KEEPING THE INTERNET
OPEN • INNOVATIVE • FREE

www.cdt.org

CENTER FOR DEMOCRACY
& TECHNOLOGY

1634 I Street, NW
Suite 1100
Washington, DC 20006

P +1-202-637-9800
F +1-202-637-0968
E info@cdt.org

Statement of **John B. Morris, Jr.**
General Counsel

Center for Democracy & Technology

before the House Committee on the Judiciary,
Subcommittee on Crime, Terrorism and Homeland Security

Hearing on "DATA RETENTION AS A TOOL FOR INVESTIGATING INTERNET CHILD PORNOGRAPHY AND OTHER INTERNET CRIMES"

January 25, 2011

Chairman Sensenbrenner, Ranking Member Scott, and Members of the Subcommittee:

On behalf of the Center for Democracy & Technology (CDT),[1] I thank you for the opportunity to testify today on data retention in the context of child pornography investigations.

CDT strongly agrees with the Subcommittee that child pornography is a horrific crime, and we have long supported increasing the resources available for its prosecution. The organization has spent extensive time examining the challenges raised by child pornography and seeking ways to fight this crime that are consistent civil liberties, and with openness, competition, and innovation on the Internet.

Mandatory data retention raises serious privacy and free speech concerns, and would also harm innovation and competition in the online context. We urge this Subcommittee to carefully consider the significant risks posed by a data retention mandate. Congress has already enacted strong data *preservation* requirements, which have proven to be effective tools for combating child pornography, without the panoply of problems raised by data *retention*. Mandatory data retention would cause significant harms and would, at the same time, not likely increase the number of child pornographers that this country is able to prosecute and put in prison.

As detailed below, we believe that mandatory data retention:

- Would harm Americans' privacy rights, both vis-à-vis the government as well as private actors. Beyond inappropriate invasion of privacy, data retention would also aggravate the problem of identity theft.

[1] The Center for Democracy & Technology is a non-profit public interest organization dedicated to keeping the Internet open, innovative and free. We have long worked to protect children in the online environment while at the same time also protecting online users' privacy and civil liberties. CDT has offices in Washington, D.C., and San Francisco.

- Would harm Americans' free speech rights and would chill Americans from accessing sensitive content online; and
- Would seriously damage competition and innovation in the Internet industry, and would harm the American industry's ability to compete in the global online market.

A vital alternative to data retention is data preservation, which avoids the risks inherent in data retention. It is, in any event, very unclear that adding a data retention regime would in fact lead to more prosecutions of child pornographers. We urge Congress to fully investigate questions about child pornography investigations before it considers imposing burdensome and costly mandates on American industry that, in turn, harm the civil liberties of American citizens.

Defining Data Retention

A starting point in any discussion of "data retention" must be to identify what is meant by the term. In the narrowest possible definition relevant to this hearing, we understand "data retention" to refer to the retention by Internet Service Providers (ISPs) of records of "IP address allocations" indicating which subscriber was assigned which "IP address" for a particular period of time. An IP address (standing for "Internet Protocol" address) is the unique numeric address (such as, for example, 143.228.146.10) used on the Internet to route communications to their proper destination. For any Internet traffic to reach the right place, it must contain the unique address of the destination computer or server.

For common residential broadband Internet access, each customer's household is assigned an IP address at the time the household turns on its service. This IP address can persist for days or weeks, but it can change (both on a regular schedule and whenever the hardware in the household is turned off or loses power). This use of "dynamic IP addresses" is an efficient and effective way for an ISP to manage its service to customers. A consequence of dynamic IP addresses, however, is that the person who is communicating with a given IP address on one day may not be the person who was using that same IP address last week or last month. To assist law enforcement, the leading ISPs in the United States have voluntarily kept records of these "IP address allocations" so that, for limited periods of time, the ISPs would be able to tell law enforcement who had a given IP address on a particular date and time.

Retention of IP address allocations by ISPs, especially if made mandatory, raises a host of serious policy and economic concerns, as discussed below (and as addressed in the testimony today from the ISP industry). But some data retention proposals have gone much farther. Some have advocated that ISPs monitor and record their users' online activities. Other proposals have suggested that *any* entity that gives temporary, dynamic IP addresses (such as coffee shops or WiFi "hotspots") be required to gather and retain data about their users. And in the Department of Commerce Online Safety and Technology Working Group ("OSTWG") process last year, law enforcement went even further to urge that any online site or service that allows users to communicate (such as blogs, social networks, and e-mail services) be required to track and retain "source data"

about every communication that any users make online.[2] These proposals raise enormous concerns.

It is critical to differentiate data *retention* from data *preservation.* A data retention mandate would affect all users, not just bad actors. By contrast, a far more targeted approach – preserving the data of suspects – can already be found in current law. Section 2703(f) of U.S. Code Title 18 permits law enforcement, without any judicial permission or notice at all, to require an ISP or other service provider to retain data – including IP address and customer identifying information – for as much as 180 days. As discussed more fully below, data preservation orders do not raise the kinds of problems raised by data retention

For the reasons set out below, we urge this Subcommittee to reject calls for mandatory data retention, whether narrow or expansive.

Risks Posed by a Data Retention Mandate

Data retention mandates would pose significant risks to individual liberties and, at the same time, would damage innovation and competition within the technology industry. The Subcommittee should carefully consider the serious costs that would flow from any law mandating that service providers track their customers' Internet usage and retain that data.

Data Retention Laws Would Harm Personal Privacy

Data retention laws threaten personal privacy at a time when the public is justifiably concerned about privacy online. A key to protecting privacy is to minimize the amount of data collected and held by ISPs and online companies in the first place. A data retention law would undermine this important principle. Mandatory data retention laws would require companies to maintain large databases of subscribers' personal information, which would be vulnerable to hackers, accidental disclosure, and government or other third party access, thereby aggravating the identity theft problem and undermining public trust in the Internet. And the longer data is maintained, the more at risk it is to compromise or disclosure. The risk of harm would be even greater if entities that do not now keep data on their customers – such as coffee shops, airports, libraries, and others offering wireless access – were required to keep information on customers who use wireless services. And if companies are forced to collect data on their customers, it is very likely that they would decide to use that data for their own commercial purposes as well.

Proposals to mandate data retention cannot be viewed in a legal vacuum, but rather must be considered in light of the very limited privacy protections that are currently afforded to the data held by service providers. The Electronic Communications Privacy

[2] "Youth Safety on a Living Internet," Online Safety and Technology Working Group (OSTWG), June 4, 2009, at 105. OSTWG was established by Congress in the "Protecting Children in the 21st Century Act," (part of the "Broadband Data Improvement Act',' Pub. L. No. 110-385), available at http://www.ntia.doc.gov/reports/2010/OSTWG_Final_Report_060410.pdf [hereafter "OSTWG Report"]. I served as a member of OSTWG and participated in the drafting of the privacy-focused portion of the data retention section of the final report.

Act (ECPA) was a forward-looking statute when enacted in 1986, specifying standards for law enforcement access to electronic communications and associated data, and affording important privacy protections to subscribers of emerging wireless and Internet technologies. But, as underscored by hearings held last year by the Constitution Subcommittee,[3] technology has advanced dramatically since 1986 and ECPA has been outpaced. The statute has not undergone a significant revision since it was enacted in 1986 – light years ago in Internet time.

Because of the out-dated and inadequate standards, data that might be required to be retained – including data that reveals highly sensitive information – could be obtained by law enforcement with almost no restrictions or limitations. This data is available with a mere subpoena and no notice need be made to the record subject. The legal process would involve no proof of specific facts, no judge, and no opportunity for the subject to object for any reason.

As a result, law enforcement requests for such inadequately protected data can target people who are likely entirely innocent. Were websites and other online services required to retain data on visitors, such information would be subject to a mere subpoena, which could, for example, be issued to require a online site to supply identifying information about every person viewing a particular Web site. Although one could argue that this would be acceptable if the web site contained child pornography, the problem is that a data retention mandate might apply to all online sites, including sites that provide sensitive or controversial – but completely lawful – content.

Not only would retained data be at risk of inappropriate and overbroad exposure to the government, but a database of retained data would also serve as a honeypot for lawyers in civil cases. As the OSTWG Report explained, looking back on early data retained by telephone companies, "private litigants soon recognized that [telephone] call record databases contained information that could facilitate investigations and litigation."[4] The exact same thing would happen with online data that might be required to be retained by ISPs, websites, and other online services, except that the online information can be dramatically more sensitive than the record of a phone call. Already, we understand that the great majority of requests that ISPs and others receive for customer information come not from the government but from private litigants in divorce cases, copyright enforcement actions, and commercial lawsuits. A data retention law would aggravate this problem, and would increase the likelihood that whistleblowers and journalists would also be among those whose records were subpoenaed.

Beyond the government and litigant access concerns, there is also a significant risk that service providers, once they were forced to build tracking databases on their customers, would decide to repurpose that data for other uses, such as behavioral advertising. There is bi-partisan interest in improving online users' "baseline" privacy rights relating to commercial uses of data, and it would be important for users to be protected in the context of any data that service providers are mandated to retain.

---

[3] *See* http://judiciary.house.gov/hearings/legislation11.html.

[4] OSTWG Report, at 101.

At a time when there is increasing concern about the privacy and security of personal information, and when there is increasing fear of governmental intrusion into our citizens' personal lives, Congress should be extremely cautious before it imposes a costly and invasive obligation that service providers monitor and track their users.

Data Retention Laws Would Harm Core Free Speech Rights

Data retention laws would threaten a core First Amendment right: the right to speak and access content anonymously. Anonymity fosters public discourse and political debate. Some of our founding fathers – including James Madison, John Jay and Alexander Hamilton – authored the Federalist Papers anonymously, publishing them under the pseudonym "Publius." The leading anti-Federalist also responded anonymously, under the name the "Federal Farmer" – and today we still are not sure of the identity of that political commentator.

The activists and political commentators of today are no more likely to deal in child pornography than those two hundred years ago. But a data retention mandate would sweep in everyone, whether or not they have committed a crime or are engaging in protected political speech that is vital to our society.

The constitutional right to anonymity is well established in our country. In *Talley v. California*, 362 U.S. 60 (1960), the Supreme Court wrote that "[a]nonymous pamphlets, leaflets, brochures and even books have played an important role in the progress of mankind." More recently, in *McIntyre v. Ohio Elections Commission*, 514 U.S. 334 (1995), the Supreme Court reiterated that anonymous speech is part of "an honorable tradition of advocacy and dissent. . . . Anonymity is a shield from the tyranny of the majority." Data retention mandates would harm the ability of commentators and dissenters to express their views anonymously.

The speech harms that would flow from a data retention mandate are not limited to political speech. At least one study has shown that data retention in Europe (which, as discussed more fully below, has a data retention rule that is under attack and is being reconsidered) has significantly diminished citizens' willingness to discuss and obtain information about mental health issues online.[5] This is *precisely* the type of vital speech that would be harmed by a data retention mandate. Congress should not be chilling the discussion of politics, mental health issues, or a vast range of other sensitive topics, when less intrusive tools are already available.

Data Retention Laws Would Harm Innovation and Competition Online

Data retention laws would also harm American consumers – and American businesses, including small businesses – because retention mandates would diminish both competition and innovation in the online context.

---

[5] *See* A.M. Arnbak, "Plenary Presentation on 'Taking on the Data Retention Directive,'" Brussels, Dec. 3, 2010, available at http://www.edri.org/files/Data_Retention_Conference_031210final.pdf (find that as a result of data retention, "half of Germans will not contact marriage counsellors and psychotherapists" via e-mail), citing FORSA, "Opinions of citizens on data retention," June 2, 2008, available at http://www.eco.de/dokumente/20080602_Forsa_VDS_Umfrage.pdf.

A threshold concern is simply one of cost. To our knowledge, ISPs have no business reason to retain IP address allocations. A mandate that all ISPs retain IP address allocations would impose significant costs on those providers. Extending a data retention mandate to the other end of Internet communications – the vast array of large and small online services that allow users to communicate with each other – would be an overwhelming and extraordinarily costly burden. Such a data retention mandate would, without question, drive some providers out of business.

Three scenarios can help illustrate the types, and magnitude, of the business harms that would flow from data retention mandates:

Scenario: Mandate on all online service providers such as e-mail, chat, blogging, and social networking websites to retain "source data" tracking the origins of all user communications: This type of mandate would impose a devastating burden on any website – large or small – that allows users to communicate (and, accordingly, it would certain discourage at least some U.S.-based sites from offering user interaction capability that would trigger the federal mandate). The magnitude of the proposed mandate is breathtaking. As one example, in mid-2009 users on Facebook posted one billion chat messages *per day,*[6] all of which would have to be tracked in a database; a mandate on Facebook alone would likely require that company to add more than *one trillion* entries to a mandated retention database every year.[7] The cost of creating and maintaining such a database would be hard for any company to handle, but a retained data mandate would be especially hard on small and innovative websites seeking to compete with the larger players. Most successful sites on the Internet began as small start-ups and a retention mandate on online companies would certainly chill (or drive offshore) the development of new sites and services.

Scenario: Mandate on any entity that provides Internet access using dynamic IP addresses to retain IP address allocations: Some proposals for data retention have called for mandates on *any* entity that provides access to the Internet to retain dynamic IP address allocations. Yet this type of data retention would burden many small retail businesses and other establishments (such a coffee shops and libraries) that seek to attract customers by offering free wireless Internet access. Such a mandate would also create additional privacy and identity theft risks arising from the mandated storage of personal information by large numbers of retail businesses.[8] And smaller businesses would be particularly hard hit, as they would likely be less able to comply with a federal mandate than would the large national chain shops.

---

[6] *See* "Chat reaches 1 billion messages sent per day," June 15, 2009, at http://www.facebook.com/note.php?note_id=91351698919&id=9445547199.

[7] Facebook's user base has more than doubled since the one billion chat message mark was hit in 2009, and thus it is likely that the chat message count has at least doubled. On top of that, Facebook reports that users post more than a billion other pieces of content to the site each day. *See* "Statistics," at http://www.facebook.com/press/info.php?statistics. Collectively, this equals in the neighborhood of 1.1 trillion separate user communications that Facebook would have to track in a data retention database each year.

[8] The privacy risks cannot be overstated. The small businesses that would be bound by the requirement that they retain significant amounts of personally identifiable data about their customers who access the Internet would become targets of ID thieves, particularly if the businesses lack the sophistication necessary to protect sensitive data. These risks would likely lead many users to decide not to use the Internet services in the first place.

Scenario: Mandate on small ISPs to retain IP address allocations: Today, in an effort to assist law enforcement, the leading broadband ISPs voluntarily retain IP address allocations for limited periods of time. But there still are smaller ISPs, competing with the major ISPs, and most of those ISPs do not have (and could not afford to maintain) tracking databases and the 24/7 law enforcement response offices that larger ISPs operate. Some of these ISPs are small businesses that might be driven from business by an additional federal mandate to retain data.

Any data retention law would be burdensome and costly, requiring investments in storage equipment and design costs, and forcing service providers to incur large annual operating costs. Currently, Internet access is relatively affordable and therefore available to many. The costs associated with mandated data retention would be passed on to consumers, inhibiting efforts to expand Internet access. For online services – many of which are currently free – data retention costs could draw sites' business models into question, or lead companies to seek ways to monetize the data they are forced to collect (by selling it, for example, to behavioral advertising firms).

And, by increasing costs on a broad range of service providers, data retention mandates would reduce competition in Internet access and online services. This reduced competition would likely lead to higher costs and less innovation. At the end of the day, data retention mandates would entrench larger providers, to the detriment of innovators and users.

By increasing costs on Internet access and online services, data retention mandates would harm American businesses, and they would likely drive services overseas to markets that do not have burdensome "source data" retention mandates. The United States has been the leading engine of innovation on the Internet, but costly federal mandates could make this country unfriendly to innovation and new services. Exciting new online services would still be developed – but not as frequently in the United States.

Moreover, a "source data" mandate would be devastating to the American industry's ability to compete in the burgeoning global "cloud computing" market. Few foreign corporations would trust American providers if they were required by the U.S. government to monitor and record data about every communication made over the cloud computing service. Indeed, it is possible that laws in foreign countries would prohibit their companies from using U.S.-based services subject to a data retention mandate.

Congress should reject calls for burdensome federal mandates on a range of service providers to track and retain data on their customers.

Data *Preservation* is an Appropriate Alternative to Data Retention

As noted above, there is an important alternative to data retention: data preservation orders. Current law already allows holding data about *criminal suspects* (rather than retaining data on *all* users of any given system).[9] It permits law enforcement (or any

[9] 18 U.S.C. § 2703(f), Requirement to preserve evidence, provides:

> **(1) In general.** – A provider of wire or electronic communication services or a remote computing service, upon request of a government entity, shall take all necessary steps to preserve records and other evidence in its possession pending the issuance of a court order o other process.

other governmental entity), without any judicial permission or notice at all, to require both ISPs and online service providers to retain data – including IP address and customer identifying information – for 90 days (with an additional 90 days available on request). No supervisory approval is required, nor is any finding (even within the requesting agency) of specific facts that the records to be preserved are relevant to an investigation.

In the child pornography context, data preservation is *automatic* in cases where service providers report possible child pornography to the National Center for Missing and Exploited Children (NCMEC).[10] Whenever a provider sends a child pornography report to NCMEC, the provider must automatically preserve the data to give law enforcement enough time to open an investigation (and, if appropriate, obtain lawful process to demand the preserved data).

From a privacy and civil liberties perspective, the benefits of this approach are enormous: data about only the tiny fraction of individuals who have fallen under criminal suspicion is subject to a data preservation requirement. Everyone else would continue to enjoy the same level of privacy he or she would otherwise enjoy regardless of the law enforcement investigation. Under a data preservation regime, service providers can focus their attention and scarce resources on competition and innovation, rather than building tracking databases full of customer information.

Some countries have rejected data retention mandates in favor of the data preservation approach taken to date in the U.S. In November 2010, the Canadian Department of Justice called for new investigative tools – including data preservation authority – and it specifically rejected data retention because of its overbroad impact.[11] In Europe, many countries and courts are backing away from the European Union data retention mandates that were enacted (but not fully implemented) a few years ago. At least three national courts have questioned the validity of a data retention regime,[12] and another (the Irish High Court) has referred to the European Court of Justice a case that could call into question the validity of the entire European data retention scheme itself.[13]

For all of the reasons discussed in this testimony, data preservation is far preferable to a blanket mandate that extensive data on *all* users should be retained.

---

(2) Period of retention. – Records referred to in paragraph (1) shall be retained for a period of 90 days, which shall be extended for an additional 90-day periods upon a renewed request by the governmental entity.

[10] *See* 18 U.S.C. § 2258A(h).

[11] *See* "Backgrounder: Investigative Powers for the 21st Century Act," Nov. 2010, at http://www.justice.gc.ca/eng/news-nouv/nr-cp/2010/doc_32567.html.

[12] For example, the German Constitutional Court found that aspects of the German retention law violated the fundamental right to privacy. *See* Vorratsdatenspeicherung Bundesverfassungsgericht, 2 March 2010, 1 BvR 256/08. The Romanian Constitutional Court went further and invalidated general mandatory data retention as a violation of the Romanian Constitution and EU law. *See* Decision no.1258, Romanian Constitutional Court, 8 October 2009. Unofficial translation by Bogdan Manolea and Anca Argesiu at http://www.legi-internet.ro/fileadmin/editor_folder/pdf/decisionconstitutional-court-romania-data-retention.pdf. *See also* http://www.edri.org/edrigram/number7.20/romania-data-retention-law-unconstitutional. *See also* EDRi, "Bulgarian Court Annuls a Vague Article of the Data Retention Law," EDRi-gram No. 6.24, December 17, 2008, http://www.edri.org/edri-gram/number6.24/bulgarian-administrative-case-data-retention.

[13] Digital Rights Ireland v. Minister for Communications and others, No. 2006/3785P §108. *See also* EDRi, "Irish Court Allows Data Retention Law to be Challenged in ECJ," EDRi-gram No. 8.10, May 19 2010, http://www.edri.org/edrigram/number8.10/data-retention-ireland-ecj.

In the Face of the Serious Risks and Costs of Data Retention, Congress Should Carefully Investigate What Benefits There Would Be, If Any, in the Prosecution of Child Pornography Cases

Although no data is publicly available, and law enforcement officials have consistently refused to release information of this type when asked, a common perception in the child safety world is that law enforcement agencies already have far more child pornography cases on their plates than they can investigate and prosecute. In other words, even if a vast data retention regime were imposed on the American Internet industry, and even if data were retained for a lengthy period of time, law enforcement agencies would *still* not be able to investigate and prosecute more child pornography cases.

Moreover, it appears that many of the cases that are not being pursued because of a lack of resources are very recent reports (not older cases). In 2008, in testimony before the Senate Judiciary Committee, Special Agent Flint Waters of the Wyoming Internet Crimes Against Children Task Force testified about a broad range of *very current* cases that were not being pursued because of a lack of law enforcement resources.[14]

Congress recognized in 2008 the critical problem of a lack of resources to investigate child pornography cases, and it responded by authorizing for appropriation an additional $300 million (over five years) aimed at increasing prosecution of child pornographers. Unfortunately, to our knowledge, none of these funds have actually been appropriated.

In light of the continuing critical lack of resources to prosecute child pornography cases, and in light of all of the problems raised by data retention as detailed above, Congress should not impose huge costs on the Internet industry to implement a data retention regime.

As part of the OSTWG process, one of the OSTWG subcommittees addressing child pornography discussed the important need for Congress – before it takes additional action in this area – to learn the critical facts about the timing of and resources available to the investigation and prosecution of child pornography cases. As an Addendum to the OSTWG report, *see* OSTWG Report at 92-94,[15] I suggested a detailed series of questions that Congress should ask to inform any further policy decisions. We urge this Subcommittee to review those questions, and obtain answers to them.

Conclusion

Mandatory data retention is a risky and costly path to go down, and one that is all the more problematic because once Congress opens the door to mandating that service providers amass huge tracking databases documenting citizens' Internet usage, it will be hard to close it. If Congress were to impose data retention on even just a narrow category of service providers, and even for a narrow category of crimes, there would be

---

[14] Waters wanted to emphasize that he was not criticizing law enforcement: "I would like to be clear, I am NOT saying law enforcement isn't doing enough with what they have. I am saying they could do so much more if they only had the resources." *See* Testimony of Special Agent Flint Waters before the Senate Committee on the Judiciary Subcommittee on Crime and Drugs, April 16, 2008, at http://judiciary.senate.gov/pdf/08-04-16WatersTestimony.pdf.

[15] *See* OSTWG Report, http://www.ntia.doc.gov/reports/2010/OSTWG_Final_Report_060410.pdf, at 92-94.

a strong and inevitable push to broaden the scope and reach of data retention. Congress should not cross this risky line.

CDT appreciates the opportunity to testify today and we look forward to working with the Subcommittee on these issues.

For more information, contact John Morris, jmorris@cdt.org, Greg Nojeim, gnojeim@cdt.org, or Jim Dempsey, jdempsey@cdt.org, or at (202) 637-9800.