Association for Computing Machinery (ACM)
US Public Policy Council of ACM (USACM)

usacm.acm.org
facebook.com/usacm
twitter.com/usacm

October 4, 2013

Review Group on Intelligence and Communications Technologies
Office of the Director of National Intelligence
Washington, D.C. 20511

Dear Members of the Review Group:

We are writing to submit comments in connection with the September 4, 2013 request posted online by the Office of the Director of National Intelligence. We are the U.S. Public Policy Council of ACM (USACM): a group of technical experts representing ACM (the Association for Computing Machinery) a major technical and professional society with over 110,000 members involved in all aspects of computing and information technology. ACM's members have decades of experience in the development, implementation and use of computing systems, with consideration for ensuring the privacy and security of those systems. USACM stands ready to provide further assistance beyond these comments, should you request it.

We offer the following comments intended to address technical questions and assumptions connected to how the nation can "…employ its technical collection capabilities in a manner that optimally protects our national security and advances our foreign policy while respecting our commitment to privacy and civil liberties, recognizing our need to maintain the public trust, and reducing the risk of unauthorized disclosure." (We urge you to include "unnecessary disclosure" in this charge.) We are specifically *not* addressing policy issues concerning the scope and legal basis for these capabilities. Given the classified nature of these programs, and the incomplete information publicly available, our comments are necessarily high-level in nature. Should you wish further explanation or have follow-up questions, please do not hesitate to contact our Public Policy Office through the contact information at the end of this letter; several of our members have the highest security clearances.

Computing can allow for the collection, analysis and exposure of information in volumes and ways that were not previously possible (or imaginable). But computing alone cannot provide expert judgment on the meaning, utility, correctness, and uses of that information. Without taking the proper steps to implement and execute policies for controlling the collection, security, integrity, audit, accuracy, and use of information, application of computing tools may lead to policy problems that are more difficult to solve than the problems those tools were intended to address.

As even the strongest technical controls may be programmed incorrectly, fail, or be subverted, human oversight will almost certainly be required to meet appropriate policy goals. So, while we support the implementation of strong technical controls, decision makers should not rely solely on such controls to implement security and privacy policies.

As a component of your own study, we recommend conducting an independent systems-

ACM US Public Policy Council (USACM)
1828 L Street NW, Suite 800
Washington, DC 20036

Tel: +1-212-626-0541
Fax: +1-202-667-1066

acmpo@acm.org
usacm.acm.org

engineering analysis of the data collection and analysis structure(s) developed - intentionally and unintentionally - by the programs you are examining. This would include the specific technical requirements of these systems, the operational assumptions underpinning the programs, and the practices involved; policy is often embedded within, and limited by, technical considerations. Limited access to information and its classified nature make it impossible for complete outsiders to effectively conduct such an analysis. At a minimum, such an analysis might be conducted by independent agencies within the government, and reviewed by cleared legal and technical personnel outside of the agencies involved.

With such an analysis in hand it should be simpler to observe and consider the complex trade-offs involved among national security interests, technical capabilities, privacy and civil liberties principles, and other factors of concern. Critical assessment of the programs, including whether or not less invasive methods and practices could achieve the same results, can help address the concerns of many parties. *We doubt that a meaningful policy analysis can be conducted without a corresponding technical analysis.*

In the comments that follow, our assumption is that the ubiquity and importance of our information infrastructure is such that intelligence or national security access to that infrastructure must involve appropriately balancing multiple interests that include

- the ability to detect and investigate threats to the country;
- the need to secure our infrastructure against attackers;
- the right to individual privacy; and
- the need for continued technical innovation.

There should be a full and open public dialog about the tradeoffs involved across these interests, with the risks and benefits carefully and fully explored. While it may not always be possible to have an unclassified dialog, ensuring some kind of outside review of the tradeoffs will help make sure risks and benefits are carefully considered.

We also are basing our comments on our best understanding of the nature and goals of the various systems under examination; we do not have particular insight into their scope or nature beyond what has been presented in the press. Insofar as our understanding is (necessarily) incomplete, we trust you will adjust our comments accordingly.

Limitations of Computing

Before discussing technical issues specific to the surveillance programs being considered by the Board, it is worth taking time to note some limitations of computing and computing systems. Policymakers often presume that computer systems can do more to achieve a desired policy

ACM US Public Policy Council (USACM)
1828 L Street NW, Suite 800
Washington, DC 20036

Tel: +1-212-626-0541
Fax: +1-202-667-1066

acmpo@acm.org
usacm.acm.org

objective than is actually the case. Part of USACM's work is dedicated to ensuring that policymakers understand what computing can and cannot do in support of policy objectives.

As with any other technology, computers are fallible. They can break down or produce erroneous results – especially if they are not maintained properly. Software has bugs, and computer systems have many threats to contend with both from inside and outside of the organizations where they are used. These threats include, but are not limited to: viruses, botnets, other malware, poorly patched software, faulty hardware, improperly implemented security measures, and users who fail to maintain their systems. One of the more pernicious issues is that of an "insider threat" – someone within an organization who takes advantage of lax security or evades controls and other procedures to release a great deal of protected information. Alternatively, someone with great authority may be able to insist that access controls be overridden and audit results be ignored or made unavailable for review without the proper policies and technical controls in place to prevent this. All of these problems — outsider attacks, insider abuses, and administrative excesses — have been reported for even highly-classified, controlled systems.

Even when a set of desired controls has been determined, implementing those controls may be difficult in every place where the data resides. In a few cases (reportedly for some NSA surveillance data), highly sensitive data is kept segregated. But in other cases, the data flows under analysis are mixed and diffused until they are no longer separable. In addition, various collected items are stored in data formats used by specialized systems in a number of different places, which results in poor data coherence. Few organizations are well-positioned to track all of their data flows. Even if they can, it is difficult to transform high-level policies on data controls into appropriate rules that are executable in each specialized system.

A database system cannot, by itself, determine if collected information is accurate or reliable. Bad or incomplete data will lead to flawed searches and results, which will contribute to flawed decisions.[1] For instance, an electronic employment eligibility verification system will be able to confirm that people with certain credentials can legally work in the United States. But if the databases checked by this system contain bad or incomplete data, the system's ability to verify employment eligibility suffers.[2]

[1] We noted this in 2003, in a letter to Congressional leaders, when analyzing the Total Information Awareness initiative — which has some uncanny similarities to the systems you are now examining. Congress shut down that program, and we suggest that you may wish to examine its profile and objections to it to see if there are additional lessons to be learned.

[2] A few of our members have provided Congressional testimony on this topic in recent years, and we would be happy to share this if you would like to see it.

ACM US Public Policy Council (USACM)
1828 L Street NW, Suite 800
Washington, DC 20036

Tel: +1-212-626-0541
Fax: +1-202-667-1066

acmpo@acm.org
usacm.acm.org

Technical Issues

Identification of persons subject to surveillance

The programs under consideration putatively require that collected information can only be returned for analysis for specific persons in certain circumstances.[3] The ability to check these circumstances may depend on several distinct factors such as successfully locating a communications device, linking that device to a particular individual or individuals, and knowledge about the individual's status. Each has considerable uncertainty which technology will never completely remove.

Technologies exist that can determine the location of particular communications devices, but there are also means of intentionally circumventing those technologies or otherwise concealing the location of a particular individual and/or device for various reasons, whether they be for privacy-enhancing or malicious purposes.

Network analysis, which often relies on the use of metadata, can be very effective in finding active nodes. As connections can easily be erroneous, caution is necessary in interpreting the results. While the use of social network analysis to identify a moderate number of individuals for the next stage of an investigation seems reasonable, one must avoid declaring massive numbers of people "suspects" whose data may be freely collected or used. Part of the search criteria should include an assessment of the minimum degree of separation needed to conduct an effective search to avoid false positives, such as those caused by spoofed IP addresses or botnet activity, or by errors in processing.

Policies that assume collected data is perfect ignore critical issues that affect the desired outcomes of searching that data. Policies need to deal explicitly with uncertainty, expressing confidence in particular assertions. Implementation of these policies will address specific technologies, and those measures need to be reviewed by privacy officers.[4, 5]

[3] Meaning that the search software will examine other records, but only return the information that fits within the authorized parameters.

[4] Assertions might, for example, be about location, or who owns a device, whether that owner is a non-US person, or whether the owner is linked to terrorism. Each of these assertions is probabilistic. Debate about the appropriate levels of protection for U.S. persons and non-U.S. persons is outside our technical expertise, and is not addressed here.

[5] The strength of the links between people is another assertion, and not all links are equal in meaning. For instance, both parties communicating with the same suspected terrorist is a much different link than both parties communicating with a major airline.

ACM US Public Policy Council (USACM)
1828 L Street NW, Suite 800
Washington, DC 20036

Tel: +1-212-626-0541
Fax: +1-202-667-1066

acmpo@acm.org
usacm.acm.org

Data minimization

The breadth of data collection involved in these programs raises concerns about the ability to effectively secure this data and maintain the privacy of persons whose information is collected. If the NSA is to collect and store data within its own infrastructure, we would encourage it to follow generally accepted fair information practices of minimization. We distinguish this notion of minimization from how minimization may be used in the signals intelligence context by focusing on the need for information collected and used. This includes both *data minimization* – collecting only the information needed for a particular purpose, and *use limitation* – avoiding the use of information for additional purposes not connected to the ones for which data was initially collected.[6] This minimization would reduce the amount of information collected and the associated demands on personnel and infrastructure to effectively manage its storage, analysis and access. Careful consideration should also be given to expiration of data: data becomes less reliable as it ages, while continuing to tempt parties to use it for other purposes.

A massive dataset is a large target, especially if the dataset is stored or accessible via network-connected systems. To the extent practical, data searched should be kept where it originated (e.g., at the communication providers' data centers, if it is kept at all), and both searches and analyses distributed across those locations. This has already been shown to be feasible for certain statistical computations in health care. But such distributed work is not easy. Analyzing social networks for paths that span multiple private systems might include parties who are not currently under investigation. If the data is not centralized, the query load and latency to find only targets of investigation will increase, and might make such analyses infeasible. Investigation is needed to determine how to make such distributed work practical.

Conducting searches and analyses across different storage locations requires that all sources respond promptly to queries, that queries are adapted so they do not reveal the query-submitter's interest to data holders or those eavesdropping on them, that analyses are amenable to being done in pieces and assembled later, and that rich connections between classified and critical commercial systems are tightly maintained without reducing their security. That this problem is difficult does not mean it is insoluble.

Decentralizing this information has its advantages, in protecting from large scale misuse – but also inhibits some legitimate uses.

For getting information on a particular person, there are probably dozens of sources (at least) and no standard form for querying. It will be laborious but not impossible to go to each of them to inquire about a particular suspect. It seems unwise to ask the private sector to build in easy trap doors for such inquires. As others have repeatedly pointed out, this inserts vulnerabilities outsiders might exploit.

[6] This is sometimes referred to as "mission creep," and has allegedly occurred with data collected by the NSA being used by the DEA and other agencies.

ACM US Public Policy Council (USACM)
1828 L Street NW, Suite 800
Washington, DC 20036

Tel: +1-212-626-0541
Fax: +1-202-667-1066

acmpo@acm.org
usacm.acm.org

Many systems are built for batch processing, not for handling queries about individuals. To insist that each source respond to queries on individuals, system owners may be required to invest in additional capabilities. We suggest that any arguments about cost efficiencies include meaningful costs of privacy and compliance assurance.

Data access controls

Members of the Privacy and Civil Liberties Oversight Board expressed an interest during the Board's July 2013 workshop in finding technical means to limit the uses of collected data. We agree with Dr. Steve Bellovin who indicated during the workshop that limiting access is primarily a policy question rather than a problem in search of a technical solution. Policies can be implemented by technical means, but choosing particular technologies cannot guarantee limitations on use of collected information.

Because the systems under examination collect a great deal of data that may be outside the scope of specific, authorized searches, the temptation for decision makers to expand the scope of what is authorized is perhaps greater than in other circumstances. There will also be the temptation, if only for expediency, for operational personnel to take advantage of gaps in enforcement. Only consistent implementation of access controls and logging, with regular audits and timely, forceful corrective actions can ensure that inappropriate use (as well as outsider access) of collected data is minimized (not eliminated).

If data of a particular sort is segregated, one can impose new policies. But often, the data gets sliced and diced by category, and copies get forwarded into a spaghetti-tangle of systems. With the best intentions in the world, it is difficult to impose a policy on each kind of data, or to audit compliance over the tangle of systems. It is certainly harder to have a comprehensive understanding of these systems, making violations of procedure easier to miss. The underlying message is: If you want to control your systems, you need to impose some simplicity, e.g., uniform tagging. (Document tagging is cheap, but if small data records have tags, the overhead can be substantial.)

We have considerable empirical evidence of abuses of authorities and capabilities related to government access to communications and information, most recently in the context of National Security Letters. Improved oversight, while always desirable, cannot substitute for more substantive policy and technical controls on any intercept functionality. This includes audit trails that cannot be modified or erased and that are subject to routine (including, possibly, automated) review and analysis. Appropriate controls play an important role in enabling effective oversight.

Some pundits have suggested that anonymizing or otherwise obscuring some collected information can address privacy concerns. This approach is problematic. Advances in data collection and analysis make re-identifying supposedly anonymized information easier. In the same way that security threats and countermeasures have a short life before they are countered,

ACM US Public Policy Council (USACM)
1828 L Street NW, Suite 800
Washington, DC 20036

Tel: +1-212-626-0541
Fax: +1-202-667-1066

acmpo@acm.org
usacm.acm.org

anonymizing techniques may be effective for only a short period of time. Much in the same way that metadata can reveal much about a particular individual and his or her habits, data and behavioral patterns connected to an anonymized record can reveal much more than could reasonably be expected for data that is supposed to be 'anonymous.'

Mandated intercept capabilities can introduce cybersecurity vulnerabilities

Our information infrastructure is under constant threat from outside attackers seeking to exploit security vulnerabilities for a variety of purposes, including inflicting physical damage, identity theft, criminal fraud, and intelligence collection. These threats are growing in sophistication and magnitude. Requirements that new or existing technologies that are part of this infrastructure include so-called "backdoors"—special access mechanisms outside of normal access controls—for intelligence gathering will create *new* security vulnerabilities that outside attackers will exploit.

Examples of these unintended vulnerabilities have been described in the media and the academic literature. Systems with built-in intercept functionality have been subverted, and systems were infiltrated by parties outside of those for whom the designed backdoors were intended. A prime example of the unintended consequences of such backdoors is the so-called "Athens Affair." In 2005, it was discovered that built-in intercept functionality of four switching computers of the Athens-based Vodafone-Panafon cell phone network had been subverted some months before.

Indeed, the federal government is in the process of committing massive resources to efforts to enhance the cyber-security of both federal systems and information infrastructure controlled by the private sector. Wide-scale introduction of backdoors represents a fundamental tension, a risk-risk tradeoff, in which security risk associated with the unmonitored use of information infrastructure will be traded for additional security risk associated with cyber attacks. This proposed tradeoff should be subjected to in-depth, systematic analysis, including consideration of the abilities of some actors to defeat intercept functionality (e.g., through client-based encryption) and the abilities of others (e.g., sophisticated criminals or spies) to exploit that same functionality.

*Design demands can limit innovation***.**

For many years, law enforcement's ability to tap into our information infrastructure was essentially opportunistic. The infrastructure was not designed to be secure and law enforcement could obtain legally authorized access in a technically straightforward manner. However, the continuous evolution of commercial communications systems caused this to change, resulting in legal constraints for systems design and implementation as a result of the Communications Assistance for Law Enforcement Act (CALEA). Insufficiently considered during the debate over CALEA was the adverse effect such legal intrusion into communications and computer

ACM US Public Policy Council (USACM)
1828 L Street NW, Suite 800
Washington, DC 20036

Tel: +1-212-626-0541
Fax: +1-202-667-1066

acmpo@acm.org
usacm.acm.org

technology might have on technological and commercial advancement.

Further intervention into the design of information infrastructure may have the unintended side effect of limiting technical innovation, including the development of new and more effective security mechanisms. Regardless of the specifics, mandated intercept capabilities could very well compromise functionality, performance, privacy, and/or security, resulting in inferior commercial information systems compared to systems developed in jurisdictions without such legal constraints. If such intervention must take place, proposed access mechanisms should undergo broad, thorough, and transparent technical review by appropriate bodies.

Technical assistance for FISA courts

Many computing professionals from outside of government have experience and current clearances in classified environments, and would be able to assist the FISA courts in their reviews and deliberations by answering technical questions about the programs and cases the courts review. The courts may wish to devise an ongoing method of obtaining needed technical advice from computing professionals who possess the necessary clearances (and who are independent of those agencies) to appear before the court. This takes on more urgency, as the FISA court does not benefit from the usual adversarial discovery and argument present in other courts.

Cryptography and Communications Infrastructure

We must express our deep disapproval of activities alleged in the press disclosures that national surveillance policy undermined standards and standards bodies, introduced "backdoors" into commercial products, and compromised commercial communications infrastructure. Although we acknowledge the important and difficult mission of US intelligence agencies, we are dismayed at the vast extent of what is alleged to have taken place. The Internet, computing, science, and our profession are all are based on a foundation of trust in products, services, publication, and in conduct; the activities alleged in the disclosures represent a large-scale subversion and corruption that undercuts that trust. The alleged noxious activities are directly contrary to ideals we hold dear as U.S. citizens, are contrary to principles expressed by the U.S. Department of State for use of the Internet (e.g., the principles of the Freedom Online Coalition), appear to be in conflict with the UN Charter of Human Rights (especially Article 12), and go against "best practices" as embodied in the Codes of Conduct of many professional organizations (e.g., our own ACM Code of Conduct).

Competitiveness

For years to come, the results of actions alleged in the press may hurt our economy, our national reputation, and the ability of computing professionals in the U.S. to interact with our colleagues internationally. Continued concerns over the extent of these compromising actions raises the possibility that companies and consumers will seek out products and services that are not produced by companies that are seen as complicit in those actions. Data, and the business

ACM US Public Policy Council (USACM)
1828 L Street NW, Suite 800
Washington, DC 20036

Tel: +1-212-626-0541
Fax: +1-202-667-1066

acmpo@acm.org
usacm.acm.org

activity associated with it, will migrate from U.S. locations and businesses, boosting the economic activity of other countries at the expense of the U.S.

Undercutting the trust placed in American computing companies, products, services and personnel risks much, including the ability to access necessary intelligence information. A deterioration of this global trust in computing could lead to a loss of interoperability as consumers opt to purchase only those goods and services that they could trust not to be exploited by foreign governments. It also undercuts the nation's initiatives that utilize computing technology to support freedom abroad and encourage civil society initiatives around the world.

We urge strong oversight mechanisms and rigid safeguards to prevent the possibility of any activities that could further damage the trust relationships we have noted above.

We appreciate the opportunity to provide comment to the Review Group on Intelligence and Communications Technologies. Should you have any questions on the above, or need additional information, please contact our Public Policy Office at acmpo@hq.acm.org, or at 212-626-0541.

For the USACM,

Eugene H. Spafford, Ph.D.
Chair, U.S. Public Policy Council
Association for Computing Machinery

ACM US Public Policy Council (USACM)
1828 L Street NW, Suite 800
Washington, DC 20036

Tel: +1-212-626-0541
Fax: +1-202-667-1066

acmpo@acm.org
usacm.acm.org