns implicated by data mining.[54]

A. DEVELOPING TECHNOLOGIES TO ADDRESS PRIVACY CONCERNS

Taipale contends that use of developing technologies can mitigate the privacy concerns of data mining.[55] Preliminarily, he argues that "rule-based processing," and a "distributed database architecture" can address data aggregation concerns.[56] Rule-based processing is the idea that "rules" can be attached to data as a method of enforcing privacy restrictions.[57] To the extent that a computer is used to query distributed databases, it would be required to "negotiate access and permitted uses with each database."[58] Therefore queries marked as pursuant to a warrant would have different permissions attached to it than one labeled as pursuant to "subpoena or administrative authorization."[59] Rule-based processing could also allow for data that can be labeled with meta-data indicating how it is to be processed.[60] Thus "even if a data item is removed or copied to a central database, it retains relevant rules by

50. *Id.* at 126.

51. *Id.* at 137–39; *see also* Taipale, *supra* note 47, at 17 ("[The] practical problem with efforts to simply block particular government research and development projects or outlaw specific technologies over privacy concerns is that 'the genie is already out of the bottle.' Resisting developments that have already occurred or will occur elsewhere regardless of whether any particular government project . . . is shut down seems futile and counter-productive.").

52. Taipale, *supra* note 49, at 140.

53. *Id.* at 137–38.

54. Taipale, *supra* note 47, at 70 (quoting Paul Rosenzweig, *Proposals for Implementing the Terrorism Information Awareness System*, HERITAGE FOUND. (Aug. 7, 2003), http://www.heritage.org/research/reports/2003/08/proposals-for-implementing-the-terrorism-information-awareness-system).

55. Taipale, *supra* note 49, at 199.

56. *Id.*

57. Taipale, *supra* note 47, at 75–76.

58. *Id.* at 76.

59. *Id.*

60. *Id.*

which it must be processed."[61] Taipale notes:

> [A] data item may be returned in encrypted form in which only subsequent processing under a warrant or pursuant to a security clearance is permitted. Alternatively, a particular data item may be labeled as belonging to a U.S. citizen or to a foreign national, or its original source labeled as from a particular government or commercial database. In each case different procedures developed to enforce particular policy decision and privacy concerns would apply in its subsequent processing.[62]

Furthermore, "strong tamper-proof audit mechanisms" must be structured into the data mining technology itself to prevent governmental abuse in addition to attacks by outsiders.[63] Such protections could include the use of a privacy appliance—"'hardware and software that sits on top of a database, which is controlled by some appropriate oversight authority, and has mechanisms to enforce access rules and accounting policy.'"[64] Taipale argues this type of mechanism would prevent the tampering of data because access to restricted data would trigger an "immutable digital record" documenting the offense.[65] Similarly, self-reporting data technologies—"data that, when accessed, reports its location and who is accessing it to an automated log or central tracking system"—can further accountability.[66]

Taipale makes strong, rational arguments. If our intelligence analysts actually implemented all of his suggestions, perhaps the concerns about government data mining would be mitigated. However, Snowden's disclosures of government data mining practices reveal that these ideas are far from being implemented.

IV. THE SNOWDEN REVELATIONS

On June 5, 2013, *The Guardian* published classified government documents detailing an immense data collection program.[67] Edward Snowden, a former contractor with the National Security Agency (NSA), was responsible for the newspaper's disclosure.[68] The documents revealed that the NSA ordered Verizon to turn over millions of telephone records, which included "'metadata', or

---

61. *Id.*

62. *Id.*

63. *Id.* at 80.

64. *Id.* (quoting INFO. AWARENESS OFFICE, REPORT TO CONGRESS REGARDING THE TERRORISM INFORMATION AWARENESS PROGRAM: IN RESPONSE TO CONSOLIDATED APPROPRIATIONS RESOLUTION, 203, PUB. L. NO. 108-7, DIVISION M, § 111(B) at A-13 (2003)).

65. *Id.*

66. *Id.*

67. Glenn Greenwald, *NSA Collecting Phone Records of Millions of Verizon Customers Daily*, THE GUARDIAN (June 6, 2013), http://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order.

68. *Edward Snowden*, BIOGRAPHY.COM, http://www.biography.com/people/edward-snowden-21262897 (last visited Apr. 17, 2015).

transactional information" revealing, "the numbers of both parties on a call . . . location data, call duration, unique identifiers, and the time and duration of all calls."[69] Since then, a plethora of classified information has been disclosed, igniting a public outcry and renewing a national debate about the legality and necessity of data mining programs for use in counterterrorism endeavors.[70] The disclosures revealed numerous government abuses, forcing explanations from the NSA,[71] technology companies including Google[72] and Facebook,[73] and the President of the United States.[74]

A. THE 215 PROGRAM

The Foreign Intelligence Surveillance Court (FISC) authorized the NSA to conduct a bulk surveillance data-mining program (the "215 Program"), pursuant to Section 215 of the USA PATRIOT Act ("Patriot Act").[75] Under the 215 Program, the NSA is authorized to collect millions of phone records ("metadata"), but not the actual content of the associated phone calls.[76] Acting upon this authority, the NSA collected these records in bulk, rather than limiting their collection to those suspected of wrongdoing.[77] The idea behind the program was that by aggregating large amounts of data on millions of Americans, the NSA could use such data to detect a pattern of activity typically associated with terrorism and find the individuals who "fit the pattern."[78] The NSA required Verizon to disclose, on an ongoing daily basis "all call detail records, or 'telephony metadata' created by Verizon for communications (i) between the United States and abroad; or (ii) wholly within the United States, including local

69. Greenwald, *supra* note 67.

70. David Cole, *'No Place to Hide' by Glenn Greenwald, on the NSA's sweeping efforts to 'Know it All'*, WASH. POST (May 12, 2014), http://www.washingtonpost.com/opinions/no-place-to-hide-by-glenn-greenwald-on-the-nsas-sweeping-efforts-to-know-it-all/2014/05/12/dfa45dee-d628-11e3-8a78-8fe50322a72c_story.html.

71. *Press Statement on 30 July 2013*, NAT'L SEC. AGENCY CENTRAL SEC. SERV. (July 31, 2013), http://www.nsa.gov/public_info/press_room/2013/30_July_2013.shtml.

72. Larry Page & David Drummond, *What the . . . ?*, GOOGLE OFFICIAL BLOG (June 7, 2013), http://googleblog.blogspot.com/2013/06/what.html.

73. Mark Zuckerberg, FACEBOOK (June 7, 2013), https://www.facebook.com/zuck/posts/10100828955847631.

74. *Transcript of President Obama's Speech On NSA Reforms*, NPR (Jan. 17, 2014), http://www.npr.org/blogs/itsallpolitics/2014/01/17/263480199/transcript-of-president-obamas-speech-on-nsa-reforms.

75. PRIVACY & CIVIL LIBERTIES OVERSIGHT BD., REPORT ON THE TELEPHONE RECORDS PROGRAM CONDUCTED UNDER SECTION 215 OF THE USA PATRIOT ACT AND ON THE OPERATIONS OF THE FOREIGN INTELLIGENCE SURVEILLANCE COURT 8 (2014) [hereinafter PCLOB REPORT].

76. *Id.*

77. Greenwald, *supra* note 67.

78. Timothy B. Lee, *Everything You Need to Know About the NSA's Phone Records Scandal*, WASH. POST (June 6, 2013), http://www.washingtonpost.com/blogs/wonkblog/wp/2013/06/06/everything-you-need-to-know-about-the-nsa-scandal/.

telephone calls."[79] In describing the highly unusual nature of the 215 Program, *The Guardian* reported that the "[FISA] court orders typically direct the production of records pertaining to a specific named target who is suspected of being an agent of a terrorist group or foreign state, or a finite set of individually named targets."[80] Traditionally, FISA required that:

> [A]ny government agency seeking to use electronic surveillance for foreign intelligence purposes inside the United States had to obtain a warrant from the FISC. For such a warrant to be issued, the government had to show 'probable cause to believe that the target of the electronic surveillance' is an agent of a foreign power.[81]

The collection effort under the 215 Program is not limited by any such requirement.[82]

The disclosures immediately generated a wave of controversy, particularly regarding the unlimited scope of the program. Following the leaks, a bipartisan group of thirteen U.S. Senators tasked two committees to conduct investigations into the data mining programs and to report back on their findings.

1. The Privacy and Civil Liberties Oversight Board Report

Members of the Senate asked the Privacy and Civil Liberties Oversight Board (PCLOB) to investigate the data mining programs and to provide a recommendation "so that the public and the Congress can have a long overdue debate about [the] important privacy concerns."[83] On January 23, 2014, the PCLOB issued its findings and recommendations in a 238-page report.[84] Most notably, it recommended that the government put an end to the telephony metadata program.[85] The PCLOB concluded:

> The Section 215 bulk telephone records program lacks a viable legal foundation under Section 215, implicates constitutional concerns under the First and Fourth Amendments, raises serious threats to privacy and civil liberties as a policy matter, and has

79. *Verizon Forced to Hand Over Telephone Data—Full Court Ruling*, THE GUARDIAN (June 5, 2013), http://www.theguardian.com/world/interactive/2013/jun/06/verizon-telephone-data-court-order (providing a copy of the order signed by a judge of the Foreign Intelligence Surveillance Court).

80. Greenwald, *supra* note 67.

81. PRESIDENT'S REVIEW GROUP, LIBERTY AND SECURITY IN A CHANGING WORLD: REPORT AND RECOMMENDATIONS OF THE PRESIDENT'S REVIEW GROUP ON INTELLIGENCE AND COMMUNICATIONS TECHNOLOGIES 65–66 (2013) [hereinafter REVIEW GROUP REPORT]. "FISA denied the President the previously assumed authority to engage in foreign intelligence surveillance inside the Untied States without judicial supervision." *Id.* at 66.

82. *But see id.* at 98. Before the NSA can access the collected metadata, the FISC requires it to provide facts demonstrating there is a "reasonable, articulable suspicion (RAS) that the selection term to be queried . . . is associated with a specific foreign terrorist organization." *Id.* (internal quotation marks omitted).

83. Letter from Sen. Tom Udall et al. to the Privacy and Civil Liberties Oversight Board (June 12, 2013), *available at* http://www.pclob.gov/library/Letter-Senate_letter_to_PCLOB-Jun2013.pdf.

84. PCLOB REPORT, *supra* note 75.

85. *Id.* at 16.

> shown only limited value. As a result the Board recommends that the government end the program.[86]

Additionally, it recommended that the government purge the data collected under the 215 Program once it had ended.[87]

Echoing sentiments Jeff Jonas expressed in 2006, the PCLOB found that the 215 program showed only minimal value in protecting the country from terrorist threats.[88] It explained:

> Based on the information provided to the Board, including classified briefings and documentation, we have not identified a single instance involving a threat to the United States in which the program made a concrete difference in the outcome of a counterterrorism investigation. Moreover, we are aware of no instance in which the program directly contributed to the discovery of a previously unknown terrorist plot or the disruption of a terrorist attack.[89]

In his testimony before the Senate Committee on the Judiciary seven years prior to this report, Taipale countered the pseudo-technical arguments against data mining and argued that it was a useful tool for counterterrorism.[90] He further argued that data mining should not be burdened with a standard of perfection before it has been given an opportunity to be researched and developed.[91] However, after seven more years of research and development, there is still no credible evidence supporting the effectiveness of this technology.

2. The President's Review Group

On August 12, 2013, President Obama commissioned a high-level group of experts to:

> [A]ssess whether, in light of advancements in communications technologies, the United States employs its technical collection capabilities in a manner that optimally protects our national security and advances our foreign policy while appropriately accounting for other policy considerations, such as the risk of unauthorized disclosure and our need to maintain the public trust.[92]

In a 308-page document released on December 12, 2013, the

---

86. *Id.*

87. *Id.* (noting that the recommendation that the government end the program was supported by a three person majority of the five member Board, although the other recommendations were unanimous).

88. *Id.* at 11.

89. *Id.*

90. *Balancing Privacy and Security: The Privacy Implications of Government Data Mining Programs: Hearing Before the S. Comm. on the Judiciary*, 110th Cong., 1st Sess. 12–15 (2007) (statement of Kim Taipale, Founder & Exec. Dir., Center for Advanced Studs. in Sci. & Tech. Pol'y).

91. *Id.* at 14.

92. White House Office of the Press Sec'y, *Presidential Memorandum—Reviewing Our Global Signals Intelligence Collection and Communications Technologies*, WHITE HOUSE (Aug. 12, 2013), https://www.whitehouse.gov/the-press-office/2013/08/12/presidential-memorandum-reviewing-our-global-signals-intelligence-collec.

President's Review Group made forty-six recommendations, including a determination that the government should not collect and store mass aggregated data for data mining purposes.[93] A major concern of the President's Review Group was the fear that citizens who believe the government is unreasonably scrutinizing them may be "inhibited from contributing fully to the social and cultural life of their communities, and may even alter their purely private and perfectly legal behavior for fear that discovery of intimate details of their lives will be revealed and used against them in some manner."[94] The Reporters Committee for Freedom of the Press also echoed this fear of chilling speech in its *amicus curiae* brief in *Am. Civil Liberties Union v. Clapper*,[95] stating that the mass data collection program threatened to chill reporter-source communications by making confidential "sources increasingly wary of contacting journalists."[96]

The fact that a group of experts specifically commissioned by the President recommended that the program be terminated brings much credibility to the Jonas and Harper arguments. Although the Presidential Review Group's focus was less on the effectiveness on data mining, its general impression was the same—the benefits of data mining are not worth the harms it produces.

B. THE 702 PROGRAM

In June 2013, the *Washington Post* reported that, pursuant to the PRISM[97] surveillance program, the NSA and FBI were tapping into the servers of nine U.S. Internet service providers including Microsoft, Yahoo, Google, Facebook, and Apple to extract "audio and video chats, photographs, e-mails, documents, and connection logs that enable analysts to track foreign targets."[98] In response to the PRISM leaks, James Clapper, Director of National Intelligence, released a fact sheet in which he asserted, "PRISM is not an

93. REVIEW GROUP REPORT, *supra* note 81, at 108.

94. *Id.* at 111 (quoting NAT'L RES. COUNCIL OF THE NAT'L ACAD. OF SCI., PROTECTING INDIVIDUAL PRIVACY IN THE STRUGGLE AGAINST TERRORISTS: A FRAMEWORK FOR PROGRAM ASSESSMENT 2–3 (2008)).

95. 959 F. Supp. 2d 724 (S.D.N.Y. 2013).

96. Brief Amicus Curiae of the Reporters Committee for Freedom of the Press and 18 News Media Organizations in Support of Plaintiffs' Motion for a Preliminary Injunction at 2, Am. Civil Liberties Union v. Clapper, 959 F. Supp. 2d 724 (S.D.N.Y. 2013), *available at* https://www.rcfp.org/sites/default/files/aclu-v-clapper-sdny.pdf.

97. *See* Benjamin Dreyfuss & Emily Dreyfuss, *What is the NSA's PRISM Program? (FAQ)*, CNET (June 7, 2013), http://www.cnet.com/news/what-is-the-nsas-prism-program-faq/ ("PRISM stands for 'Planning Tool for Resource Integration, Synchronization, and Management,' and is a 'data tool' designed to collect and process 'foreign intelligence' that passes through American servers.").

98. Barton Gellman & Laura Poitras, *U.S., British Intelligence Mining Data From Nine U.S. Internet Companies in Broad Secret Program*, WASH. POST (June 7, 2013), http://www.washingtonpost.com/investigations/us-intelligence-mining-data-from-nine-us-internet-companies-in-broad-secret-program/2013/06/06/3a0c0da8-ceb f-11e2-8845-d970ccb04497_story.html.

undisclosed collection or data mining program."[99] However, Director Clapper's assertions cannot be taken at face value given his poor record of candor.[100] Under PRISM, "the government collects the content of electronic communications, including phone calls and emails, where the targets are reasonably believed to be non-U.S. persons located outside the United States."[101] Section 702 of the Foreign Intelligence Surveillance Act (FISA) Amendments of 2008[102] furnishes the basis for the program.[103] Although the NSA may not intend to monitor the content of U.S. citizens' electronic communications, "collections of communications involving U.S. persons may still occur, either where those individuals are in communication with non-U.S. persons or where they are mistakenly believed to be non-U.S. persons."[104] This type of governmental overreach runs counter to the structurally safe and sufficiently safeguarded programs proposed by Taipale.

V. THE DISCLOSURES SUPPORT DATA MINING OPPOSITION

The information disclosed regarding the 215 Program and the 702 Program reveal that the NSA has very little oversight or accountability for its actions, while indiscriminately collecting, storing, and analyzing vast amounts of data from millions of American citizens. The leaks reveal a number of government abuses, a lack of meaningful procedural or substantive safeguards, an absence of strict congressional oversight, and a persistent pattern of ineffectiveness in addressing terrorism.

A. THESE PROGRAMS EXEMPLIFY GOVERNMENT ABUSE

A substantial part of Taipale's argument in support of data mining was that implementing structural and internal safeguards and strong congressional oversight would minimize privacy intrusions.[105] However, the Snowden revelations and related disclosures demonstrate that the NSA has continued to engage in a number of abuses in connection with these data mining programs, despite any structural safeguards or oversight already built into the system.

---

99. DIR. OF NAT'L INTELLIGENCE, FACTS ON THE COLLECTION OF INTELLIGENCE PURSUANT TO SECTION 702 OF THE FOREIGN INTELLIGENCE SURVEILLANCE ACT 1 (2013), *available at* http://www.dni.gov/files/documents/Facts%20on%20the%20Collection%20of%20Intelligence%20Pursuant%20to%20Section%20702.pdf.

100. *See* Andrew Rosenthal, *Making Alberto Gonzales Look Good*, N.Y. TIMES (June 11, 2013), http://takingnote.blogs.nytimes.com/2013/06/11/making-alberto-gonzales-look-good/?_php=true&_type=blogs&_r=0.

101. PCLOB REPORT, *supra* note 75, at 1.

102. 50 U.S.C. § 1881(a) (2012).

103. PCLOB REPORT, *supra* note 75, at 1.

104. *Id.* at 1 n.2.

105. *See supra* Part III.A.

According to published internal audits and classified documents, the NSA has failed to abide by privacy guidelines thousands of times every year since 2008.[106] This resulted in the unauthorized surveillance of American citizens.[107] Additionally, the NSA instructed its analysts on what to disclose in its reports to the Justice Department and the Office of the Director of National Intelligence.[108] Specifically, the classified document revealed instructions to the analysts to "remove details and substitute more generic language in reports to the Justice Department and the Office of the Director of National Intelligence."[109] For example, NSA agents withheld disclosure of errors relating to the unintentional surveillance of Americans.[110] In another incident, the NSA did not make the FISC aware of a newly developed data collection technique until months after its implementation.[111] The FISC subsequently found the technique to be unconstitutional.[112] While the *Washington Post* noted that most of the unreported collection of American citizen's telephone calls and emails were the result of unintended analyst error,[113] other violations were more severe. These included "unauthorized access to intercepted communications, the distribution of protected content and the use of automated systems without built-in safeguards to prevent unlawful surveillance."[114]

These are just a few of the infractions that have been disclosed. It would not be unreasonable to assume there are many more abuses that remain concealed. While mistakes are inevitable in executing any type of complex program, intentionally hiding information and making inaccurate disclosures to the organizations in charge of overseeing the operation is unacceptable.

B. LACK OF PROCEDURAL AND SUBSTANTIVE SAFEGUARDS

The secrecy that shrouded the 215 and 702 programs make it difficult to appropriately scrutinize and evaluate the complexities of the technology and any embedded safeguards. However, the

106. Barton Gellman, *NSA Broke Privacy Rules Thousands of Times Per Year, Audit Finds*, WASH. POST (Aug. 15, 2013), http://www.washingtonpost.com/world/national-security/nsa-broke-privacy-rules-thousands-of-times-per-year-audit-finds/2013/08/15/3310c554-05ca-11e3-a07f-49ddc7417125_print.html.

107. *Id.*

108. Barton Gellman & Matt DeLong, *What to Say, and Not to Say, to Our 'Overseers,'* WASH. POST (Aug. 15, 2013), http://apps.washingtonpost.com/g/page/national/what-to-say-and-not-to-say-to-our-overseers/390/.

109. Gellman, *supra* note 106.

110. *Id.*

111. *Id.*

112. Bill Chappell, *Secret Court: NSA Surveillance Program Was Unconstitutional*, NPR (Aug. 21, 2013), http://www.npr.org/blogs/thetwo-way/2013/08/21/214212847/nsa-culled-tens-of-thousands-of-u-s-emails-yearly-fisa-opinion-says.

113. Gellman, *supra* note 106.

114. *Id.*

Snowden disclosures made it clear that the existing safeguards are not the robust and effective protections Taipale envisioned.

For example, PRISM program analysts would type in search terms ("selectors") to estimate the likelihood that a target is foreign.[115] The selectors produce a result with a confidence level of at least 51 percent.[116] A test that produces a confidence level this low cannot be considered very stringent. The *Washington Post* published confidential training materials, which revealed that the NSA instructed PRISM program analysts to make reports of any data mistakenly collected on U.S. citizens, but added "it's nothing to worry about."[117]

The FISC oversees and reauthorizes the continuance of the 215 Program every ninety days.[118] The FISC, equipped with only five lawyers responsible for reviewing compliance violations, meets *ex parte* in classified proceedings and usually does not publish any of its opinions.[119] Additionally, only government attorneys appear before the court.[120] The Obama administration has tried to assure the public that the FISC "serves as a check on the administration's surveillance power,"[121] but given the high degree of secrecy in its operations, this affirmation has done little to temper down public concerns.[122] Further undermining the President's statement is the fact that since 2008, the FISC has only denied two of the 8,000 surveillance requests submitted by the Obama administration.[123] Recognizing the one-sided nature of the FISC's operations, the PCLOB suggested that Congress enact legislation "enabling the FISC to hear independent views, in addition to the government's views . . . ."[124]

As an inevitable consequence of overseeing an extremely complicated intelligence program, the FISC's oversight is only as strong as the information the NSA provides to it. Judge Reggie B. Walton, the chief judge of the FISC admitted:

> "The FISC is forced to rely upon the accuracy of the information that is provided to the Court . . . [t]he FISC does not have the capacity to investigate issues of noncompliance, and in that respect the FISC is in the same position as any other court when it comes to enforcing [government] compliance with its orders."[125]

---

115. Gellman & Poitras, *supra* note 98.

116. *Id.*

117. *Id.*

118. *Id.*

119. *Id.*

120. *Id.*

121. Jennifer Granick & Christopher Sprigman, *The Secret FISA Court Must Go*, THE DAILY BEAST (July 24, 2013), http://www.thedailybeast.com/articles/2013/07/24/the-secret-fisa-court-must-go.html.

122. *Id.*

123. *Id.*

124. PCLOB REPORT, *supra* note 75, at 17.

125. Carol D. Leonning, *Court: Ability to Police U.S. Spying Program Limited*, WASH.

Judge Walton's admissions directly contradict the repeated assurances by the Obama administration that the FISC exercises robust oversight.[126]

These revelations do little to support Taipale's vision of data mining programs that operate under meaningful safeguards. The NSA's lackadaisical approach to ensuring its analysts fully comply with legal rules and the FISC's rubberstamping of government surveillance requests cannot seriously be considered adequate safeguards for government data mining practices.

C. LACK OF STRICT CONGRESSIONAL OVERSIGHT AND REVIEW

The extent of Congress' knowledge of the 215 Program is still unclear. At least two members of Congress claim that Congress was mislead into authorizing the program.[127] In 2011, Senators Ron Wyden and Mark Udall alleged that officials in the Department of Justice mislead Congress by analogizing "tangible-things" orders with grand jury subpoenas.[128]

The authority of the 215 Program was based on statutory language that allowed the NSA to "obtain business records that are relevant to an ongoing terrorism investigation."[129] The legislation never mentioned that the purpose of the 215 Program was to legalize bulk collection of business records.[130] In 2005, when Congress revised and reauthorized the statute, the public record also did not indicate that it would allow for bulk collection.[131] This seems to support the argument that Congress was kept in the dark about the specific purposes of the 215 Program.

FISA requires the Attorney General to "fully inform" the Senate and House Intelligence and judiciary committees of government activities conducted under Section 215.[132] Additionally, the statute requires the government to provide copies of "all decisions, orders, or opinions of the FISC . . . that include significant construction or interpretation" of FISA provisions to the congressional committees.[133] The PCLOB noted that, "By the time Section 215 was up for renewal in 2011, it was known to *some* members of Congress that the statute was being used to support bulk collection, and the

---

POST (Aug. 15, 2013), http://www.washingtonpost.com/politics/court-ability-to-police-us-spying-program-limited/2013/08/15/4a8c8c44-05cd-11e3-a07f-49ddc7417125_story.html (quoting U.S. District Judge Reggie B. Walton).

126. *Id.*

127. David S. Kris, *On the Bulk Collection of Tangible Things*, 1 LAWFARE RES. PAPER SERIES 1, 54 (2013).

128. *Id.*

129. Lee, *supra* note 78.

130. PCLOB REPORT, *supra* note 75, at 197.

131. *Id.*

132. *Id.* at 202.

133. *Id.* (internal quotation marks omitted).

DOJ provided Congress with a classified description of the NSA's telephone and Internet bulk collection programs."[134] David Kris, former Attorney General for National Security, U.S. Department of Justice, opined that the government clearly "met its legal disclosure obligations to Congress."[135] However, Kris admitted that "[a]lthough the House Intelligence Committee did notify Members of the House of the classified documents and briefings in 2010 (when it was led by Chairman Sylvestre Reyes), it may not have done so in 2011 (when it was led by Chairman Mike Rogers)."[136]

Despite the insistence by NSA defenders, including President Obama, that Congress was aware of the 215 Program, and that "'these programs are subject to congressional oversight and congressional reauthorization and congressional debate,'"[137] "members of Congress, including those in Obama's party, have flatly denied knowing about them."[138] Senator Richard Blumenthal stated, "'The revelations about the magnitude, the scope and scale of these surveillances, the metadata and the invasive actions surveillance of social media Web sites were indeed revelations to me.'"[139]

Other members of Congress were unable to seek out information concerning the NSA programs and FISC court rulings on the legality of the programs. House Representative Morgan Griffith of Virginia sent out multiple requests seeking FISC reviews, critiques of the NSA, and access to the FISA court ruling compelling Verizon to turn over its phone records to the NSA.[140] Representative Griffith wanted this crucial information prior to submitting a vote on "whether to defund the NSA's bulk collection program," but never received a response.[141] House Representative Alan Grayson of Florida similarly asked the House Intelligence Committee for FISA court opinions directing bulk collection and documents concerning the 702 Program.[142] Chairman of the House Intelligence Committee, Representative Mike Rogers, responded that Grayson's request was denied by a Committee "voice vote."[143] Grayson's staff member

134. *Id.* at 197 (emphasis added).

135. Kris, *supra* note 127, at 41. "Apart from briefings for, and documents submitted to, the four designated committees, the record shows that classified briefings were offered to all Members of Congress." *Id.* at 47.

136. *Id.* at 51 n.199.

137. Sam Stein, *NSA Surveillance Program Oversight: White House, Congress Point Fingers at Each Other*, HUFFINGTON POST, http://www.huffingtonpost.com/2013/06/07/nsa-surveillance-program-oversight_n_3405716.html (last updated June 8, 2013) (quoting President Obama).

138. Glenn Greenwald, *Members of Congress Denied Access to Basic Information About NSA*, THE GUARDIAN (Aug. 4, 2013), http://www.theguardian.com/commentisfree/2013/aug/04/congress-nsa-denied-access.

139. *Id.* (quoting Sen. Richard Blumenthal).

140. *Id.*

141. *Id.*

142. *Id.*

143. *Id.*

responded by informing the Chairman that Congressman Grayson "'discussed the committee's decision with Ranking Member [Dutch] Ruppersberger on the floor last night, and he told the Congressman that he was unaware of any committee action on this matter.'"[144] Grayson was puzzled at "how a voice vote denying him access to these documents could have taken place without the knowledge of the ranking member on the Committee, and asked: 'can you please share with us the recorded vote, Member-by-Member?'"[145] The Committee responded, "Thanks for your inquiry. The full Committee attends Business Meetings. At our July 18, 2013 Business Meeting, there were seven Democrat Members and nine Republican Members in attendance. The transcript is classified."[146]

On July 31, 2013, Representative Justin Amash told CNN "I, as a member of Congress, can't get access to the court opinions. I have to beg for access, and I'm denied it if I—make that request."[147] Moreover, in a hearing before Congress Senator Ron Wyden asked James Clapper, Director of National Intelligence, whether the NSA collects "any type of data at all on millions or hundreds of millions of Americans."[148] Clapper answered, "No sir" and then added "not wittingly."[149] Clapper came under intense scrutiny when the public learned that this statement was in fact a lie.[150]

It is unlikely that meaningful congressional oversight and control over government data mining programs can be achieved when the congressional oversight officials are being intentionally deceived by fallacious statements from those in charge of our intelligence programs. The House and Senate Intelligence Committees are primarily responsible for overseeing the NSA, but as Glenn Greenwald notes, "the Intelligence Committees in both houses of Congress are filled with precisely those members who are most slavishly beholden to, completely captured by, the intelligence community over which they supposedly serve as watchdogs. Many receive large sums of money from the defense and intelligence industries."[151]

---

144. *Id.* (quoting a staff member of Rep. Alan Grayson).

145. *Id.*

146. *Id.* (quoting the House Intelligence Comm.).

147. *Transcripts: The Situation Room*, CNN (July 31, 2013), http://edition.cnn.com/TRANSCRIPTS/1307/31/sitroom.02.html.

148. Andrew Rosenthal, *Making Alberto Gonzales Look Good*, N.Y. TIMES (June 11, 2013), http://takingnote.blogs.nytimes.com/2013/06/11/making-alberto-gonzales-look-good/?_php=true&_type=blogs&_r=0.

149. *Id.*

150. Fred Kaplan, *Fire James Clapper*, SLATE (June 11, 2013), http://www.slate.com/articles/news_and_politics/war_stories/2013/06/fire_dni_james_clapper_he_lied_to_congress_about_nsa_surveillance.html.

151. Glenn Greenwald, *Major Opinion Shifts, in the US and Congress, on NSA Surveillance and Privacy*, THE GUARDIAN (July 29, 2013), http://www.theguardian.com/commentisfree/2013/jul/29/poll-nsa-surveillance-privacy-pew.

A critical component of Taipale's argument was that vigorous congressional oversight and review could minimize the inherent risks of a massive government data-mining program.[152] However, it is impossible for Congress to exercise meaningful oversight and review if it is not unequivocally and fully informed about the data mining programs, which it is entrusted with authorizing. Even if Congress was fully aware of the existence and breadth of the 215 Program, the blind rubberstamping of Executive power[153] (as was witnessed in the immediate aftermath of the September 11th attacks) can hardly be considered effective oversight and control.

D. DATA MINING REMAINS INEFFECTIVE AT PREDICTING TERRORIST ACTIVITY

Jonas criticized data mining's effectiveness for governmental counterterrorism endeavors.[154] Though obviously imperfect, one method to test the effectiveness of data mining is to scrutinize any actual terrorist attacks that have occurred since analysts have used these techniques. For instance, predictive data mining did nothing to prevent the horrific bombings carried out by Chechen brothers, Tamerlan and Dzhokhar Tsarnaev at the Boston Marathon in April of 2013.[155] Russian intelligence previously warned the FBI about the eldest Tsarnaev brother in 2011, stating that he "'was a follower of radical Islam and a strong believer,' and 'had changed drastically since 2010.'"[156] Russian officials asked twice in 2011 for information on Tamerlan, but after the FBI's initial investigation did not reveal any evidence of terrorist-like behavior, it ceased its inquiry.[157]

Some congressional officials criticized the FBI for failing to appropriately follow up on these leads. House Representative and head of the House Homeland Security Committee, Michael McCaul, noted "[w]e talked a lot about connecting the dots and stovepipes after 9/11. . . . Here we are 12 years later and it's still not working."[158] House Representative Peter King echoed these statements and noted, "'The Russians did not give the FBI enough information, but the FBI shouldn't expect the Russians to give the information. . . . The FBI should have done more on its own, and in addition to that, shared

152. *See* Taipale, *supra* note 47, at 19.

153. *See* Gellman & Poitras, *supra* note 98.

154. *See supra* Part II.

155. *Boston Marathon Terror Attack Fast Facts*, CNN (Apr. 14, 2014), http://www.cnn.com/2013/06/03/us/boston-marathon-terror-attack-fast-facts/.

156. Bill Hutchinson, *Russia Hid Information from FBI on Boston Marathon Bomber: Report*, N.Y. DAILY NEWS (Apr. 10, 2014), http://www.nydailynews.com/news/national/russia-hid-info-fbi-boston-marathon-bombers-article-1.1751611 (quoting Russian officials).

157. Pete Williams, Erin McClam & Tracy Connor, *What Did the FBI and CIA Know About Bombing Suspects, and When?*, NBC NEWS (Apr. 24, 2013), http://usnews.nbcnews.com/_news/2013/04/24/17899652-what-did-the-fbi-and-cia-know-about-bombing-suspects-and-when?lite.

158. *Id.*

their information with the Boston Police.'"[159] House Representative and author of the Patriot Act, Jim Sensenbrenner, suggested that the NSA's data mining programs might have actually obscured clues that could have prevented the attacks.[160] Representative Sensenbrenner stated, "Sometimes too much information means that what you are looking for is actually a very small needle in a very large hay stack.... You can be drowned in too much information."[161]

Given the breadth of information available to our intelligence agencies concerning the Tsarnaev brothers, it is not unreasonable to wonder why data mining did nothing to anticipate the Boston attacks. This scenario is yet another example of the need for "[b]etter interagency information sharing, investigatory legwork in pursuit of genuine leads, and better training."[162] While the government's central argument in support of data mining is that it prevents terrorism,[163] Nawaf al-Hazmi, Khalid al-Mihdhar, and Tamerlan Tsarnaev are painful reminders that the terrorists who threaten the United States are often already known to its security agencies. The aggregation of vast amounts of information needlessly wastes precious government resources that could be used to more effectively fine-tune the government's ability to follow available leads.

Furthermore, as terrorist activity shifts from being the product of organized, international networks into smaller individualized attacks, it seems reasonable to presume that, over time, any algorithm created to predict a specific pattern of terrorist behavior becomes increasingly unreliable. In response to the Boston attacks, Bruce Riedel, Director of the Intelligence Project at the Brookings Institution, stated that "[w]e are likely to see this as the future face of terrorist threats to the United States."[164] Mitchell Silber, a former intelligence official in the New York Police Department, similarly added, "[t]his more pedestrian, bare-bones terrorism is out there, and it's going to be very difficult to detect."[165]

CONCLUSION

Since June of 2013, Edward Snowden has provided the public

159. *Rep. King: FBI Dropped the Ball Before Boston Bombing*, FOX NEWS (Apr. 10, 2014), http://www.foxnews.com/on-air/americas-newsroom/2014/04/10/rep-king-fbi-dropped-ball-boston-bombing (quoting Rep. Peter King).

160. Lauren Fox, *Sensenbrenner: Data Mining Missed Red Flags in Boston Bombing*, U.S. NEWS (June 10, 2013), http://www.usnews.com/news/articles/2013/06/10/patriot-act-author-sensenbrenner-data-mining-led-to-missed-signals-in-boston-bombing.

161. *Id.*

162. Jonas & Harper, *supra* note 3, at 5.

163. *See supra* note 24.

164. Siobhan Gorman, Gary Fields & Devlin Barrett, *Boston Attack Renews Fears About Homegrown Terrorism*, WALL ST. J. (Apr. 20, 2013), http://online.wsj.com/news/articles/SB10001424127887324763404578433113880189762.

165. *Id.*

with an ample amount of information allowing for a long overdue national debate regarding government data-mining practices. The disclosed 215 and 720 Programs pale in comparison to the effective, structurally safe and dutifully supervised programs Kim Taipale envisioned. If the government cannot effectively structure and oversee data mining to minimize privacy and civil liberties harms, it is better to follow the recommendations of the PCLOB report and discontinue the use of government data mining. The information disclosed by Snowden regarding the interworking of the government's data mining programs supports the concerns of data-mining opponents, and undermines the credibility of the heavily relied upon counterterrorism justification.