BRENNAN
CENTER
FOR JUSTICE

Brennan Center for Justice
At New York University School of Law

Washington, D.C. Office
1730 M Street, N.W.
Suite 413
Washington, D.C. 20036
202.249.7190 Fax 202.223.2683

April 4, 2014

Big Data Study
Office of Science and Technology Policy
Eisenhower Executive Office Building
1650 Pennsylvania Ave. NW.
Washington, DC 20502
Via email: bigdata@ostp.gov

Re: Big Data Request for Information

To whom it may concern:

On March 4, 2014, the Office of Science and Technology Policy (OSTP) issued a request for public comment "on the ways in which big data may impact privacy, the economy, and public policy."[1] This request is part of the OSTP's "comprehensive review of how 'big data' will affect how Americans live and work,"[2] which is expected to be released by the end of April. We are pleased that the White House and OSTP are considering these issues, and we welcome the prospect of a comprehensive review of both the benefits and pitfalls of the use of "big data" by government and private entities.

On March 3, John Podesta told a forum at the Massachusetts Institute of Technology that the White House's review of recently revealed government surveillance programs was occurring on a "somewhat separate track," but that the big

[1] Request for Information, 79 FR 12251 (Mar. 4, 2014), *available at* https://www.federalregister.gov/articles/2014/03/04/2014-04660/government-big-data-request-for-information#h-7.
[2] *Deadline Extension: There is Still Time to Join the Conversation on Big Data and Privacy*, WHITE HOUSE OFFICE OF SCIENCE AND TECHNOLOGY POLICY (Mar. 31, 2014, 7:10 PM), http://www.whitehouse.gov/blog/2014/03/31/deadline-extension-there-still-time-join-conversation-big-data-and-privacy (also indicating that comments are now due April 4, 2014).

data review "may help inform intelligence policy going forward."[3] We thus write to focus on the use of "big data," including data mining, for counterterrorism and other national security purposes.

There is substantial consensus among scientists and national security experts that the use of big data and data mining in the counterterrorism arena is ineffective. A drive to amass information on the scale of "big data" may even be counterproductive in the national security context to the extent it overwhelms intelligence agencies with unhelpful information. In addition, there are heightened risks to privacy and civil liberties associated with the accumulation of deep databases of information for counterterrorism and other national security purposes. In light of the potential impact of OSTP's proposals on intelligence policy, it is critical that OSTP's review acknowledge the limited utility of big data analytics in the counterterrorism context. At the very least, OSTP should affirm that principles for the use of big data that may be beneficial in certain contexts may be inappropriate, underinclusive, or overbroad in other contexts, particularly with respect to national security.

These comments address the following issues, which are responsive to Questions 1, 2, and 4 of the Request for Information:

- Public policy implications of the collection, storage, analysis, and use of big data;
- Uses of big data that raise significant public policy concerns; and
- Distinctions between policy frameworks for use of big data by different sectors – specifically, by counterterrorism-focused and other national security agencies.

Pattern-Based Data Mining Has Limited Value in the Counterterrorism Context

One potential use for "big data" is data mining, or "pattern prediction": analyzing a store of data to tease out patterns connected to certain behaviors, and then looking for matching patterns in other datasets to predict other instances in which those behaviors are likely to occur.[4] When it comes to counterterrorism, however, a

[3] Kate Tummarello, *'Big Data' review to focus on private sector*, THE HILL TECH. BLOG (Mar. 3, 2014, 12:47 PM), http://thehill.com/blogs/hillicon-valley/technology/199710-white-house-big-data-review-to-focus-on-companies.

[4] MARY DEROSA, CTR. FOR STRATEGIC AND INT'L STUDIES, DATA MINING AND DATA ANALYSIS FOR COUNTERTERRORISM 4 (2004), *available at* http://csis.org/files/media/csis/pubs/040301_data_mining_report.pdf; K. A. Taipale, *Data Mining and Domestic Security: Connecting the Dots to Make Sense of Data*, 5 COLUM. SCI. & TECH. L. REV. 1, 22-23 (2003), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=546782 ("Data mining

study commissioned by the Department of Defense concluded that "there is no credible approach that has been documented … to accurately anticipate" terrorist threats.[5] Put another way, there is no known way to effectively identify a potential terrorist by pattern analysis.

Credit card companies are probably the best-known and most successful users of the pattern-matching model. Their success in detecting credit card fraud is due to a number of factors that are almost entirely lacking in the counterterrorism context: the massive volume of credit card transactions provides a rich body of data; a relatively high rate of credit card fraud means the model can be tested and refined; regular and identifiable patterns accompany the fraud (such as testing a card at a gas station to ensure that it works and then immediately purchasing more expensive items); and the cost of a false positive — what happens when the system erroneously concludes that a card has been stolen — is relatively minimal: a call to the owner and, at worst, premature closure of a legitimate account.[6]

By contrast, there have been — statistically speaking — a relatively small number of attempted or successful terrorist attacks, which means that there are no reliable "signatures" to use for pattern modeling.[7] Even if the number of attacks were to rise significantly, it is improbable that they would exhibit enough common characteristics to allow for successful modeling. Indeed, government agencies and experts who have engaged in rigorous empirical studies of "radicalization" have

---

generally identifies patterns or relationships among data items or records that were not previously identified (and are not themselves data items) but that are revealed in the data itself. Thus, data mining extracts information that was *previously unknown*.") (internal citations omitted).

[5] JASON, MITRE CORP., RARE EVENTS § 1.5, at 8 (2009), *available at* http://www.fas.org/irp/agency/dod/jason/rare.pdf. A National Academies of Science report echoed this finding, determining that terrorist identification via data mining (or by "any other known methodology") was "neither feasible as an objective nor desirable as a goal of technology development efforts." NAT'L RESEARCH COUNCIL, PROTECTING INDIVIDUAL PRIVACY IN THE STRUGGLE AGAINST TERRORISTS: A FRAMEWORK FOR PROGRAM ASSESSMENT 3-4 (2008), *available at* http://epic.org/misc/nrc_rept_100708.pdf.

[6] *See* Bruce Schneier, *Why Data Mining Won't Stop Terror*, WIRED (Mar. 9, 2006), http://archive.wired.com/politics/security/commentary/securitymatters/2006/03/70357; *see also* Richard Barrington, *2011 Credit Card Facts and Statistics*, INDEXCREDITCARDS (Jan. 10, 2011), http://www.indexcreditcards.com/finance/creditcardstatistics/2011-report-on-credit-card-usage-facts-statistics.html (noting that as of 2010, there were nearly 1.5 billion credit cards in circulation in the United States, and nearly 55 million credit card transactions every day).

[7] *See* JEFF JONAS & JIM HARPER, CATO INST., POLICY ANALYSIS NO. 584: EFFECTIVE COUNTERTERRORISM AND THE LIMITED ROLE OF PREDICTIVE DATA MINING (2006), *available at* http://www.cato.org/publications/policy-analysis/effective-counterterrorism-limited-role-predictive-data-mining ("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. Unlike consumers' shopping habits and financial fraud, terrorism does not occur with enough frequency to enable the creation of valid predictive models").

concluded that there is no particular pathway to terrorism or a common terrorist profile.[8]

Moreover, a counterterrorism data-mining program would look not just at a single type of data, such as credit card transactions, but "trillions of connections between people and events": merchandise purchases, travel preparations, emails, phone calls, meetings, business arrangements, and more.[9] It is close to impossible to identify coherent patterns that could be used to predict terrorist activity within this welter of data.

In addition, the adverse consequences of a false positive are vastly more damaging to an individual in the counterterrorism context. As security expert Bruce Schneier has suggested, given the almost overwhelming amount of data available, the most accurate imaginable system would still generate on the order of "1 billion false alarms" — that is, emails, meetings, associations, phone calls, and other items falsely tagged as terrorism-related — "for every real terrorist plot it uncovers."[10] A person falsely suspected of involvement in a terrorist scheme will become the target of long-term scrutiny by law enforcement and intelligence agencies. She may be placed on a watchlist or even a no-fly list, restricting her freedom to travel and ensuring that her movements will be monitored by the government. Her family and friends may become targets as well.

And unlike credit card fraud, a conclusion of possible terrorist involvement is more likely to be influenced by activities that may be protected by the First Amendment, such as email or phone communications, political activism, religious involvement, or connections to certain ethnic groups. In short, there is a reason the Cato Institute has warned that data mining for counterterrorism purposes "would waste taxpayer dollars, needlessly infringe on privacy and civil liberties, and misdirect the valuable time and energy of the men and women in the national security community."[11]

---

[8] FAIZA PATEL, BRENNAN CTR. FOR JUSTICE, RETHINKING RADICALIZATION 8 (2011), *available at* http://www.brennancenter.org/sites/default/files/legacy/RethinkingRadicalization.pdf; MARC SAGEMAN, LEADERLESS JIHAD: TERROR NETWORKS IN THE TWENTY-FIRST CENTURY 72 (2008); Clark McCauley & Sophia Moskalenko, *Mechanisms of Political Radicalization: Pathways Toward Terrorism*, 20 TERRORISM & POLITICAL VIOLENCE 415, 418 (2008); RICHARD ENGLISH, TERRORISM: HOW TO RESPOND 52 (2009).

[9] Schneier, *supra* note 6.

[10] *Id*. The Department of Defense JASON study described this problem as the high risk of "false alarm rates … in the face of massive clutter." JASON, *supra* note 5.

[11] JONAS & HARPER, *supra* note 7, at 1.

Collection of Large Amounts of Data May Be Detrimental to National Security

Second, the large-scale collection of information by national security agencies has been repeatedly associated with failures of intelligence by those agencies. Thus, while there may be many circumstances in which the accumulation of large quantities of data is critical – to facilitate credit card fraud detection, for instance, or to enable government agencies to track pandemics and safeguard public health – the wholesale collection of data by national security agencies is likely to undermine rather than enhance effective analysis.

To be sure, "big data" is qualitatively different from "large data," and not enough is publicly known about these agencies' databases to accurately assess whether they qualify as "big data." Regardless of the technical classification, however, the risks associated with far-reaching information collection will surely be magnified if and when big data is deployed in the national security context.

For instance, experts concluded that an overabundance of data contributed significantly to the failure of the intelligence community to intercept the so-called "underwear bomber" — the suicide bomber who nearly brought down a plane to Detroit on Christmas Day 2009. As an official White House review of the attempted attack observed, a significant amount of critical information was available to the intelligence agencies but was "embedded in a large volume of other data."[12] Similarly, the independent investigation of the FBI's role in the shootings by U.S. Army Major Nidal Hasan at Fort Hood concluded that the "crushing volume" of information was one of the factors that hampered accurate analysis prior to the attack.[13]

Officials across a range of agencies have echoed this assessment. As one veteran CIA agent described it, "The problem is that the system is clogged with information," most of which "isn't of interest."[14] A former official in the Department of Homeland Security branch that handled information coming from fusion centers (state- or regional-based centers that collect, analyze and share threat-related

[12] THE WHITE HOUSE, SUMMARY OF THE WHITE HOUSE REVIEW OF THE DECEMBER 25, 2009 ATTEMPTED TERRORIST ATTACK 3 (n.d.), *available at* http://www.whitehouse.gov/sites/default/files/summary_of_wh_review_12-25-09.pdf.
[13] *Lessons from Fort Hood: Improving Our Ability to Connect the Dots: Hearing Before the Subcomm. on Oversight, Investigations, and Mgmt. of the H. Comm. on Homeland Security*, 112th Cong. 2 (2012) (statement of Douglas E. Winter, Deputy Chair, William H. Webster Comm'n on the Fed. Bureau of Investigation, Counterterrorism Intelligence, and the Events at Fort Hood, Texas on Nov. 5, 2009).
[14] David Ignatius, *A Breakdown in CIA Tradecraft*, WASH. POST, Jan. 6, 2010, *available at* http://articles.washingtonpost.com/2010-01-06/opinions/36805490_1_cia-base-cia-veteran-agency-officers.

information among the federal government, the state, and other partners) characterized the problem as "a lot of data clogging the system with no value."[15] Even former Defense Secretary Robert Gates has acknowledged that a decade after 9/11, one would need to ask, "'Okay, we've built tremendous capability, but do we have more than we need?'"[16]

In addition, centralized storehouses of data are particularly vulnerable to intentional as well as inadvertent security breaches, which can have significant consequences both for operational effectiveness and for the victims of the breaches. In mid-2008, for instance, it was revealed that the director of the secretive Strategic Technical Operations Center at the Marine Corps' Camp Pendleton had been feeding reams of classified federal surveillance files to a local terrorism task force.[17] These information-sharing breaches may become more common as more data is aggregated and available through a single access point. In fact, the federal Government Accountability Office has reported a significant increase in data breaches since 2006, with more than a third of the incidents in 2011 involving personally identifiable information, and the compilation of massive databases is likely to further amplify the risk.[18]

Large-Scale Collection of Information Threatens Civil Liberties and Freedom of Expression

Finally, the collection and retention of personal information on a vast scale poses well-recognized risks to privacy, invites abuse, and chills freedom of expression and dissent. As the Senate's Church Committee recognized over four decades ago, "The massive centralization of … information creates a temptation to use it for improper purposes, threatens to 'chill' the exercise of First Amendment rights, and is inimical to the privacy of citizens."[19] Of course, the nature of "big data"

---

[15] STAFF OF THE SUBCOMM. ON INVESTIGATIONS OF THE S. COMM. ON HOMELAND SEC., 112TH CONG., FEDERAL SUPPORT FOR AND INVOLVEMENT IN STATE AND LOCAL FUSION CENTERS 27 (Comm. Print 2012), *available at* http://www.hsgac.senate.gov/subcommittees/investigations/media/investigative-report-criticizes-counterterrorism-reporting-waste-at-state-and-local-intelligence-fusion-centers.

[16] Dana Priest & William Arkin, *Top Secret America: A Hidden World, Growing Beyond Control*, WASH. POST, July 19, 2010, *available at* http://projects.washingtonpost.com/top-secret-america/articles/a-hidden-world-growing-beyond-control/.

[17] *See* Rick Rogers, *Records Detail Security Failure in Base File Theft*, SAN DIEGO UNION-TRIBUNE, May 22, 2008, *available at* http://www.utsandiego.com/uniontrib/20080522/news_1n22theft.html; *Law Enforcement Records Sought in Stolen Pendleton Surveillance Documents*, ACLU OF SAN DIEGO (July 15, 2008), http://www.aclusandiego.org/law-enforcement-records-sought-in-stolen-pendleton-surveillance-documents-massive-number-of-files-stolen-according-to-press-report/.

[18] U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-12-961T, FEDERAL LAW SHOULD BE UPDATED TO ADDRESS CHANGING TECHNOLOGY LANDSCAPE 13 (2012), *available at* http://www.gao.gov/assets/600/593146.pdf (statement of Gregory Wilshusen).

[19] SELECT COMM. TO STUDY GOV'T OPERATIONS WITH RESPECT TO INTELLIGENCE ACTIVITIES, FINAL REPORT, S. REP. NO. 94-755, bk. III, at 778, *available at* http://www.intelligence.senate.gov/churchcommittee.html.

is that it may not be immediately susceptible to targeted searching or retrieval. Nevertheless, attempts to tackle and fix the big data problem will inevitably result in the easier availability of large stores of information, and the risks of abuse will rise as well.

At a minimum, these significant drawbacks mean that any collection of personal information by intelligence and law enforcement agencies must be justified by a significant benefit. There is ample reason to question whether the gathering of information without any basis to suspect wrongdoing is useful to counterterrorism efforts in any manner.[20] Certainly, as noted above, the indiscriminate collection of information for pattern analysis – the premise of a "big data" approach – is not a useful counterterrorism tool. At the same time, the more data that is collected, the greater the potential for abuse, chilling effect, and privacy intrusion.

These risks are not merely theoretical. For instance, recent disclosures about the National Security Agency have revealed both inadvertent and intentional misuses of the agency's broad surveillance authority[21] – problems that were revealed only after repeated assurances that the agency was operating in strict conformance with applicable legal standards.[22] The agency also succeeded in obtaining approval to search its storehouses of data for information about Americans without a warrant, and recently confirmed that it has conducted such warrantless searches.[23] While this

---

[20] *See, e.g.*, EMILY BERMAN, BRENNAN CTR. FOR JUSTICE, DOMESTIC INTELLIGENCE: NEW POWERS, NEW RISKS (2011), *available at* http://www.brennancenter.org/publication/domestic-intelligence-new-powers-new-risks; PATEL, *supra* note 8; RACHEL LEVINSON-WALDMAN, BRENNAN CTR. FOR JUSTICE, WHAT THE GOVERNMENT DOES WITH AMERICANS' DATA (2013), *available at* http://www.brennancenter.org/sites/default/files/publications/What%20Govt%20Does%20with%20Data%20100813.pdf.

[21] *See, e.g.*, Barton Gellman, *NSA Broke Privacy Rules Thousands of Times per Year, Audit Finds*, WASH. POST, Aug. 15, 2013, *available at* http://articles.washingtonpost.com/2013-08-15/world/41431831_1_washington-post-national-security-agency-documents; Adam Gabbatt and agencies, *NSA Analysts 'Wilfully Violated' Surveillance Systems, Agency Admits*, GUARDIAN, Aug. 24, 2013, *available at* http://www.theguardian.com/world/2013/aug/24/nsa-analysts-abused-surveillance-systems; Chris Strohm, *Lawmakers Probe Willful Abuses of Power by NSA Analysts*, BLOOMBERG, Aug. 24, 2013, *available at* http://www.bloomberg.com/news/2013-08-23/nsa-analysts-intentionally-abused-spying-powers-multiple-times.html; Press Release, Office of Sen. Chuck Grassley, Grassley Presses for Details about Intentional Abuse of NSA Authorities (Aug. 28, 2013), *available at* http://www.grassley.senate.gov/news/Article.cfm?customel_dataPageID_1502=46858.

[22] *See, e.g.*, Dan Farber, *President Obama Outlines Four NSA Reform Initiatives*, CNET (Aug. 9, 2013, 1:13 PM) http://news.cnet.com/8301-13578_3-57597814-38/president-obama-outlines-four-nsa-reform-initiatives/ (quoting President Obama as saying that NSA "programs are operating in a way that prevents abuse"); Edward Moyer, *NSA Admits to Some Deliberate Privacy Violations*, CNET (Aug. 23, 2013, 1:08 PM), http://news.cnet.com/8301-13578_3-57599916-38/nsa-admits-to-some-deliberate-privacy-violations/ (noting that earlier that month, then-NSA Director Keith Alexander said that "no one has willfully or knowingly disobeyed the law or tried to invade your civil liberties or privacy").

[23] *See* Spencer Ackerman and James Ball, *NSA performed warrantless searches on Americans' calls and emails – Clapper*, GUARDIAN, Apr. 1, 2014 *available at* http://www.theguardian.com/world/2014/apr/01/nsa-surveillance-loophole-americans-data; Julian

warrantless access was approved by the FISA Court, it highlights the fact that information collected for one purpose may eventually be used for another, making it particularly critical that close attention be paid to the accumulation of information that may be susceptible to abuse down the line.

Additionally, in the years after 9/11, the Federal Bureau of Investigation improperly gathered, recorded, and retained information about individuals' First Amendment-protected activities, often leading to targets' inclusion in federal databases from which it became almost impossible to escape.[24] Fear of inclusion in such databases – and potential scrutiny based on religion, national origin, or other suspect category – may lead to self-censorship or other chilling effects. Indeed, this phenomenon has been documented in the context of one local surveillance program.[25]

Even innocuous information is vulnerable to abuse, often for petty reasons. For instance, a special agent with the U.S. Commerce Department was indicted for and pled guilty to misusing a federal database to track a former girlfriend and her family. The agent had previously threatened to kill the girlfriend or have her and her family deported, and he accessed the database over 150 times in a one-year period to monitor her movements.[26] Recent reports by the FBI's Office of Professional Responsibility depict FBI employees misusing government databases to look up friends working as exotic dancers and conduct searches on celebrities they "thought were hot."[27] Many additional examples have been documented on the state level.[28]

Hattem, *Lawmakers incensed over NSA 'loophole,'* THE HILL TECH. BLOG (Apr. 1, 2014, 5:04 PM), http://thehill.com/blogs/hillicon-valley/technology/202350-lawmakers-incensed-over-nsa-loophole.

[24] OFFICE OF THE INSPECTOR GEN., U.S. DEP'T OF JUSTICE, A REVIEW OF THE FBI'S INVESTIGATIONS OF CERTAIN DOMESTIC ADVOCACY GROUPS 176, 182-84 (2010), *available at* http://www.justice.gov/oig/special/s1009r.pdf (describing investigations of PETA, Greenpeace, and Catholic Worker, among others).

[25] *See, e.g.*, MUSLIM AM. CIVIL LIBERTIES COAL. ET AL., MAPPING MUSLIMS: NYPD SPYING AND ITS IMPACT ON AMERICAN MUSLIMS 29-32, 40-45 (2013), *available at* http://www.law.cuny.edu/academics/clinics/immigration/clear/Mapping-Muslims.pdf (describing the consequences of the New York Police Department's monitoring of the city's Middle Eastern and South Asian population, including alienating Muslims from their mosques and religious communities, hindering social activism and political debate, and prompting student groups on monitored campuses to ban constitutionally-protected political discussions in group spaces).

[26] *United States v. Robinson*, No. 5:07-cr-00596-JF (N.D. Cal. Aug. 25, 2009); Henry K. Lee, *Ex-Agent Indicted in Misuse of Database*, S.F. GATE (Sept. 19, 2007, 4:00 AM), http://www.sfgate.com/bayarea/article/Ex-agent-indicted-in-misuse-of-database-2522021.php.

[27] Scott Zamost & Kyra Phillips, *FBI Misconduct Reveals Sex, Lies and Videotape*, CNN (Jan. 27, 2011, 10:07 AM), http://articles.cnn.com/2011-01-27/us/siu.fbi.internal.documents_1_fbi-employees-occasional-employee-fbi-s-office?_s=PM:US.

[28] *See, e.g.*, Danielle Bell, *Ottawa Cop Demoted for Database Misuse*, OTTAWA SUN, Sept. 26, 2012, *available at* http://www.ottawasun.com/2012/09/26/ottawa-cop-demoted-for-misuse-of-data-bases (senior staff sergeant accessed police databases 169 times over nearly four years for personal reasons); *Former Montreal Detective Used Police Database to Help Mafia*, TORONTO SUN, Nov. 22, 2012, *available at* http://www.torontosun.com/2012/11/22/former-montreal-detective-used-police-database-

Conclusion

In sum, the accumulation of information on a "big data" scale appears to be both ineffective and counterproductive in the national security context. It also poses particular risks to privacy, civil liberties, and freedom of expression and association. While OSTP may not be explicitly addressing the use of big data in the national security context, the principles it articulates are likely to be influential in the national

---

to-help-mafia# (Montreal police detective used a police database to run license plates and pass information to members of an organized crime syndicate); Christine Hauser, *Sergeant Said to Misuse Terror-Watch Database*, N.Y. TIMES, Nov. 21, 2008, at A31, *available at* http://www.nytimes.com/2008/11/21/nyregion/21sergeant.html; *see also* Sewell Chan, *Police Sergeant Guilty of Misusing Terror Database*, N.Y. TIMES, Jan. 14, 2009, http://cityroom.blogs.nytimes.com/2009/01/14/police-sergeant-pleads-guilty-to-misusing-database/ (reporting that a New York City police sergeant illicitly used a state database to retrieve information from a national terrorist watch list for an acquaintance involved in a child-custody case); Lee, *supra* note 26 (special agent with U.S. Commerce Department indicted in 2007 by federal grand jury for misusing a federal database to track a former girlfriend and her family; agent had previously threatened to kill the girlfriend or have her and her family deported, and he accessed database over 150 times in a one-year period to monitor her movements); Jessica Lussenhop, *Is Anne Marie Rasmusson Too Hot to Have a Driver's License?*, CITY PAGES (Feb. 22, 2012), http://www.citypages.com/2012-02-22/news/is-anne-marie-rasmusson-too-hot-to-have-a-driver-s-license/ (over a hundred officers from eighteen agencies across Minnesota accessed the driving records of a female ex-police officer to look at her picture and glean personal details about her, claiming the practice was common place despite state laws requiring all searches to have an investigative purpose); Tom Lyons, *The Odd Loose Ends in Database Misuse*, SARASOTA HERALD-TRIBUNE, Oct. 11, 2012, at BNV1, *available at* http://www.heraldtribune.com/article/20121011/ARCHIVES/210111025 (secretaries at Florida state attorney's office accessed driver and vehicle information database, limited to official police and prosecutorial use, to perform unauthorized searches for information on candidate for state attorney); Allison Manning, *Cops Criticized for 'Misuse' of Databases*, POLICEONE.COM (Apr. 2, 2012), http://www.policeone.com/police-products/software/Data-Information-Sharing-Software/articles/5360910-Cops-criticized-for-misuse-of-databases/ (officials misusing police databases in Ohio included police officer who looked up a woman's personal information and stopped her car more than a dozen times, police officer who "threw items into the front yard of two people he looked up," and three deputies who looked up the "wife of a man with whom one of the deputies had a dispute"); *Former Montgomery Co. Officer Guilty of Police Database Misuse*, DAILY RECORD (Apr. 27, 2011, 4:46 PM), http://thedailyrecord.com/2011/04/27/former-montgomery-co-officer-guilty-of-police-database-misuse/ (former police officer accessed law enforcement databases to assist her drug-dealing fiancé); Levi Pulkkinen, *IRS Worker Caught Snooping on Ex, Others*, SEATTLEPI.COM (Apr. 23, 2012, 9:44 PM), http://www.seattlepi.com/local/article/IRS-worker-caught-snooping-on-ex-others-3498550.php (IRS technician who had previously looked up her ex-husband's tax return pled guilty to misusing her access to IRS databases to review other people's personal information, including a relative with whom she had had a falling out); Aaron Rupar, *In Minneapolis, Private Information Database Abuse 'Endemic,' Attorney Says*, CITY PAGES (Sept. 26, 2012, 12:27 PM), http://blogs.citypages.com/blotter/2012/09/in_minneapolis_private_information_database_abuse_endemic_attorney_says.php (employees in Minneapolis's department of housing charged with accessing driver's license databases for personal purposes; one of the employees also shared his log-in information with other employees); *Utah Launches Investigation of Leak of Immigrants' Information*, CNN (July 22, 2010, 4:12 PM), http://www.cnn.com/2010/US/07/22/utah.attorney.general/index.html?eref=rss_latest&utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+rss%2Fcnn_latest+%28RSS%3A+Most+Recent%29 (employees of state's Department of Workforce Services generated and circulated a list of 1,300 state residents whom they falsely accused of being illegal immigrants).

security and intelligence arena. Accordingly, we urge OSTP to recognize that while basic principles such as transparency, accountability, oversight, and privacy protections are likely to be relevant regardless of the context, other principles or uses of big data may have unintended consequences if imported from one area to another. We also ask that OSTP recognize the pitfalls associated with the utilization of big data in the national security context, to help ensure that those uses are analyzed with particular attention to the concerns articulated above.

Please do not hesitate to contact us if we can be of further assistance as the review proceeds. Liza Goitein may be reached at elizabeth.goitein@nyu.edu or 202-249-7192, and Rachel Levinson-Waldman may be reached at rachel.levinson.waldman@nyu.edu or 202-249-7193.

Sincerely,

Elizabeth Goitein
Co-Director, Liberty and National Security Program

Rachel Levinson-Waldman
Counsel, Liberty and National Security Program