NORTH CAROLINA JOURNAL OF LAW & TECHNOLOGY
11 N.C. J.L. & TECH. ON. 259 (2010)

"APPS.GOV": ASSESSING PRIVACY IN THE CLOUD COMPUTING ERA

Shahid Khan [*]

An information technology project "Apps.Gov" was announced by the Federal Government on September 15th, 2009. Google, Inc. is the principal contractor and is building a special data center facility for the project. In December, Google discovered a breach at another one of it data centers, which led to public concerns about cyber security and even cyber war. Privacy concerns, however, have received comparatively less attention. Given the extraordinary degree to which government data may be aggregated and placed within a technology infrastructure that is still evolving, an outdated system of privacy laws may further widen the gap between technology and privacy policy. It is especially unclear whether agency privacy officers will be able to police privacy and interpret the Privacy Act of 1974 in the context of cloud computing. This Recent Development looks at privacy issues arising from Apps.Gov and the use of privacy impact assessments as a way of dealing with them.

I. INTRODUCTION

In December 2009, Google, Inc. discovered a sophisticated attack on some of its computers by hackers.[1] Because it is perhaps the foremost purveyor of Internet search technologies, Google's vulnerability made instant headlines and even prompted a high-level diplomatic exchange between the U.S. and China, the country from where the attacks are believed to have originated.[2] Consider

[*] J.D. Candidate, University of North Carolina School of Law, 2011

[1] *Google "May Pull Out of China After Gmail Cyberattack,"* BBC NEWS, Jan. 13, 2010 http://news.bbc.co.uk/2/hi/business/8455712.stm (last visited Apr. 8, 2010) (on file with the North Carolina Journal of Law & Technology).

[2] Google released information about the attack on its official corporate blog in January. Posting of David Drummond to The Official Google Blog, http://googleblog.blogspot.com/2010/01/new-approach-to-china.html (Jan. 12,

a similar breach of computer security, except this time it is at a facility run by a private contractor on behalf of the federal government and houses data belonging to various agencies. Such a facility may contain hundreds of thousands, perhaps millions, of records. As security experts attempt to trace the attack's origins and pinpoint the vulnerability in the computer systems that were exploited, numerous individuals are put on alert for identity theft and many others report hijacked credit cards and intrusive activity on their credit reports.

On September 15th, 2009, Vivek Kundra, the Federal Chief Information Officer ("CIO"), announced the launch of Apps.Gov,[3] an Internet portal where federal agencies can purchase the use of software services much like ones found on Google's website.[4] The decision was partly motivated by the desire to use tax dollars more efficiently[5] and partly motivated by the desire to leverage new

---

2010, 15:00) (on file with the North Carolina Journal of Law & Technology). Within a short time the news was widely circulated on the Internet. A few days later, on January 21, 2009, Secretary of State Hillary Clinton, spoke of the attacks in the context of freedom of access to information. Hillary Clinton, Secretary of State, Remarks on Internet Freedom (Jan. 21, 2010), http://www.state.gov/secretary/rm/2010/01/135519.htm (last visited Feb. 21, 2010) (on file with the North Carolina Journal of Law & Technology). In response, China recently decried what it saw as American "information imperialism." Clifford Coonan, *The Google War: China Calls US an "Informational Imperialist,"* THE INDEPENDENT, Jan. 23, 2010, http://www.independent.co.uk/news/ world/asia/the-google-war-china-calls-us-an-information-imperialist-1876409. html (on file with the North Carolina Journal of Law & Technology).

[3] General Services Administration, Apps.Gov, https://apps.gov/cloud/ advantage/main/start_page.do (last visited Apr. 12, 2010) (on file with the North Carolina Journal of Law & Technology).

[4] Google offers E-mail ("Gmail") and document processing ("Google Docs") services, among others.

[5] Total spending on information technology for fiscal year 2009 amounted to $74.2 billion. Vivek Kundra stated cost cutting as one of the goals of implementing Apps.gov. Posting of Vivek Kundra to The White House Blog, http://www.whitehouse.gov/blog/streaming-at-100-in-the-cloud/ (Sept. 15, 2009, 12:09 EST) (on file with the North Carolina Journal of Law & Technology). The administration is also under pressure to reduce the federal deficit which, for fiscal year 2009, was a record 1.44 trillion dollars. Posting of David Jackson to

technology to improve and enhance "electronic government" (E-Government).[6] Complementing the push toward E-Government is a greater emphasis on openness and collaboration in agency decision-making by the current administration. On December 9, 2009, the Obama administration issued a much-awaited Open Government directive to federal agencies requiring them to take steps to increase transparency.[7]

Not much, however, has been said about the government's privacy policies and its efforts to ensure compliance with privacy laws. What little discussion[8] of privacy has taken place has been couched in the discourse of security, in particular cyber security.[9]

The Oval: Tracking the Obama Presidency, http://content.usatoday.com/communities/theoval/post/2009/10/620000005/1 (Oct. 16, 2009, 15:54 EST) (on file with the North Carolina Journal of Law & Technology).

[6] According to U.S. law:

> [T]he use by the Government of web-based Internet applications and other information technologies . . . to (A) enhance the access to and delivery of Government information and services to the public, other agencies, and other Government entities; or (B) bring about improvements in Government operations that may include effectiveness, efficiency, service quality, or transformation.

44 U.S.C. § 3601(3) (2006).

[7] On December 8, 2009, the Office of Management and Budget (OMB) was instructed by the President to issue a much-awaited memorandum on Transparency and Open Government. The three principles necessary to achieve open government were listed as transparency, participation, and collaboration. PETER R. ORSZAG, OFFICE OF MGMT. AND BUDGET, MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES, 1 (2009) http://www.whitehouse.gov/omb/assets/memoranda_2010/m10-06.pdf (on file with the North Carolina Journal of Law & Technology).

[8] *See, e.g.*, Linda McGlasson, BankInfoSecurity, Mass. Privacy Law: Are You Compliant? (Feb. 26, 2010), http://www.bankinfosecurity.com/articles.php?art_id=2241 (on file with the North Carolina Journal of Law & Technology). Clicking on the hyperlink "Privacy" on the Government Information Security website produced two articles, one of which did not even mention privacy. Tom Field, GovInfoSecurity, Marcus Ranum: The Biggest Security Threats Getting the Least Attention, Dec. 30, 2009, http://www.govinfosecurity. com/articles.php?art_id=2032 (on file with the North Carolina Journal of Law & Technology).

[9] "Cyber security," in the view of one legal scholar, is a "concept that arrived on the post Cold War agenda in response to a mixture of technological

Even legislative activity has focused on security: two bills proposed in the last year, Senator Tom Carper's Information and Communication Enhancement Act and Senator Gillibrand's International Cybercrime Reporting and Cooperation Act, have focused on improving cyber security and dealing with cyber crime.[10] Compounding this is the view among computer experts that security is paramount to achieving privacy.[11] Alongside such

---

innovations and changing geopolitical conditions." The term, therefore, implies much more than the security of individual privacy. According to the Copenhagen School theory of securitization, security is "a speech act that *securitizes*, that is constitutes one or more referent objects, historically the nation or the state, as threatened to their physical or ideational survival and therefore in urgent need of protection. Lene Hansen & Helen Nissenbaum, *Digital Disaster, Cyber Security, and the Copenhagen School*, 53 INT'L STUD. Q. 1155 (2009), *available at* http://www.nyu.edu/projects/nissenbaum/papers/Digital%20Disaster,%20Cyber%20Security%20and%20the%20Copenhagen%20School.pdf.

[10] Since last year, two bills have been proposed to improve cyber security. Senator Tom Carper's bill to reform the Federal Information Security Management Act ("FISMA"), the United States Information and Communications Enhancement Act ("ICE Act"), was introduced in the Senate on April 28, 2009, but remains in committee. The original bill proposed the establishment of a National Office for Cyberspace in the Executive Office of the President that would have overseen federal government-wide "implementation of policies, standards, and guidelines on information security." S. 921, 111th Cong. § 3553(c)(2) (2009).

Senator Kirsten Gillibrand and Senator Orrin Hatch introduced the International Cybercrime Reporting and Cooperation Act on March 23, 2010 which seeks greater cooperation with other countries to fight cybercrime. *Proposed US Law Would Single Out Cybercrime Havens*, REUTERS, Mar. 23, 2010, http://www.reuters.com/article/idUS190003768320100324 (on file with the North Carolina Journal of Law & Technology).

[11] According to Tim Mather, a computer security expert, "[y]ou can have security without privacy, but you cannot have privacy without security." TIM MATHER ET AL., CLOUD SECURITY AND PRIVACY 145 (2009). *But see* Bruce Schneier, *What Our Top Spy Doesn't Get: Security and Privacy Aren't Opposites*, WIRED NEWS, Jan. 24, 2008, http://www.schneier.com/essay-203.html (on file with the North Carolina Journal of Law & Technology) (arguing that security and privacy are not either or alternatives; if they were, countries with high levels of government surveillance would be very secure, but that is not the case).

discussion have been calls for greater vigilance to guard against attacks similar to the ones[12] that occurred at Google.[13]

This Recent Development examines how privacy issues may be implicated by the technology—"cloud computing"—that enables Apps.Gov. Specifically, it analyzes the gap between law and technology in the context of Privacy Impact Assessments (PIAs), a tool for agency officials provisioned in the E-Government Act of 2002 (E-Gov Act). Part II briefly describes the technical underpinnings of cloud computing, especially virtualization technology, and the relevant legislative background. Part III discusses the legal issues that may arise out of the necessarily incongruent mapping of existing legal definitions onto an emerging technology. As a concrete example of how privacy analysis may break down in the context of Apps.Gov, the use of PIAs is considered. Part III also discusses some of the near-term consequences of this incongruence on privacy and some possible solutions.

II. BACKGROUND

A. *Cloud Computing*

Cloud computing is the technology that enables and is at the heart of the Apps.Gov project. It is a subcategory of the broader

---

[12] *See supra* note 1.

[13] On February 2, 2010, the Director of National Intelligence stated:

> The United States confronts a dangerous combination of known and unknown vulnerabilities, strong and rapidly expanding adversary capabilities, and a lack of comprehensive threat awareness. Malicious cyber activity is occurring on an unprecedented scale with extraordinary sophistication. . . . Sensitive information is stolen daily from both government and private sector networks, undermining confidence in our information systems, and in the very information these systems were intended to convey.

Annual Threat Assessment of US Intelligence Community for the S. Select Comm. on Intelligence, 111th Cong. 2 (2010) (statement of Dennis Blair, Director of National Intelligence) *available at* http://www.dni.gov/testimonies/20100202_testimony.pdf.

genus of information technology ("IT").[14] According to the National Institute of Standards and Technology ("NIST"), the federal technology agency that has been closely studying cloud computing for the purpose of providing guidance on securing unclassified government systems, the term refers to "a model for enabling convenient, on-demand network access to a *shared* pool of *configurable* computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."[15] Whereas most computers are owned and managed by consumers today, cloud computing offers a model in which computing resources—generically termed "cloud services"—are owned and managed by an organization which rents or leases it to consumers over the Internet.[16] In this way, virtualization software

[14] "The branch of technology concerned with the dissemination, processing, and storage of information, esp. by means of computers." OXFORD ENGLISH DICTIONARY 945 (2d ed. 1989)

[15] PETER MELL & TIM GRANCE, NIST, THE NIST DEFINITION OF CLOUD COMPUTING 1 (2009), http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc (emphasis added) (on file with the North Carolina Journal of Law & Technology). Like other technical buzzwords, cloud computing is susceptible to many definitions none of which are likely to be entirely inaccurate. Another, less technical, definition comes from a market research report:

> Cloud computing represents a new way to deploy computing technology to give users the ability to access, work on, share, and store information using the Internet. The cloud itself is a network of data centers—each composed of many thousands of computers working together—that can perform the functions of software on a personal or business computer by providing users access to powerful applications, platforms, and services delivered over the Internet.

JEFFREY F. RAYPORT & ANDREW HEYWARD, MARKETSPACE, ENVISIONING THE CLOUD: THE NEXT COMPUTING PARADIGM ii (2009), http://market spaceadvisory.com/cloud/ Envisioning-the-cloud.pdf (on file with the North Carolina Journal of Law & Technology). The NIST definition, nevertheless, is considered to be one of the best working definitions for its breadth and comprehensiveness. J. Nicholas Hoover, *NIST Team Deeply Studying Cloud Computing*, INFORMATION WEEK, June 3 2009, http://www.informationweek .com/news/government/enterprise-architecture/showArticle.jhtml?articleID=217 701603 (on file with the North Carolina Journal of Law & Technology).

[16] The online retailer Amazon.com, for example, provides cloud services via its Amazon Elastic Cloud Compute (EC2) network. Consumers are able to

drives the aggregation of computing resources and its delivery in the fashion of a metered utility.[17]

Like other technologies, cloud computing will have its users[18] and providers.[19] Using their broadband Internet connections, users will be able to access cloud services without being tied to the particular computer that holds the program or the data they want to use. The data and programs will be managed by the provider. Although they may not be aware of it, most Americans already have some experience with cloud computing.[20] They may have accessed webmail programs (e.g. Hotmail, Gmail, or Yahoo! Mail), backed up computer files online, stored personal videos online (e.g. YouTube), used online applications (e.g. Google Documents and Adobe Photoshop Express), and visited social networking sites (e.g. Facebook and Twitter).[21]

According to NIST, there are three service model categories providers can fall into.[22] The first, Software-as-a-Service ("SaaS"), provides software to users.[23] The user does not manage or control the actual physical computer networks belonging to the provider.[24]

---

purchase storage space and computing power based on a tiered monthly rate. Amazon Web Services, Amazon Elastic Compute Cloud (EC2), http://aws.amazon.com/ec2/ (last visited Apr. 12, 2010) (on file with the North Carolina Journal of Law & Technology).

[17] *See infra* note 64 (discussing the evolution of computing).

[18] A user is typically a customer of a cloud computing service. This may be an individual, a corporation, or any government agency.

[19] A provider is the entity that provides the cloud computing service and may be an individual, a corporation, or any other business organization.

[20] According to a survey by the Pew Internet and American Life Project, around sixty-nine percent of online Americans have used cloud computing services ("cloud services") in some form or another. JOHN B. HORRIGAN, PEW INTERNET & AMERICAN LIFE PROJECT, USE OF CLOUD COMPUTING APPLICATIONS AND SERVICES 1 (2008), http://www.pewinternet.org/~/media//Files/Reports/2008/PIP_Cloud.Memo.pdf.pdf (on file with the North Carolina Journal of Law & Technology).

[21] *Id.*

[22] MELL & GRANCE, *supra* note 15.

[23] The term "infrastructure" is used here to refer to the physical computing hardware and associated software owned by the provider.

[24] Most consumers use cloud services in this model.

The second, Platform-as-a-Service ("PaaS") gives users limited control over the software so long as it does not interfere with the physical infrastructure of the provider's network.[25] The third, Infrastructure-as-a-Service ("IaaS"), provides users control over limited parts of the cloud infrastructure. Although the physical network remains under the control of the provider, the consumer can now decide what operating systems to run and how to allocate memory and computing resources.

B. *Virtualization Software*

Virtualization software is what makes cloud computing possible. The concept of virtualization is not new at all nor is it solely associated with cloud computing.[26] Almost any time software is used, hardware—a computer's central processing unit, memory, and input/output devices—is virtualized, or hidden, in order to allow users to interact and make use of a computer.[27] Virtualization software does for a network of computers what an operating system does for a computer—it hides the "bricks and mortar" of computing hardware beneath a layer of software to create a "virtual infrastructure."[28] Upon this virtual infrastructure,

[25] Social networking applications available on Apps.Gov would utilize this model.

[26] The concept of virtualization originates from the 1960s when it was pioneered by IBM as a way of providing users of its mainframe computers with "concurrent, interactive access." Susana Nanta & Tzi-cker Chiueh, SUNY at Stony Brook, A Survey on Virtualization Technologies 1 (2005), http://www.ecsl.cs.sunysb.edu/tr/TR179.pdf (last visited Apr. 12, 2010) (on file with the North Carolina Journal of Law & Technology).

[27] All computers come equipped with a central processing unit (CPU), memory, and input/output (I/O) devices. In order to interact with a computer, software is required to abstract the hardware and present an interface that most individuals can understand and use. The first layer of abstraction software is the computer's operating system. Software like Microsoft Word or Internet Explorer are commonly referred to as "application" software—software that has particular uses.

[28] "Virtualization is a technology that combines or divides computing resources to present one or many operating environments using methodologies like hardware and software partitioning or aggregation, partial or complete machine simulation, emulation, time-sharing and many others." Nanta &

virtual computers, or "virtual machines,"[29] can be created by an additional layer of virtualization software.[30] A virtual machine interacts with the virtual infrastructure and is composed of the physical resources provided to it by the virtual infrastructure. The power of cloud computing is to maximize the efficiency of a given network of physical computers by rapidly allocating resources for computational tasks and then releasing those resources back to the virtual infrastructure so that they can be reallocated.

VMWare, Inc. ("VMWare") is a leading vendor of virtualization software.[31] The company advertizes a range of software products for cloud computing.[32] vSphere is installed on a network of computers and creates a virtual infrastructure.[33] Another product, VMWare ESXi,[34] creates the virtual machines that run on top of the virtual infrastructure created by vSphere.

Chiueh, *supra* note 26, at 1. Virtualization can be applied to an entire computer and also to components, including the computer's central processing unit (CPU) or memory. *Id*. at 2.

[29] The concept of a virtual machine is not new either. *Id.* Virtualization provided each mainframe user with a virtual machine which gave that user the illusion that he/she was interacting with the computer directly. *Id*. A virtual machine is a self-contained copy of the computer's operating system and, therefore, could be experimented with without affecting other users. *Id.*

[30] It can be used to combine, as in the case of Apps.Gov, the resources of an entire network of computers.

[31] VMWare, http://www.vmware.com (last visited Apr. 10, 2010) (on file with the North Carolina Journal of Law & Technology).

[32] VMWare, Server & Datacenter Virtualization Products, http://www.vmware.com/products/datacenter-virtualization.html (last visited Apr. 12, 2010) (on file with the North Carolina Journal of Law & Technology).

[33] VMWare, VMWare vSphere 4, http://www.vmware.com/products/vsphere/ (last visited Apr. 12, 2010) (on file with the North Carolina Journal of Law & Technology).

[34] VMWare, VMWare ESXi, http://www.vmware.com/products/esxi/ (last visited Apr. 12, 2010) (on file with the North Carolina Journal of Law & Technology).

C. *Apps.Gov*[35]

Apps.Gov is a portal where federal agencies will be able to purchase cloud services from a variety of providers, primarily Google.[36] There are two broad categories of services offered to agencies: cloud offerings[37] and social media tools.[38] The Cloud Computing Program Management Office within the General Services Administration ("GSA") coordinates its activities to leverage the benefits of cloud computing technology across the government.[39] It has developed "federal-friendly"[40] Terms of

[35] For purposes of this Recent Development, Apps.Gov will refer to not only the Government Services Administration (GSA) website, but more broadly to the use by federal agencies of cloud services.

[36] Google has not made publicly available information about the exact location of the physical data centers that will host these services, except that they will be in the United States and will be shared by all federal agencies. However, we do know that the cloud will be a "community cloud," whereby several organizations share the cloud's physical infrastructure. Posting of Michael Glotzbach to Official Google Enterprise Blog, http://googleenterprise .blogspot.com/2009/09/google-apps-and-government.html (Sept. 15, 2009, 11:45 AM) (on file with the North Carolina Journal of Law & Technology). The "community," presumably limited to federal agencies, all share similar security, mission, policy, and compliance concerns.

[37] These include business applications (analytics, asset management, business intelligence, business management, financial, medical, etc) and productivity applications (collaboration/meeting software, document and content management, and other office software tools and suites, etc).

[38] The Apps.Gov website describes these as follows: "Social media can take many different forms, including Internet forums, social blogs, wikis, podcasts, photos, videos, rating and bookmarking. Technologies include: blogs, video hosting, photo-sharing, wall-postings, email, instant messaging, and music sharing. Applications are computer programs or services available over the web." Apps.Gov, Frequently Asked Questions, https://apps.gov/cloud/ advantage/main/start_page.do (follow "Cloud FAQs" hyperlink; then follow "What are social media applications?" hyperlink) (last visited Apr. 12, 2010) (on file with the North Carolina Journal of Law & Technology).

[39] Apps.Gov, Frequently Asked Questions, https://apps.gov/cloud/advantage /main/start_page.do (follow "Cloud FAQs" hyperlink; then follow "What is the role of GSA in supporting the Federal Cloud Computing Initiative?" hyperlink) (last visited Apr. 12, 2010) (on file with the North Carolina Journal of Law & Technology). According to the FAQ posted on the website, the GSA is responsible for:

Service ("TOS") agreements that each agency can tailor to its specific needs, expectations, and practices.[41]

D. *The Privacy Act of 1974*

The Privacy Act regulates the government's use of personal information.[42] It applies to all agencies working with personal information contained in a system of records. The Privacy Act reins in agency use of personal information by requiring them to give notice before establishing any record systems,[43] restricts the transfer of records between agencies,[44] and gives individuals the right to review and correct any such records.[45] The Act also applies to any system of records whenever an agency "provides by a contract for the operation by or *on behalf of* the agency of a system of records to accomplish an agency function."[46]

> [T]he coordination of GSA's activities with respect to the Initiative via its Program Management Office (CC PMO). GSA and the CC PMO are focusing on implementing projects for planning, acquiring, deploying and utilizing cloud computing solutions for the Federal Government that increase operational efficiencies, optimize common services and solutions across organizational boundaries and enable transparent, collaborative and participatory government.

Id.

[40] The GSA recognized that typical vendor TOS agreements do not comply with federal laws and the specific needs of federal users. Hence, it developed amended TOS agreements which agencies could use as a starting point for their negotiations with providers. General Services Administration, Apps.Gov Frequently Asked Questions, https://apps.gov/cloud/advantage/main/start_page.do (follow "Cloud FAQs" hyperlink; then follow "Why is the government negotiating terms of services agreements with social media providers?" hyperlink) (last visited Apr. 12, 2010) (on file with the North Carolina Journal of Law & Technology).

[41] *Id.*

[42] U.S. GOV'T ACCOUNTABILITY OFFICE, PRIVACY ACT: OMB LEADERSHIP NEEDED TO IMPROVE AGENCY COMPLIANCE, GAO-03-304, at 5 (2003), *available at* http://www.gao.gov/new.items/d03304.pdf.

[43] 5 U.S.C. § 552a(e)(4) (2006).

[44] *Id.* § 552a(c).

[45] *Id.* § 552a(d).

[46] *Id.* § 552a(m) (emphasis added).

The basic building block of analysis under the Privacy Act is the system of records. A "record" is any "item, collection, or grouping of information" that contains an individual's "name, or the identifying number, symbol, or other identifying particular assigned to the individual, such as a finger or voice print or a photograph."[47] A "system of records," is any information that can be "retrieved by the name of the individual or by some identifying number, symbol, or other identifying particular assigned to the individual."[48]

E. *E-Government Act and the Federal Information Security Management Act*

In addition to the Privacy Act, the E-Government Act ("E-Gov Act")[49] also contributes to the protection of personal information. Under § 208 of the Act, agencies are required to conduct PIAs. A PIA is a report prepared by an agency's privacy officers analyzing how personal information is used in a federal information system.[50]

Title III of the E-Gov Act, entitled the Federal Information Security Management Act ("FISMA"),[51] directed agencies to create information security policies.[52] It institutionalized a risk management approach[53] to information security and required the

---

[47] *Id.* § 552a(a)(4).
[48] *Id.* § 552a(a)(5).
[49] E-Government Act of 2002, Pub. L. No. 107-347, 116 Stat. 2899.
[50] *See infra* Part III.C.
[51] 44 U.S.C. §§ 3541–49 (Supp. II 2004). FISMA consolidates fragments of earlier computer security laws scattered across the U.S. Code such as the Computer Security Act of 1987, the Clinger-Cohen Act and the Paperwork Reduction Acts of 1980. H.R. REP. NO. 107-787, pt. 1, at 54 (2002).
[52] *See, e.g.*, 44 U.S.C. § 3544(a)(1) (2006).
[53] FISMA introduced oversight and accountability by requiring agencies to use a cost-benefit analysis approach when assessing threats to information security. Each agency is called upon to assess the risk to its systems, determine the magnitude of the harm that would occur if its systems were compromised, and then to implement safeguards which are proportionate to the identified risks to protect against such breaches. 44 U.S.C. § 3543(a)(2) (2006) (requiring agencies to "identify and provide information security protections *commensurate with the risk and magnitude of the harm* resulting from the

NIST to develop minimum standards of compliance.[54] Such a framework would cover both federal information systems and also any systems "used or operated . . . by a contractor of an agency or other organization on behalf of an agency."[55]

F. *Privacy Impact Assessments*

PIAs were introduced in the E-Gov Act under § 208[56] to assess the impact of information technology projects on privacy.[57] They "explain how an agency takes into account privacy considerations when purchasing and creating new information systems, and when initiating collections of information."[58] Agencies are required to

---

unauthorized access, use, disclosure, disruption, modification, or destruction of" federal government data) (emphasis added).

[54] According to the statute:

> The purpose of FISMA is to permanently authorize a government-wide risk-based approach to information security . . . and to further strengthen Federal Information security by requiring compliance with minimum mandatory management controls for securing information and information systems, clarifying and strengthening current management and reporting requirements, and strengthening the role of [NIST].

Id.

[55] 44 U.S.C. 3544(a)(1)(A)(ii) (Supp. II 2004).

[56] E-Government Act of 2002, Pub. L. No. 107-347, § 208, 116 Stat. 2899, 2921–23.

[57] S. REP. NO. 107-174 (2002). During Senate hearings, the Center for Democracy and Technology applauded the use of PIAs:

> In particular mandatory privacy impact assessments in all government IT projects will be highly beneficial. For instance, if an agency wanted to share information with another agency, this bill would force it to consider whether this amalgamation of data adequately protected the people the agency served. These assessments would be similar to the environmental impact assessments that agencies must perform before embarking on such projects. These have been successful in making agencies accountable for the decisions with regard to building projects. It is hoped that privacy impact assessments will have similar consequences.

E-Government Act of 2001: Hearing Before the S. Comm. on Governmental Affairs, 107th Cong. 253 (2001) (statement of The Center for Democracy and Technology).

[58] S. REP. NO. 107-174, at 28.

conduct PIAs when they utilize, either by developing or procuring, information technology for such purposes.[59] Not all information collection is subject to this requirement, only information that is in "identifiable form" i.e. information that discloses the personal identities of individuals.[60]

The contents of a PIA must address:

(I) what information is to be collected; (II) why the information is being collected; (III) the intended use of the agency of the information; (IV) with whom the information will be shared; (V) what notice or opportunities for consent would be provided to individuals regarding what information is collected and how that information is shared; (VI) how the information will be secured; and (VII) whether a system of records is being created under section 552a of title 5, United States Code (commonly referred to as the "Privacy Act").[61]

Of these, (VI) and (VII) present the most interesting legal questions. To date there has been no FISMA litigation where the information was maintained by a private contractor on behalf of an agency. The "system of records" analysis under the Privacy Act has remained very controversial since the age of modern computer databases and will only become more so in the age of virtual machines.

III. DISCUSSION

A. *The Technology-Law Divide Presented by Apps.Gov*

The E-Gov Act requires PIAs to be "commensurate with the size of the *information system* being assessed, the sensitivity of information that is in an identifiable form in that system, and the

[59] E-Government Act of 2002, Pub. L. No. 107-347, § 208(b)(1)(A), 116 Stat. 2899, 2921–22..

[60] Identifiable information is "any representation of information that permits the identity of the respondent to whom the information applies to be reasonably inferred by either direct or *indirect* means." *Id.* § 502(4), 116 Stat. at 2962 (emphasis added).

[61] *Id.* § 208(b)(2)(B)(ii), 116 Stat. at 2922.

risk of harm from unauthorized release of that information[.]"[62] The first task for privacy officers, therefore, will be to interpret what constitutes an information system.

Our current understanding of computers is such that an information system has roughly corresponded to a physical machine i.e. a computer. This has made the task of privacy oversight relatively straightforward because the boundaries have been physical ones. This understanding was valid as far back as the Paperwork Reduction Act of 1980 and as late as 2002 when the E-Gov Act was passed.

At present, however, we are in the midst of a significant change in our understanding of computing that threatens to shift the ground beneath our current policy discourse.[63] This change reflects a long-term cyclical shift in computing from centralized to personal and back to centralized.[64] It used to be in the 1960s and 1970s that computers were owned only by large corporations where they assisted management in decision-making.[65] With the emergence of the personal computer in the late 1970s and 1980s computing power suddenly became affordable to the ordinary consumer and its use became widespread.[66] Individuals and small businesses could now tap into the power of spreadsheets and document processing software. This decentralization of computing

---

[62] *Id.*, § 208(b)(2)(B)(i), 116 Stat. at 2922.

[63] *See, e.g.*, *Clouds and Judgment*, THE ECONOMIST, Oct. 25, 2008, at 17 (examining the paradox presented by a technology that turns computing into a "borderless utility" while being subject to laws that are "mainly local").

[64] In the 1960s and 1970s most of the world's computing power was found in mainframe computers owned by large corporations which were shared by numerous users. The 1980s and 1990s marked a shift to personal computers with the introduction of personal computers and hand-held devices. Cloud computing is swinging the pendulum back to centralized model. DAVID WYLD, IBM CTR. FOR THE BUSINESS OF GOVERNMENT, MOVING TO THE CLOUD: AN INTRODUCTION TO CLOUD COMPUTING IN GOVERNMENT 16 (2009) http://www.businessofgovernment.org/pdfs/WyldCloudReport.pdf (on file with the North Carolina Journal of Law & Technology).

[65] *See also infra* text accompanying notes 69–76 (discussing the 1980 amendments to the definition of an "information system" as a "management information system" to its broader present-day definition).

[66] WYLD, *supra* note 64, at 16.

power continued into the 1990s with the explosion in mobile computing and personal hand-held devices.[67] We are now at the beginning of a return to a more centralized model of computing. Computer hardware has become so commoditized, and the Internet has made distance so irrelevant, that it makes sense economically to aggregate computing power in the hands of large corporations and accessing that power much as one would access a public utility.[68]

B. *The "Information System"[69] as a Unit of Assessment*

The most recent definition of information system[70] comes from the 1995 revisions to the Paperwork Reduction Act.[71] In those revisions the definition was broadened from its original definition of "management information system"[72] to the current definition of

---

[67] *Id.*

[68] Although we have not yet reached the utility computing stage, some argue that the stage has been set for the emergence of metered computing services. *Id.* Features of cloud computing, such as the on-demand access to pooled information technology services, have been used to draw historical comparisons to electricity and power generation. *Id*. at 56. Changes in the production and distribution of electricity in the early twentieth century led users to throw out their own generators and start buying electricity on a metered basis from large electricity producers. *Id.*

Two drivers of that earlier development are also seen to be operating with respect to computing today. *Id*. at 57. The first is wasted computing power—most computing power is owned and managed by corporate and individual users and a lot of it remains unused. The fragmented and unused computing capacity is providing strong incentives for centralizing its supply. *Id*. at 58 (quoting Nicholas G. Carr, *The End of Corporate Computing*, MIT SLOAN MGMT. REV., Spring 2005, at 67, 73) ("The history of commerce has repeatedly shown that redundant investment and fragmented capacity provide strong incentives for centralizing supply.").

The second driver, commoditization, has made computers the building blocks of even larger "computers." Computer hardware can now support a wide variety of software from different vendors and there are less of the compatibility issues that plagued earlier software. WYLD, *supra* note 64, at 58.

[69] 44 U.S.C. § 3502(8) (2006).

[70] *Id.*

[71] *Id.*. § 3501.

[72] *Id*. § 3502(14) (1994), *amended by* Paperwork Reduction Act of 1995, Pub. L. No. 104-13, § 3502(8), 109 Stat. 163, 166.

any "discrete set of information resources organized for the collection, processing, maintenance, use, sharing, dissemination, or disposition of information."[73] An "information resource"[74] includes "information and related resources, such as personnel, equipment, funds, and information technology."[75] The expanded definition was intended to update the law by recognizing the full range of uses information systems beyond the functions they served in corporations.[76]

Although the theory and practice of virtualization technology is not new, virtualization software such as VMWare[77] breaks new ground in the manner and degree to which computing resources are virtualized. There is the first degree of virtualization—a kind of "bare metal" virtualization that allows the resources of a single physical computer to be shared by multiple virtual machines. Each virtual machine believes it is accessing the underlying physical computer directly but is in fact sharing it with other virtual machines. A second degree of virtualization—virtual infrastructure—virtualizes the resources of more than one physical computer. Virtual machines in a virtual infrastructure may believe they are accessing a single physical computer directly, but are in fact sharing the resources of multiple computers with other virtual machines. Virtualization blurs the physical boundaries between computers to the point of erasing them. It allows virtual machines to use the resources of multiple physical computers with the result that the data storage and processing capabilities of a virtual machine are dispersed among multiple physical computers.[78]

---

[73] *Id.*

[74] *Id.* § 3502(6).

[75] *Id.*

[76] According to a House Report accompanying the amendment, information systems today "serve a much broader range of purposes than just providing management information." H.R. REP. NO. 104-37 (1995).

[77] VMWare, *supra* note 31.

[78] Data dispersal has been discussed by senior computer scientists at NIST. However, most of the discussion has been limited to understanding how might the laws of foreign jurisdictions, e.g. the European Union, impact cloud computing. Peter Mell & Tim Grance, Nat'l Inst. of Standards and Tech., Info. Tech. Lab., Effectively and Securely Using the Cloud Computing Paradigm,

This blurring has legal implications as well because virtualization software may inadvertently allow agencies to offload privacy responsibilities if the software acts as an intermediary between the government and third parties. FISMA applies to information systems that are owned or operated on or behalf of an agency.[79] In order for FISMA to apply, therefore, privacy officers will have to distinguish between components of virtualization software that are operating on behalf of an agency and those that are not. Similarly, the Privacy Act applies to system of records that are operated "by or on behalf of" an agency.[80] Here too, privacy officers will have to distinguish between components of virtualization software that operate on behalf of an agency and those that do not, a task made more challenging by our current understanding of computers. An important threshold question when conducting PIAs, therefore, will be whether something equivalent to an information system exists and whether such a system is being used by or on behalf of an agency.

The definition of an information system[81] and the accompanying definition of an information resource[82] are arguably broad enough to describe either a virtual machine or the set of physical computers whose resources are shared by that virtual machine with other virtual machines. Because only one agency will use a particular virtual machine, interpreting an information system to refer to a virtual machine has the added advantage of neatly aligning technical and administrative boundaries. It is not clear from reports how information system is being interpreted by agency officials.[83]

(May 15, 2009), http://csrc.nist.gov/news_events/HIPAA-May2009_workshop/presentations/6-051908-new-technologies-cloud-computing.pdf (on file with the North Carolina Journal of Law & Technology). There has not been much, if any, discussion of how data dispersal impacts privacy assessment under U.S. laws.

[79] *Supra* note 57.

[80] 5 U.S.C. § 552a(m) (2006).

[81] *See supra* text accompanying note 73.

[82] *See supra* text accompanying note 74–75.

[83] Google has applied for the certification and accreditation of the suite of software tools "Google Apps" under FISMA, but it is not clear whether such

Because the resources of a single physical computer may be shared by more than one virtual machine, it cannot be said definitively whether that particular physical computer is being operated "on behalf of *an* agency."[84] If an information system is being operated by or on behalf of more than one agency, it is not clear which agency's privacy standards or procedures should that system be regulated by. As each agency manages its own privacy policies and within any agency a privacy officer may also be responsible for information security policy, conflicting interpretations may arise. Assigning responsibility to one agency to apply a single privacy standard may create a conflict of interest.[85]

C. *Conducting PIAs in the Cloud*

Although there are seven elements of a PIA, the biggest impact in terms of interpreting the Privacy Act will likely be elements VI and VII—how information on the cloud will be secured and whether the collection of information on the cloud results in the creation of a "system of records" under the 1974 Privacy Act. Agencies will have more experience when answering elements I through V, and discussion of these elements is therefore omitted.

A PIA must include details on how information will be secured.[86] FISMA[87] states that it is applicable to "information

---

certification will also audit Google's physical computers. Glotzbach, *supra* note 36 (describing Google's efforts to obtain FISMA certification).

[84] *Supra* note 57 (emphasis added).

[85] Agency officials are charged with ensuring their particular agency is complying with privacy laws, not other agencies. This may be a bigger concern for agencies whose CIO doubles as a privacy officer and wants to ensure only a minimum level of compliance. *See, e.g.*, U.S. GOV'T ACCOUNTABILITY OFFICE, PRIVACY: AGENCIES SHOULD ENSURE THAT DESIGNATED SENIOR OFFICIALS HAVE OVERSIGHT OF KEY FUNCTIONS, No. GAO-08-603, at 7 (2008), *available at* http://www.gao.gov/new.items/d08603.pdf (on file with the North Carolina Journal of Law & Technology) (describing the varying requirements for senior privacy officials at different agencies, including situations where the CIO doubles as a privacy officer).

[86] *See supra* text accompanying note 63.

[87] Section 208 does not provide any guidance on what constitutes sufficient the security of an information system. However, title III of the same legislation

systems used or operated by an agency or *by a contractor of an agency* or other organization *on behalf of* an agency."[88] As discussed in the previous section, there is some ambiguity in how "on behalf of"[89] could be interpreted depending how an information system is defined.

Apportioning responsibility for FISMA compliance is complicated in both interpretations of the "on behalf of"[90] language of the statute. The virtual machine interpretation has the seeming advantage of delegating responsibility to the agency that uses the virtual machine. However, virtual machines cannot be so easily separated from the physical infrastructure as it may seem. Because virtual machines are supported by so many layers[91] of software, it is hard to distinguish the layer at which the software stops being a part of the information system operated on behalf of an agency and starts belonging to the cloud provider.[92] The "brick and mortar" interpretation[93] of an information system would be plagued by its own set of complications. Because the physical computers in a

---

(FISMA) states that its purpose is to "provide for [the] development and maintenance of minimum controls required to protect Federal . . . *information systems*" and it defines "information security" to include "means for protecting *personal privacy*."

[88] *Supra* note 57 (emphasis added).

[89] *Id.*

[90] *Id.*

[91] *See* discussion *supra* Part II.B. The different cloud computing service models, *see supra* text accompanying notes 22–27, also means that there may be more than one third party supplying components of the cloud. In the SaaS service model of cloud computing there may be as many as three parties: the provider of the physical infrastructure (in the case of Apps.Gov, Google), the provider of the application software which runs on the infrastructure (e.g. Facebook), and the provider of the virtualization software (e.g. VMWare). *Id.*

[92] A conservative approach to security would require all the layers of software to be compliant with FISMA. However, this would create overlapping domains of responsibility between different agencies since they all share the same physical and virtual infrastructure. A less burdensome approach would only require the virtual machine to be FISMA compliant. The danger here is that this may not provide enough protection in a shared computing environment. Some balance is required, but there is little guidance for privacy officers on such matters.

[93] *See supra* text accompanying note 28.

cloud's virtual infrastructure are shared by multiple virtual machines, such an approach will require determining which agency a particular physical computer is used "on behalf of."[94]

Accountability in the case of a breach at Apps.Gov may require a court to inquire whether a PIA was conducted and if so, whether it was adequate.[95] Google may argue that its duties under FISMA should be limited to that part of the infrastructure of which it has retained control under the contractual agreements[96] it signed; any software operating on its infrastructure is beyond its control. Much would depend on the nature of the breach itself. A fact-intensive inquiry would likely ensue, the results of which may not be conclusive given the complexity of virtualization.

The Privacy Act applies to systems of records operated by contractors on behalf of agencies.[97] It does not apply, however, if the records are not maintained on behalf of the government, such as in the case of information brokers.[98] It is this proviso that has allowed the government to gain access to large private repositories

---

[94] *Supra* note 57.

[95] As stated in a recent law review article, the production of PIAs does not involve much public consultation and the one instance where a privacy advocacy group submitted a FOIA request for the production of a PIA was rejected by a federal court. Kenneth A. Bamberger & Deirdre K. Mulligan, *Privacy Decisionmaking in Administrative Agencies*, 75 U. CHI. L. REV. 75, 87 (2008) (citing Elec. Privacy Info. Ctr. v. Transp. Sec. Admin., 2006 WL 626925, at *10 (D.D.C. Mar. 12, 2006)).

[96] FOIA requests filed by the Electronic Privacy Information Center have resulted in the production of a number of agreements the GSA signed with various cloud service providers, e.g. Yahoo!, Google, and YouTube. Elec. Privacy Info. Ctr.—Cloud Computing, http://epic.org/privacy/cloudcomputing/ (under bulleted item "EPIC Forces Disclosure of Government Contracts with Social Media Companies, Privacy Terms Missing") (last visited Apr. 13, 2010) (on file with the North Carolina Journal of Law & Technology). None of the agreements explicitly discuss privacy policies to be adopted by the providers or assignment of responsibility in case of a breach. *Id.*

[97] 5 U.S.C. § 552a(m) (2006).

[98] *See* DANIELLE K. CITRON, FULFILLING GOVERNMENT 2.0'S PROMISE WITH ROBUST PRIVACY PROTECTIONS 17 (March 2010), http://groups.law.gwu.edu/LR/ArticlePDF/Arguendo_Citron.pdf (on file with the North Carolina Journal of Law & Technology).

of personal information.[99] The same proviso could now also be used to access personal information collected through an intermediation mechanism that may inadvertently be created by Apps.Gov.[100] As in the case of FISMA, the boundary between an agency's information systems and that of the provider will become harder to distinguish in a virtualized cloud environment. For example, if a social networking application such as Facebook generates personal information it will have to be determined what part of the virtual infrastructure it resides on. Depending on whether that component is considered a part of an information system operated or used on behalf of an agency or a part of the provider's virtual infrastructure rights under the Privacy Act may or may not be available.[101]

Up till now computer databases have been maintained by individual agencies on their own premises.[102] Physical barriers

---

[99] Information resellers like ChoicePoint, Inc. have created massive personal information databases over the years. They collect, collate, and sell consumer information to a wide range of buyers, including the federal government. *See, e.g.*, Daniel J. Solove, *Access and Aggregation: Public Records, Privacy and the Constitution*, 86 MINN. L. REV. 1137, 1151 (2002); *see also* Joseph T. Thai, *Is Data Mining Ever a Search Under Justice Stevens's Fourth Amendment?* 74 FORDHAM L. REV. 1731, 1737 (2006). ChoicePoint's databases, as of 2001, contained ten billion records and the company had contracts to share its data with at least thirty-five federal agencies. *Id*. (citing Glenn R. Simpson, *Big Brother-in-Law: If the FBI Hopes to Get The Goods on You, It May Ask ChoicePoint*, WALL ST. J., Apr. 13, 2001, at A1).

Incidentally, the growth of an information marketplace has been encouraged by a utilitarian approach to privacy. For example, the Privacy Act does not apply to private corporations. In addition, the Supreme Court has held that the Fourth Amendment does not apply once private information is made public. *See, e.g.*, Katz v. United States, 389 U.S. 347, 351 (1967) ("What a person knowingly exposes to the public, even in his own home or office, is not a subject of Fourth Amendment protection.").

[100] *See* discussion *supra* Part III.B.

[101] *See* CITRON, *supra* note 98 (describing the potential threat to privacy created by the Government accessing an individual's social media data on Government 2.0 websites).

[102] The interagency sharing of information has been subject to the Privacy Act and requires the sharing to be governed by "matching agreements" or qualify under the "routine use" exemption. Matching agreements govern computer

have existed between databases belonging to different agencies. The question of whether a system of records has been maintained on behalf of an agency, therefore, has rarely arisen. This will likely change, however, if databases start to reside on a virtual infrastructure such as the one behind Apps.Gov.[103] The challenge of establishing that a system of records was operated on behalf of an agency and the aggregation[104] of potentially large collections of personal information will present a new landscape to agency officials charged with overseeing compliance with privacy laws.

D. *Consequences of the Divide*

Leaving the gaps between our current privacy laws and cloud technology unexamined will likely weaken the government's ability to assure the public of privacy protection which in turn would obstruct the Open Government objectives as outlined by the

---

matching programs and must specify the purpose for which the program was created, the data that will be shared, the time limits during which the exchange will occur, the procedures for verifying the shared data, and procedures for retention and destruction of the data. 5 U.S.C. § 552(o) (2006).

[103] Google is building separate data centers dedicated solely for the government's use. The centers will be located within the United States and will not be shared with other Google customers. Glotzbach, s*upra* note 36.

[104] The concerns are not entirely dissimilar to the ones that led to the Privacy Act. One of the principal motivating factors for the Privacy Act was the increasing use of computers to establish large "databanks"—"a collection of information about individuals assembled in one place for easy access by a number of users." ALAN F. WESTIN & MICHAEL A. BAKER, NATIONAL ACADEMY OF SCIENCES, DATABANKS IN A FREE SOCIETY: COMPUTERS, RECORD-KEEPING AND PRIVACY 9 n.1 (1972); *see also* Solove, *supra* note 99 (describing the contemplated plans for a National Data Center that had to be scrapped in response to public concerns about privacy):

> The Johnson Administration had contemplated creating a National Data Center that would combine information collected by various federal agencies into one large computer database, but the plan was scrapped after a public outcry. In the early 1970s, John E. Holt, an official in the General Services Administration, proposed that all of the federal government's computer systems be connected in a network called FEDNET. Responding to a public outcry, Vice President Gerald Ford stopped the plan.

Id. at 1192.

current administration.[105] The principal reasons for the analytical and conceptual gaps would be privacy officials operating in a rapidly changing environment and continuing to work with little guidance on how to interpret the legal provisions governing privacy.[106]

Neither has there been sufficient litigation under either FISMA or the E-Gov Act to provide guidance on what constitutes an information system.[107] What little litigation has occurred under FISMA has only indicated the judiciary's unwillingness to make any substantive evaluations of security procedures. In *Cobell v. Norton*,[108] the beneficiaries of an Individual Indian Money Trust sued the trustee, the Department of Interior (DOI).[109] The plaintiffs alleged among other things that DOI had breached its fiduciary duty to protect personal data and requested a preliminary injunction to disconnect some of the department's computers from the Internet.[110] The DC district court granted the injunction after examining the security controls put in place by the DOI.[111] However, on appeal the DC Circuit reversed, holding that the balance of equities favored the DOI.[112]

The *Cobell* case is significant because the D.C. Circuit's opinion suggested that the judiciary may not be in a position to provide guidance on any substantive security issues. A court reviewing an action under the Privacy Act may similarly be

---

[105] *See* text accompanying note 7, *supra*.

[106] *See supra* note 39, at 14. In addition to the Privacy Act of 1974 and the Paperwork Reduction Act of 1980, a number of additional laws address the roles and responsibilities of officials within each agency to oversee privacy. The variation in these laws means that in some agencies the same individual is both the CIO and the Chief Privacy Officer. In other agencies, general counsel and other senior officials have undertaken the role of privacy officers in addition to their regular duties. *Id.*

[107] A LexisNexis search on keyword "FISMA" under federal court cases turned up only eight hits.

[108] 394 F. Supp. 2d 164 (D.D.C. 2005).

[109] *Id.*

[110] *Id.*

[111] *Id.*

[112] Cobell v. Kempthorne, 455 F.3d 301 (D.D.C. 2006).

inclined to defer to the agency official's interpretation of what constitutes an information system.[113] Despite glaring inadequacies in computer security at DOI,[114] the circuit court restricted its judgment to whether the district court's decision to grant the injunction was justified, "not whether Interior's actions have been sufficient to achieve the goals of FISMA."[115] The circuit court's analysis centered on balancing equities and hardships. It set a particularly high standard for reviewing agency security measures in such cases requiring plaintiffs to demonstrate "some imminent threat or specific reason to be concerned"[116] and was dismissive of the district court's review of the department's lax security measures, holding that "perfect security" is unattainable for computers.[117]

The rights provided for by the Privacy Act of 1974 apply only to information collections deemed to be a "system of records."[118] Despite criticism of the term's definition and its interpretation by courts, the drafters of the 2002 PIA provision[119] in the E-Gov Act decided to remain consistent with the Privacy Act and require officials to use it when conducting PIAs.[120] Only if such a record is

---

[113] This is especially so because PIAs were provisioned in the same legislation as FISMA, namely the E-Gov Act of 2002.

[114] The district court found that "no effort was made by [Bureau of Land Management] administrators to restrict, block, or deny access from the source" of attacks. *Cobell*, 394 F. Supp.2d at 165. An independent auditor found that the DOI Office of Surface Mining's "Intrusion Detection System had not been monitored or reviewed by anyone for approximately forty-five days and that an additional system was connected to the Internet for twenty-six days with no Intrusion Detection System implemented at all." *Id.*

[115] *Kempthorne*, 455F.3d at 314.

[116] *Id.* at 315.

[117] *Id.*

[118] 5 U.S.C. § 552a(5) (2006).

[119] *See* text accompanying note 63, *supra*.

[120] *See* H. R. REP. NO. 107-787, pt. 1, at 51 (2002). The committee in this House report stated its reasoning as follows:

> [T]he Committee intends that the OMB guidance on process for developing a Privacy Impact Assessment (PIA) be done in a way that allows for consistency with work done by agencies to assess privacy requirements under the Paperwork Reduction Act (PRA) and the

created will the Privacy Act place limitations on the use of personal information in it.

The definition of system of records cannot be understood from its ordinary, plain meaning.[121] Courts have interpreted it in a very specific way, generally so as to exclude rather than include what can be considered personal information.[122] The leading case for understanding whether an agency may establish a system of records is the 1996 D.C. Circuit case of *Henke v. United States Dep't of Commerce*.[123] The *Henke* court's analysis focused on whether the agency had an actual practice of retrieving a record by the person's name or other identifier.[124] The mere capability to

---

> Privacy Act of 1974, with regard to new collections of information that include personally identifiable information.

Id.

[121] *See, e.g.*, Julianne M. Sullivan, *Will the Privacy Act of 1974 Still Hold up in 2004? How Advancing Technology Has Created a Need for Change in the "System of Records" Analysis*, 39 CAL. W. L. REV. 395, 399 n.30 (2003). The definition provided in the Privacy Act is not based on the ordinary, plain meaning of the words "system of records," but is in fact a very specific type of system, with very particular rules. This distinction likely arose out of the need to create some sort of distinction between groups of records that should be accessible and those that should not.

[122] *Id.*

[123] 83 F.3d 1453 (D.D.C. 1996). The court held that mere capability to retrieve records by an individual's name is not sufficient to constitute a "system of records." The *Henke* court went further and restricted the "system of records" analysis by considering whether, regardless of the method by which a record was retrieved, the record was actually about the individual. Mere incidental naming would not be considered as such an analysis. *Id.*

By and large the view of the DC Circuit remains the norm; later decisions from other circuits and district courts have tended to follow *Henke*. However, the Fourth Circuit in *Williams v. Dept. of Veteran's Affairs*, 104 F.3d 670 (1997), criticized it for ignoring the definition's ambiguity and for deferring to the OMB's narrow construction.

[124] The court in *Henke* examined three factors to determine whether the agency had established a "system of records": what the function of the agency was, the purpose for which the information was gathered, and the agency's "actual retrieval practice and policies." *Henke*, 83 F.3d at 1461.

perform such retrievals was not sufficient to establish a system of records.[125]

Because of the way they are indexed, most modern computer databases provide users with sophisticated searching capabilities that far exceed the capabilities of older indexed collections.[126] However, as mentioned above, a system of records analysis under *Henke* focuses on an agency's actual practices; the technical capabilities of the system that holds the data are immaterial. *Henke*, therefore, protects agency use of advanced information technologies by making it harder for individuals to establish that an agency was maintaining a system of records.[127]

Because large aggregated collections of information are valuable intelligence resources,[128] it will nevertheless remain

---

[125] *Id.* at 1460.

[126] *See, e.g.*, Sullivan, *supra* note 121. Sullivan argues the court's interpretation is inadequate to police privacy violations in the age of modern computer databases. Unlike older recordkeeping systems, these databases do not have to be searched by name, social security number, or other traditional forms of indexing. Any search term can be used and the database will search all its contents to find a match.

[127] If an individual cannot establish that the agency established a system of records, the court's inquiry effectively ends. *See, e.g.*, *Fisher v. Nat'l Inst. Of Health*, 934 F. Supp. 464, 474 (D.C. Cir. 1996) (granting summary judgment to the NIH upon determining that the information held in its files did not constitute a system of records.).

[128] Aggregation of personal information in the modern computing era is inherently susceptible to misuse. *See*, Danielle K. Citron, *Reservoirs of Danger: The Evolution of Public and Private Law at the Dawn of the Information Age*, 80 S. CAL. L. REV. 241, 244 (2007). According to Professor Citron, the information storage technologies of today can be compared to water reservoirs of the Industrial Age:

> Computer databases are this century's reservoirs. Today, databases of personal identifying information in the private sector ensure the seamless flow of commerce. Social security numbers facilitate loans and instant credit. Employers and colleges use SSNs to identify employees and students. Over 1000 companies collect and sell our sensitive personal information. Databases of biometric information—fingerprint, retinal, iris, and facial images—increasingly authenticate retail transactions, secure workplaces, and provide access to corporate

important to monitor how and why agencies access such information. The consequences of allowing agencies access to personal information through the intermediation of a cloud provider may go beyond personal privacy concerns. Sophisticated information retrieval techniques such as data mining[129] have the capability of searching extremely large collections of data for trends and patterns and zeroing in on particular transactions of interest. In the language of law enforcement, such a technique provides "event-driven"[130] instead of direct, or "target-driven,"[131] surveillance. The former examines records containing transactional data whereas the latter examines records associated with a particular individual. There are, however, serious questions about the use of this technology by the government. Data mining, like any other technique based on statistical analysis, is imperfect in that it may produce false positive results—results in which an individual is incorrectly identified by the software.[132] Therefore, its

---

> computer networks. Much as water reservoirs drove the Industrial Age, computer databases fuel the Internet economy of our Information Age.

Id.

[129] Data mining refers to the use of computer-based analytic tools that sift through large collections of data searching for patterns based on statistical techniques. Information collected by organizations and stored in large databases is most often the subject of these forms of analysis. The results of the analysis lead to conclusions of overall trends, unusual activity patterns, and other user-specified parameters.

[130] An illustrative example of event-driven surveillance:

> Say, for instance, that the police know that a sniper-killer wears a particular type of shoe (thanks to mudprints near a sniper site), that he owns a particular type of sweater (because of threads found at another site), and that he reads Elmore Leonard novels (because of allusions to those books made in his communications to the police). Law enforcement understandably might want to peruse the purchase records of local shoe, clothing, and book stores as part of their investigation.

CHRISTOPHER SLOGOBIN, PRIVACY AT RISK: THE NEW GOVERNMENT SURVEILLANCE AND THE FOURTH AMENDMENT 12 (2007).

[131] Target-driven surveillance is when law enforcement has a particular suspect in mind and want more information about that person through the use of traditional criminal investigative techniques, some of which may require warrants and, therefore, harder to use than event-driven surveillance. *Id.* at 9.

[132] E.g. non-terrorists identified as terrorists. *Id.* at 194.

use for identifying individuals suspected of being terrorists has been of particular concern, especially given the impact such identification may have on that individual's civil liberties.[133] The use of private information for such uses and compliance with privacy laws has also been under review.[134]

In the absence of guidance from either OMB or other sources, in the event of misuse or misappropriation of personal information the effected parties will not be able to be notified and protect themselves. In addition, they may have no notice that their personal information exists on the cloud and have no opportunity for making corrections or seeking remedies. These are the foundational requirements of the Privacy Act and are threatened when agency officials either do not have sufficient authority or have little guidance to implement provisions such as § 208.

E. *Suggestions*

Uniformity of privacy practices differing among agencies, being forced onto the same[135] virtual infrastructure may help bring such practices into alignment. Up till now each agency has maintained its own computing systems with the result that the structure of privacy oversight and the substantive measures agencies have undertaken to protect privacy have been

---

[133] *See, e.g.*, U.S. GOV'T ACCOUNTABILITY OFFICE, DATA MINING: EARLY ATTENTION TO PRIVACY IN DEVELOPING A KEY DHS PROGRAM COULD REDUCE RISKS No. GAO-07-293 (2008), *available at* http://www.gao.gov/new.items /d07293.pdf (reviewing the Department of Homeland Security's terrorism detection program ADVICE). The possibility of identifying a non-terrorist as a terrorist through the government's use of data mining technology has been of special concern in the aftermath of the events of September 11th, 2001. The Transportation Security Administration's no-fly watch list has been the subject of a number of complaints involving such "false positive" identification. *See, e.g.*, *Green v. Transp. Sec. Admin.*, 351 F. Supp.2d 1119 (W.D. Wash. 2005).

[134] In a 2008 report by the Government Accounting Office (GAO), it was revealed that more work was needed to ensure the use of "data mining" did not compromise privacy protections. U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 133, at 2. The study found that no federal agency was complying with all the key procedures required by the Privacy Act. *Id.*

[135] *See supra* note 39, at 36.

inconsistent.[136] To the extent that agencies will begin purchasing some of their information technology needs from the same provider, there may be an incentive to implement similar measures and reduce work and cost of compliance. This process may take time to work itself out, however. In the meanwhile, it may be best for NIST to develop a "server test" or "virtual machine test" for agencies grappling with privacy issues on the cloud.[137] Although it is far from clear what, if any, standard would work best, a uniformly applied rule will reduce delays in conducting PIAs.

Agencies can also do a better job of informing the public when information is being collected over the Internet through more effective website privacy policies.[138] The agreements entered into by the GSA with various providers of cloud services do not address privacy obligations.[139]

Because of administrative and political considerations, providing guidance to privacy officers may remain unfeasible in the near term. It may also be advisable, therefore, to look at technological solutions. For example, the Privacy Preferences Project[140] under the auspices of the World Wide Web Consortium[141]

---

[136] *Id.*

[137] The "server test" found in copyright law was used in Perfect10 v. Google, Inc. 416 F. Supp. 2d 828, 838–39, to determine which party was liable for infringing the copyright owner's reproduction rights. The district court held that the owner of the server on which such a work is found is the liable. A similar server test has been used by courts to determine jurisdiction issues in cases involving the Internet.

[138] One of the issues raised by agency officials in the past is the lack of adequate attention paid to the collection of personal information over the Internet. *Supra* note 39, at 44 ("Insufficient attention may have been paid to agencies' collection and maintenance of personal information via the Internet and the conformance of these activities with the act's requirements.").

[139] EPIC has obtained copies of the agreements signed by the GSA with various vendors whose software will be available via Apps.Gov and none of the agreements mention privacy policies.

[140] P3P: The Platform for Privacy Preferences, http://www.w3.org/P3P/ (last visited Mar. 30, 2010) (on file with the North Carolina Journal of Law & Technology).

[141] World-Wide Web Consortium (W3C), http://www.w3.org/ (last visited Mar. 30, 2010) (on file with the North Carolina Journal of Law & Technology).

has developed programming language specifications that will allow a website owner to embed its privacy practices into the website and be retrievable in a form that visitors can interpret easily. By giving the public greater choice in assessing whether and what information they want to share at the point at which information may be collected, agencies are relieved of some of the decision making burden, in turn reducing the need for oversight.[142]

IV. CONCLUSION

As new business models begin to view information technology less as a strategic asset and more like a metered utility, computing resources will increasingly be supplied by large cloud providers such as Google. The Federal Government's foray into cloud computing with Apps.Gov presents new challenges to agency officials responsible for protecting against government intrusiveness. The disembodiment of computing services from the underlying hardware by virtualization technology will challenge privacy officers when conducting PIAs. Guidance on the technical implementation details of these important pieces of privacy legislation being unavailable, these officials will have to exercise their judgment in evaluating the degree of oversight that is appropriate for a technology which will dramatically shift our basic conception of computing. It will require rethinking how provisions of the Privacy and E-Gov Act are interpreted and implemented in relation to this new technology. The effective use of PIAs, especially, will only be possible if agencies can effectively monitor the boundaries created by virtualization software.

[142] However, this subverts the purpose of legislation that requires public information to assist the decision-making process of executive branch agencies.