Comments of
Access and the
Electronic Frontier Foundation

General Services Administration
Regulatory Secretariat Division (MVCB)
1800 F Street NW., Washington, DC 20405
ATTN: Ms. Flowers/Notice PCLOB 2014-04, Sunshine Act Meeting

Re: Notice-PCLOB-2014-04 Docket No. 2014-0001

Dear Esteemed Members of the Privacy and Civil Liberties Oversight Board,

Thank you for this opportunity to submit written comments on the Privacy and Civil Liberties Board's ("PCLOB") mid- and long-term agenda. We appreciate the work you have done over the past year, particularly your review of several of the U.S. government's surveillance programs and authorities. We would like to highlight five issues that we think are particularly ripe for your review and oversight over the course of the next year and beyond:

1. The scope of surveillance and potential reforms to the text of Executive Order 12333;
2. The impact of Presidential Policy Directive 28 (PPD-28) on signals intelligence collection;
3. The adequacy of surveillance programs and the problem of using big data analysis as a seminal aspect of these programs;
4. The process and authorities by which government information is classified and the problem of overclassification; and,
5. The U.S. government's adoption and implementation of the International Principles on the Application of Human Rights to Communications Surveillance.

We address each topic in turn.

1. Executive Order 12333

The PCLOB has already announced that Executive Order ("EO") 12333 is already on its agenda:

> The Board will examine EO 12333 and its implications for privacy and civil liberties. EO 12333 establishes the framework that intelligence agencies must follow when conducting intelligence activities. The Board has already begun to work with the Intelligence Community in seeking updates to agency guidelines to reflect advances in technology. The Board is also beginning to explore how to approach and assess agency activities under EO 12333 to ensure the protection of civil liberties.

However, we request that the PCLOB expand the scope of its inquiry. According to a recent *Washington Post* article by John Napier Tye, former section chief for internet freedom in the State Department's Bureau of Democracy, Human Rights and Labor, "[b]ased in part on classified facts...Americans should be even more concerned about the collection and storage of their communications under [EO] 12333 than under Section 215 [of the Patriot Act]." In a follow-

up story, Mr. Tye was interviewed by investigative reporter Charlie Savage for the New York Times. In that interview, Mr. Tye noted, “Regardless of the use rules in place today, [EO 12333] could be abused in the future.”

EO 12333 imposes no meaningful restrictions on the NSA’s ability to collect and retain information on non-U.S. persons or on U.S. persons during a foreign intelligence investigation. Although “[U.S.] persons cannot be individually targeted under [EO] 12333 without a court order,” the content and metadata of their communications can be collected and stored so long as “collection occurs outside the United States in the course of a lawful foreign intelligence investigation.” Since independent reporting indicates that around 90% of NSA-intercepted internet users in foreign investigations are ordinary civilians (including Americans) whose communications have been “incidentally” collected, EO 12333 effectively gives the NSA access to the private communications of Americans without a court order or even the barest suspicion.

Indeed, public reports appear to confirm that collection under EO 12333 includes siphoning off communications, buddy lists, and contact lists from the links between Yahoo! and Google data centers.[1] In an October 2013 Senate Judiciary Committee hearing, Director of National Intelligence General Clapper revealed that this information is used to create social networks of users.[2] At the same hearing, General Alexander noted such chaining is allowed under guidelines called the Special Procedures Governing Communications Metadata Analysis. The intelligence community has released some policy documents regarding EO 12333 and we recommend the board work with relevant parties to declassify and publicly report on any such guidelines and collection activities concerning EO 12333.

Second, there is no oversight of surveillance conducted under EO 12333. As previously stated, the NSA does not need a “warrant or court approval” to collect domestic communications under EO 12333 “and such collection never need be reported to Congress.” According to Senator Dianne Feinstein (D-CA), Chairwoman of the Senate Select Committee on Intelligence, the Committee has been unable “to ‘sufficiently’ oversee activities conducted under 12333.” Nor do companies need to consent or be given notice when their users’ data is collected.

As noted above, Congress is not properly informed of surveillance activities conducted under Executive Order 12333 even though the Intelligence Committees are mandated to “provide vigilant legislative oversight over the intelligence activities of the United States."[3] The PCLOB is empowered to "analyze and review actions the executive branch takes to protect the Nation from terrorism"[4] and we recommend the PCLOB comment on how to increase oversight of EO 12333 activities via Congress, the PCLOB, and/or public reporting. At issue with EO 12333

[1] Zetter, Kim. Report: NSA Is Intercepting Traffic From Yahoo, Google Data Centers. Wired. (October 30, 2013). http://www.wired.com/2013/10/nsa-hacked-yahoo-google-cables/

[2] U.S. Senate, Committee on the Judiciary. *Continued Oversight of the Foreign Intelligence Surveillance Act*. http://www.judiciary.senate.gov/meetings/continued-oversight-of-the-foreign-intelligence-surveillance-act.

[3] Rules of Procedure and S.Res. 400 (1976), as amended. See, http://www.intelligence.senate.gov/pdfs/11214.pdf

[4] 40 U.S.C 2000e.

spying is the inherent tension between the Constitutional power of Congress versus the Presidential power to authorize foreign surveillance. We encourage the PCLOB to explore the issue by commenting on the potential protections an updated surveillance statute could offer to inhibit the overcollection of innocent users' information using EO 12333.[5]

Third, EO 12333's distinction between domestic and foreign communications unfairly impacts users regardless of nationality. As a result, those who are not U.S. Persons are deemed to have little to no protection for the communications by virtue of their perceived location and nationality. The communications may be collected indiscriminately, which is contrary to international law and policy, as described in more detail below.

Alexander Joel, the civil liberties protection officer for the Office of the Director of National Intelligence, recently identified the PCLOB as a key player in the purported "extensive and multi-layered" oversight mechanisms in place to review activities under EO 12333. We hope that PCLOB will answer this call to action and expand its investigation into EO 12333 to explore the entire scope of surveillance conducted thereunder, the entire gamut of policies overseeing such collection, and what reforms could be implemented to narrow this authority.

2. Presidential Policy Directive 28 (PPD-28)

PPD-28 begets more questions than answers about America's signals intelligence policies and we hope to see the PCLOB engage with PPD-28's impact on signals intelligence in its upcoming agenda. PPD-28 notes the intelligence community conducts signals intelligence activities with "care and precision;" however, many of the activities revealed since June 2013 contradict such claims and should be reviewed by the PCLOB. The new layers of oversight introduced with PPD-28 rely exclusively on Executive branch officials. As we've learned from the past, executive-only oversight of signals intelligence activities is not adequate. We encourage the PCLOB to not only investigate remaining questions about the practical effect of PPD-28 on signals intelligence, but also alternative models of oversight that do not rely exclusively on Executive branch officials.

PPD-28 also directs signals intelligence be "as tailored as feasible."[6] Such a statement makes feasibility of tailoring crucial while leaving key terms undefined. In the past, we've seen the Executive branch argue for the mass collection of Internet data and for all calling records of a phone company as "reasonably tailored" in light of the collection. The PCLOB should explore under what trade-offs or time constraints the intelligence community defines feasibility. The board should also explore who decides such feasibility, what process undergirds such a feasibility assessment, and if this information will ever be made public. Practically speaking, will the same official approving signals intelligence also serve as the authority to conduct a

[5] For example updating the main surveillance statute, the Foreign Intelligence Surveillance Act could include definitions of "acquisition," clarifying the definition of "electronic surveillance," and narrowing the targeting procedures.
[6] Presidential Policy Directive No. 28, Section 1d

feasibility assessment? And who will hold the intelligence community publicly accountable if certain purposes for signals intelligence are off limits?

As the PCLOB engages with the feasibility of certain signals intelligence, it will also engage with the intelligence community's redefinition of "bulk collection." Since the Edward Snowden revelations, the intelligence community has redefined "bulk collection" to mean any mass collection of data without the use of a selection term. PPD-28 uses[7] a similar definition; however, the definition lacks nuance. The use of a predicate, such as an identifier, does not necessarily mean a collection activity should be considered "targeted" since massive amounts of unnecessary information may still be collected. We encourage the PCLOB to explore the nuances of "bulk collection" and “targeted collection” by investigating the different gradations of collection techniques. Key questions include how certain constituencies define "bulk collection" and when can signals intelligence collection with a selection term turn into "bulk collection?"

Currently, PPD-28 only initiates annual reviews of "bulk collection," but the scope of the reviews should be expanded to all kinds of mass collection activities outside of the intelligence community's definition of "bulk collection." An analysis on the different types of bulk (or “bulky”) collection by the PCLOB will facilitate its oversight of these annual reviews. The expanded scope may also help to inform its review of the efficacy of certain signals intelligence activities. Another avenue the PCLOB can explore includes reviewing the NSA's ad-hoc Privacy Impact Assessments for its surveillance programs. The PCLOB can study the degree to which such assessments include bulk collection, comment on such assessments, and work to declassify such assessments.

Any investigation into PPD-28's effect on "bulk collection" necessitates a review of PPD-28's effect on "incidental collection." The Executive argues incidental collection of Americans' communications is materially indistinguishable from other collections because incidental collection occurs during a legal acquisition process. Such an approach is incorrect because it uses legal rules made in the non-classified, low-tech world and applies them as design principles for classified, high-tech surveillance. The approach is reminiscent of *Riley v. United States*, where the government argued a search of a cell phone on a person is indistinguishable from the search of other physical property. The government must begin to recognize that new technologies create new types of searches triggering heightened privacy protections. Incidental collection as part of a Title III wiretap is materially distinguishable from incidental collection as part of bulk collection, and the PCLOB should investigate how to further mitigate, or even stop, incidental collection of communications unrelated to the target of the acquisition—especially when such incidentally collected material consists of Americans’ constitutionally protected communications.

[7] PPD-28 notes: "References to signals intelligence collected in 'bulk' mean the authorized collection of large quantities of signals intelligence data which, due to technical or operational considerations, is acquired without the use of discriminants (e.g., specific identifiers, selection terms, etc.)." Footnote 5, Presidential Policy Directive No. 28.

Incidental collection and bulk collection are part of a positive-feedback loop producing overcollection, overretention, and oversharing of innocent users' information. The PCLOB can delve into the positive-feedback loop by investigating footnote 5 in PPD-28, which excludes limitations on bulk collection from "signals intelligence data that is temporarily acquired to facilitate targeted collection." The qualification is ripe for investigation as it intentionally skips over the NSA's initial seizure of users' (non-content and content) data and implicitly assumes that temporary seizure of data comports to Constitutional requirements. Most notably, it skips over key aspects of the NSA's "upstream collection" technique. The PCLOB's review of bulk collection will be especially useful when investigating the exemptions provided by the footnote.

The footnote also touches on key words the PCLOB should explain to the public. What exactly does the Executive define as "acquisition" or "collection" in the context of PPD-28? Are there instances where the "temporarily acquired" intelligence data is retained or shared? And again, is there a threshold where "targeted collection," i.e., collection based on a discriminant, turns into "bulk collection?" Even simple phrases like "temporarily acquired" should be addressed by the PCLOB to lay the foundation for a public dialogue on these topics.

Many questions remain about the implementation of PPD-28 and we hope to see the PCLOB investigate the commitments made in the text. In particular, we hope the PCLOB sheds light on the ambiguous, and sometimes missing, definitions found in PPD-28. Lastly, PPD-28 does not bring disclosure in line with the expectations of democratic governance. The PCLOB is well placed to aid the declassification of reports, policies, and procedures relating to PPD-28. We hope the board will investigate PPD-28's effect on signals intelligence collection with as much vigor as in its previous reports.

3. Adequacy, Not Efficacy, of Signals Intelligence Collection

As the board tries to discern an appropriate methodology to use for the efficacy of intelligence programs, we believe the board should focus on the adequacy of the intelligence community's collection techniques. The Necessary and Proportionate principles define adequacy as being "appropriate to fulfill the specific legitimate aim identified." We've often heard the intelligence community use efficacy as an excuse to conduct mass surveillance. Justifying programs determining whether or not information derived from a program has stopped a terrorist attack falls woefully short of a holistic approach to evaluating the efficacy and more importantly the adequacy of counterterrorism programs.

In reviewing the adequacy of such techniques, the issue of mass surveillance and incidental collection is paramount. In addition, a number of variables should inform the PCLOB's opinion when reviewing these matters, including if the intelligence community already possessed the information before conducting a collection technique, if there were alternative routes to obtaining information other than the mass collection of data, and other variables. In any review, the PCLOB should urge the intelligence community to declassify the methodology or variables used to determine the adequacy of intelligence programs.

While we cannot prescribe an exact methodology to use in these comments, we will touch on problems with an underlying assumption of the government's surveillance programs: the power of big data analytics.

Big data analytics is central to the government's "collect it all" mentality. The PCLOB should be skeptical about claims made regarding the value of big data. For big data analysis to be valid, one must follow rigorous statistical practices. Simply “collecting it all” and then trying to extract useful information from the data by finding correlations is likely to lead to incorrect (and, depending on the particular application, harmful or even dangerous) results.

We look forward to the PCLOB investigating how well the intelligence community is using big data analytics. Many have used the Google Flu Trends example as a failure of big data.[8] Such criticisms—and the public discourse it provokes—are only possible because Google commendably put its program in the public eye. Is the intelligence community hiding comparable failures? And what are the economic and privacy costs of such work? The incredible power big data analytics has to infringe upon privacy requires a public justification of its use by intelligence agencies and the PCLOB should serve as a robust institutionalized advocate for the public's right to know.

While we do not argue that big data analysis cannot sometimes be accurate and effective, there are two major technical problems with big data analysis that we wish to highlight below in light of the PCLOB's effort to work with the intelligence community to establish a methodology for evaluating the adequacy and appropriateness of counterterrorism programs.

The first technical problem is that correlation is not causation; and sometimes, correlation is not even correlation.[9] While big data may allow one to discover new correlations in the underlying data, it doesn't follow that those correlations are meaningful. One should always be suspicious about taking any action based strictly on correlation that was not explicitly being tested for, no matter how convincing it may seem.

Big data analytics inherently involves more information, but this means there will be more false information in any given data set. This is especially true when intelligence agencies are overcollecting innocuous information involving innocent users. The collection described above contributes to a "multiple-comparisons" problem: if you have a large enough data set with enough comparisons, some comparisons that are flukes will appear statistically significant.

[8] Fung, Kaiser, http://blogs.hbr.org/2014/03/google-flu-trends-failure-shows-good-data-big-data/; Lazer, David and Kennedy, Ryan and King, Gary and Vespignani, Alessandro, Google Flu Trends Still Appears Sick: An Evaluation of the 2013-2014 Flu Season (March 13, 2014). Available at SSRN: http://ssrn.com/abstract=2408560.
[9] Marcus, Gary, and Davis, Ernest. Eight (No, Nine!) Problems with Big Data. The New York Times (April 6, 2014).

The second technical problem is the fundamental limitations of machine learning. First and foremost, "getting machine learning to work well can be more of an art than a science."[10] Unfortunately, the complicated nature of most machine-learning algorithms means that they have a variety of design parameters (variables which can be tuned to improve the algorithm's performance) which must be picked a priori by the human analyst, and which can affect the results in unexpected ways. Unfortunately this means that such algorithms can be subject to human error: even though the "machine" is "learning," it is still doing so based on inputs from the person who programmed the algorithm and is subject to that person's biases (subconscious or otherwise).

Many machine-learning techniques are fragile: if their input data is perturbed ever so slightly, the results will change significantly.[11] This is an important point to consider when the results of such machine learning algorithms could affect people's lives in dramatic ways: even the most well-trained, cross-validated, neural network can be presented with data that is just slightly perturbed, and it will output a wildly incorrect result.

The public knows that the intelligence community's reliance on big data analytics is tremendous; however, its use is shrouded in intense secrecy. We know that the intelligence community uses big data to target "burner" cell phones, but the public doesn't know the success rate of such comparisons or how many innocent users' phones were targeted for collection.[12] Many other programs may rely on similar correlative analytics. The adequacy of these algorithms and programs should be known so that the public can have an informed debate on the Executive's collection activities and whether a disproportionate and unnecessary number of users' rights are being violated via "incidental collection."

The role big data analytics plays in surveillance programs behooves the PCLOB to investigate how the intelligence community is using big data, what information it's harvesting, and how it is processing and sharing such information. In doing so, the PCLOB will also have to navigate the tumultuous waters of the Executive's overclassification problem, which we discuss below.

4. Overclassification

The Executive Branch has unilateral authority to classify information based on a finding that the disclosure of the information would damage national security. Information must fall into one or more of the following categories to be eligible for classification:

(a) military plans, weapons systems, or operations;
(b) foreign government information;
(c) intelligence activities (including covert action), intelligence sources

[10] Bradski, Gary, and Kaehler, Adrian. Learning OpenCV: Computer Vision with the OpenCV Library.O'Reilly Media, September 2008
[11] Szegedy, Christian, et al. Intriguing properties of neural networks. http://arxiv.org/abs/1312.6199
[12] Goodin, Dan. How NSA breakthrough may allow tracking of "burner" cell phones. ArsTechnica. (October 14, 2013). http://arstechnica.com/security/2013/10/how-the-nsa-breakthrough-may-allow-tracking-of-burner-cell-phones/

> or methods, or cryptology;
> (d) foreign relations or foreign activities of the United States, including confidential sources;
> (e) scientific, technological, or economic matters relating to the national security;
> (f) United States Government programs for safeguarding nuclear materials or facilities;
> (g) vulnerabilities or capabilities of systems, installations, infrastructures, projects, plans, or protection services relating to the national security; or
> (h) the development, production, or use of weapons of mass destruction.

A recent bill introduced by Rep. Bennie Thompson (D-MS) and Senator Ron Wyden (D-OR) explained that the physical amount of classified information stands somewhere between 7.5 billion and 1 trillion pages. Further, a single intelligence agency produces "...approximately 20 million four-drawer filing cabinets filled with text [of classified information]" every 1.5 years.

Bill Leonard, head of the Information Security Oversight Office and "Classification Czar" during the Bush Administration, noted that "we produce more secrets than we ever produced in the history of mankind and yet we never fundamentally re-assessed our ability to control secrets."

Currently, classification is governed by Executive Order 13526, noting that there should be oversight for overclassification and that there are strict standards for classification. By the terms of the EO, documents should be affirmatively reviewed by a valid classification authority in order to receive classified status. The EO clearly establishes a preference for public access to government documents, and government officials are instructed to err on the side of lower classification.

Government oversight bodies have deemed overclassification to be a serious problem for decades. Yet, the government continues to practice excessive secrecy, with officials publicly estimating that at least half of all information classification is unnecessary.[13]

The intelligence community has shown itself amenable to releasing highly classified material when pressed by the public, the President, or by Congress. The PCLOB should conduct an investigation into procedures for classification and determining the extent to which information is classified by default and not by affirmative review and the true extent of overclassification to the detriment of public involvement in discussions concerning the extent of government surveillance and other government activities.

The PCLOB can also increase transparency by focusing on documents describing government oversight and compliance related to the surveillance programs, and documents providing the

[13] Goitein, Elizabeth and Shapiro, David M. Reducing Overclassification Through Accountability. Brennan Center for Justice at New York University Law School.
http://www.brennancenter.org/sites/default/files/legacy/Justice/LNS/Brennan_Overclassification_Final.pdf

legal rationales under which the programs operate. It can do so by working with the DOJ and NSA to publicly identify and name documents that should be declassified and released to the public and pushing for declassification of many of the materials sent to Congress. This could involve immediate disclosure of past documents, and the recommendation of setting up a system for releasing future Congressional disclosures.

Even without agency agreement, the PCLOB could and should publish an index of such documents, perhaps identifying them by author, date, title, general subject matter, relevant statutory provision, and number of pages. Such information cannot be properly classified, and its release could help the public identify, and prioritize the release of, those documents most critical to the ongoing national debate.

Overall, the board must serve as a guiding light for the larger public. In all the documents the board reviews, it should make its own determination of the necessity of its classification level, and ask the executive to declassify the documents.

5. The International Principles on the Application of Human Rights to Communications Surveillance & the International Covenant on Civil and Political Rights

The PCLOB should analyze the extent to which the United States surveillance practices are truly in line with the country's obligations under international law. The International Principles on the Application of Human Rights to Communications Surveillance ("the Principles") provide an evaluative framework for this analysis. The Principles, which have been endorsed by more than 400 civil society organizations, and referenced in debates on legislative reform in several countries and regions (including the President's Review Group on Intelligence and Communications Technologies), set the benchmark for assessing whether surveillance laws and practices respect human rights. The Principles draw on international law and jurisprudence.

Earlier this year, the U.S. Government announced that adoption of a subset of the Principles, known as the U.S. Framework.[14] Namely, the U.S. government identified six privacy principles to govern surveillance practices: Rule of Law, Legitimate Purpose, Non-Arbitrariness, Competent Authority, Oversight, Transparency, and Democratic Accountability. These six principles mirror several of the International Principles on the Application of Human Rights to Communications Surveillance.

Despite the fact that the Principles were intended as a holistic and self-referential document, the White House stopped short of adopting all thirteen items. However, the United States surveillance practices, such as those under EO 12333, contravene the government's own privacy guidelines for signals intelligence. As Access has explained, "Despite the public commitment to these principles, U.S. communication surveillance programs do not respect the rights laid out in the U.S. Framework."

[14] See, http://www.humanrights.gov/2014/03/04/state-department-on-internet-freedom-at-rightscon/

Another important international instrument is the International Covenant on Civil and Political Rights ("ICCPR"), to which the United States is a member. Article 17 of the ICCPR holds that "no one shall be subjected to arbitrary interference with his privacy." The U.S. Framework recognizes this Article in its "Non-Arbitrariness" Principle. However, the United States believes that the terms of the ICCPR, and in particular Article 17, do not apply extra-territorially. However, this view was rendered untenable by recent reports of the UN Human Rights Committee and the UN High Commissioner for Human Rights, which clarified the ICCPR's application to apply to all citizens "regardless of the nationality or location of individuals whose communications are under direct surveillance."

The PCLOB should conduct a detailed investigation into the U.S surveillance practices as they relate to the U.S. Framework as well as international law (as evinced by the Principles), including the ICCPR, and with particular respect to the High Commissioner's report and statements on extraterritoriality. To the extent that U.S. practice does not conform with these laws and policies, the PCLOB should recommend reforms in order to bring the U.S. into compliance.

Conclusion

As we continue to learn more about the scope and scale of secret surveillance in the United States, and around the world, the PCLOB continues to become increasingly important. We thank you again for this chance to provide feedback, and would happily answer any further questions on the above issues. You can reach us at aime@accessnow.org and jaycox@eff.org.

Thank you,

Amie Stepanovich
Senior Policy Counsel
Access

Mark Jaycox
Legislative Analyst
Electronic Frontier Foundation