Testimony of Dr. Alma Whitten, Privacy Engineering Lead, Google Inc.
Senate Committee on Commerce, Science, and Transportation
Hearing on Consumer Online Privacy
July 27, 2010

Chairman Rockefeller, Ranking Member Hutchison, and members of the Committee:

I am pleased to appear before you this afternoon to discuss online privacy and the ways that Google protects our users' personal information. My name is Dr. Alma Whitten, and I am Google's Privacy Engineering Lead. I am responsible for a team of dedicated privacy and security engineers who develop and improve Google's privacy tools, like our Dashboard, and work with our other engineers and product teams to build transparency, user control, and security into Google's products.

Google is most well known for our search engine, which is available to Internet users throughout the world. We also offer dozens of other popular services, from YouTube to Gmail to Google Earth. Our products are free to individuals for personal use, supported by revenue from online advertising.

While our users benefit from our free services, Google's innovative advertising system is also helping businesses grow in a challenging economic time. In 2009 alone, our advertising products generated a total of $54 billion of economic activity for American businesses, website publishers, and non-profits. This number only covers economic activity generated by Google's search and advertising tools, including the over $5 billion of revenue we generate for online publishers in 2009. It does not include the positive economic impact of products like Gmail and others that allow consumers, entrepreneurs, and businesses of all sizes to communicate and collaborate for free – or, in the case of enterprise customers, at a lower cost than alternative platforms.

Our recent economic impact report (google.com/economicimpact) explains Google's contribution to the American economy, and features small businesses that rely on Google's advertising products to reach customers and generate revenue.

One example is OVIS, a 20 year-old cabinet hardware and woodworking supplier based in Millwood, West Virginia (www.ovisonline.com). OVIS's owner Chip Wimbauer told us that Google's advertising system is "the best way for a small business to compete and look like a big company," and that with online advertising OVIS has gone from a regional company to one that does as much business in Hawaii as it does within West Virginia. In Texas last year we created over $3 billion in economic value for over 100,000 advertisers and online publishers. And we donated almost $3 million in advertising to non-profit groups like the American Heart Association and the

Susan G. Komen Breast Cancer Foundation through our Google Grants program (information about which is available at www.google.com/grants). These types of success stories happen in every state in partnership with hundreds of thousands of businesses and numerous not-for-profit organizations.

In a time of tight budgets, we're glad to help so many small businesses and entrepreneurs find customers more efficiently and increase revenue through relevant advertising. We also take pride in building trust with users, and privacy is a core part of that effort.

At Google, privacy is something we think about every day across every level of our company. We make this effort because privacy is both good for our users and critical for our business. If we fail to offer clear, usable privacy controls and strong security, our users will simply leave. This is the basic truth that guides me in my job as Google's lead privacy engineer.

In my testimony today, I'm going to talk about three topics:

First, I'd like to discuss how Google's approach to privacy manifests itself in our products. In other words, how do we put our privacy principles into executable code? I'll provide several examples to give the Committee a tangible sense of the considerations that go into designing privacy as part of our products and the transparency, control, and security that are built into Google's products.

Second, I will discuss the challenges companies like Google face when designing for privacy and security. How do we harness the power and value of data for our users while protecting against privacy harms? How can we communicate about evolving data practices and controls to users in a meaningful way?

Third, while I'm far from a legal expert, I'll offer a bit of thought as to how Congress can help protect consumers and improve user trust in data-intensive services – including through the development of comprehensive, baseline privacy rules.

How we approach privacy at Google

When I think about privacy at Google, I start with our five privacy principles. In brief, these are:

- Use information to provide our users with valuable products and services.
- Develop products that reflect strong privacy standards and practices.
- Make the collection and use of personal information transparent.
- Give users meaningful choices to protect their privacy.
- Be a responsible steward of the information we hold.

The principles are located at www.google.com/corporate/privacy_principles.html.

Let's break these down a bit. As with every aspect of our product, we follow the motto of "Focus on the user and all else will follow." We are committing ourselves to use information only where we can provide value to our users. That's what we mean by our first principle.

For instance, **we do not sell our users' personal information**.

To further guide us, under the second principle, we aim to build privacy and security into our products and practices from the ground up. From the design phase through launch we are considering a product's impact on privacy. And we don't stop at launch – we continue to innovate and iterate as we learn more from users.

Our last three principles give substance to what we mean by privacy: We commit to *transparency*, *user control,* and *security*.

We work hard to embed privacy considerations into our culture through our principles and in the way we're organized. As Google's Privacy Engineering Lead, I'm only one of many individuals at the company who work on privacy from every angle – including technology, products, policy, and compliance initiatives. This cross-functional team, all focused on our users' privacy interests, ensures that privacy doesn't exist as a silo within Google. For example, our Privacy Council, a cross-functional group of Google employees, helps us identify and address potential privacy issues across all our products.

In just the last 18 months, we have been tackling four broad privacy issues that face our industry in a way that is consistent with our principles:

- Transparency and control in the online advertising ecosystem.
- Easy data portability for information that is processed and stored by Google.
- A comprehensive and useful dashboard of privacy and account controls for a suite of web services.
- Strong security for users of Google's services, like Gmail and Google Search.

In the next section of my testimony I'll discuss these privacy issues and illustrate how Google works to bring transparency, user control, and security to its users.

Transparency and control for interest-based advertising

The availability of Google Search and our other products – and the improvements that we make to our products on a daily basis – is funded by online advertising: by far our primary source of revenue. As we work to bring more relevant ads to our users, we continually seek to preserve transparency and user control over the information used in our ad system.

Google was not the first to offer interest-based advertising (known as IBA) online, but it was important to us that we offer clear and strong privacy controls before introducing this product. When we launched IBA, in March 2009, we included a number of groundbreaking privacy features. As Google tells its users:

> Many websites, such as news sites and blogs, use Google's AdSense program to show ads on their sites. It's our goal to make these ads as relevant as possible for you. While we often show you ads based on the content of the page you are viewing, we also developed new technology that shows some ads based on interest categories that you might find useful.

Google's interest-based ads contain notice in the actual advertisement indicating that it is a Google ad. The in-ad notice is linked to information about IBA, including our Ads Preferences Manager, which allows users to change the interest categories used to target ads or to opt-out of interest-based advertising altogether. Note that we use only non-personally-identifiable data for IBA targeting.

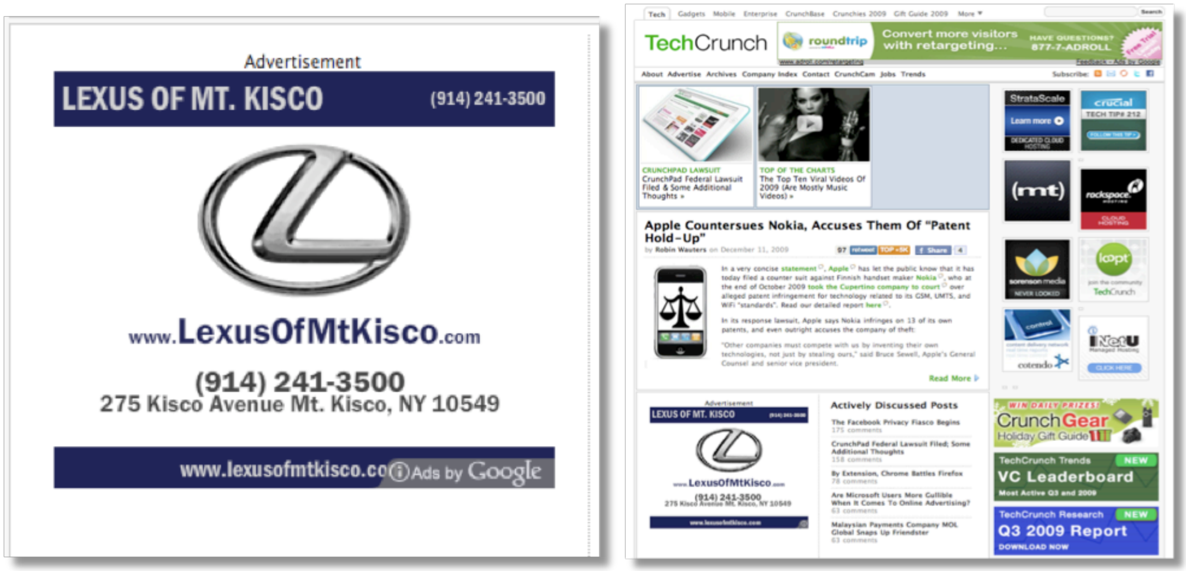
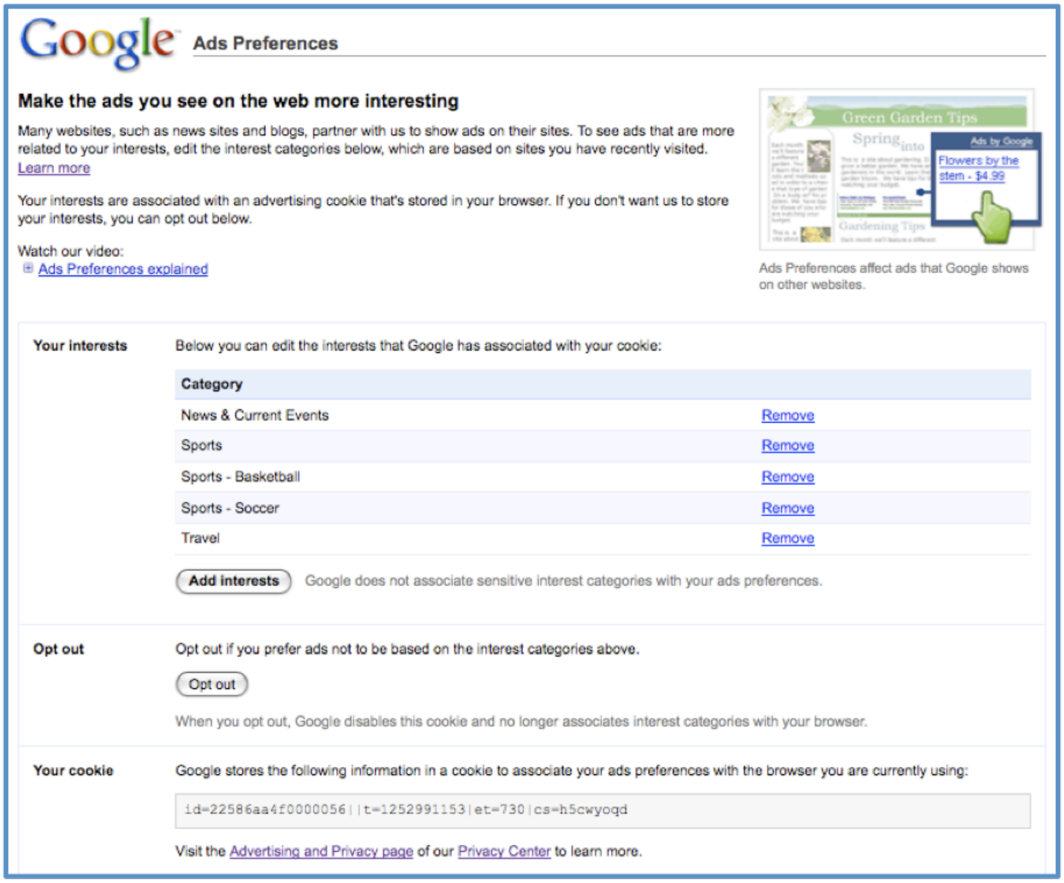


Fig. 1: Sample advertisement with in-ad privacy notice

With the launch of our Ads Preferences Manager (www.google.com/ads/preferences), Google became the first major industry player to empower users to review and edit the interest categories we use to target ads. The Ads Preferences Manager enables a user to see the interest categories Google associates with the cookie stored on her browser, to add interest categories that are relevant to her, and to delete any interest categories that do not apply or that she does not wish to be associated with.

I should also clarify that Google does not serve interest-based ads based on sensitive interest categories such as health status or categories relating to children under 13.

The Ads Preferences Manager also permits users to opt out of interest-based ads altogether. Google implements this opt-out preference by setting an opt-out cookie that has the text "OPTOUT" where a unique cookie ID would otherwise be set. We have also developed tools to make our opt-out cookie permanent, even when users clear other cookies from their browser (see www.google.com/ads/preferences/plugin). We are encouraged that others are using the open-source code for this plugin, released by Google, to create their own persistent opt-out tools.



Ads Preferences

Make the ads you see on the web more interesting

Many websites, such as news sites and blogs, partner with us to show ads on their sites. To see ads that are more related to your interests, edit the interest categories below, which are based on sites you have recently visited. Learn more

Your interests are associated with an advertising cookie that's stored in your browser. If you don't want us to store your interests, you can opt out below.

Watch our video:
⊞ Ads Preferences explained

Ads Preferences affect ads that Google shows on other websites.

Your interests

Below you can edit the interests that Google has associated with your cookie:

Category	
News & Current Events	Remove
Sports	Remove
Sports - Basketball	Remove
Sports - Soccer	Remove
Travel	Remove

Add interests — Google does not associate sensitive interest categories with your ads preferences.

Opt out

Opt out if you prefer ads not to be based on the interest categories above.

Opt out

When you opt out, Google disables this cookie and no longer associates interest categories with your browser.

Your cookie

Google stores the following information in a cookie to associate your ads preferences with the browser you are currently using:

```
id=22586aa4f0000056||t=1252991153|et=730|cs=h5cwyoqd
```

Visit the Advertising and Privacy page of our Privacy Center to learn more.

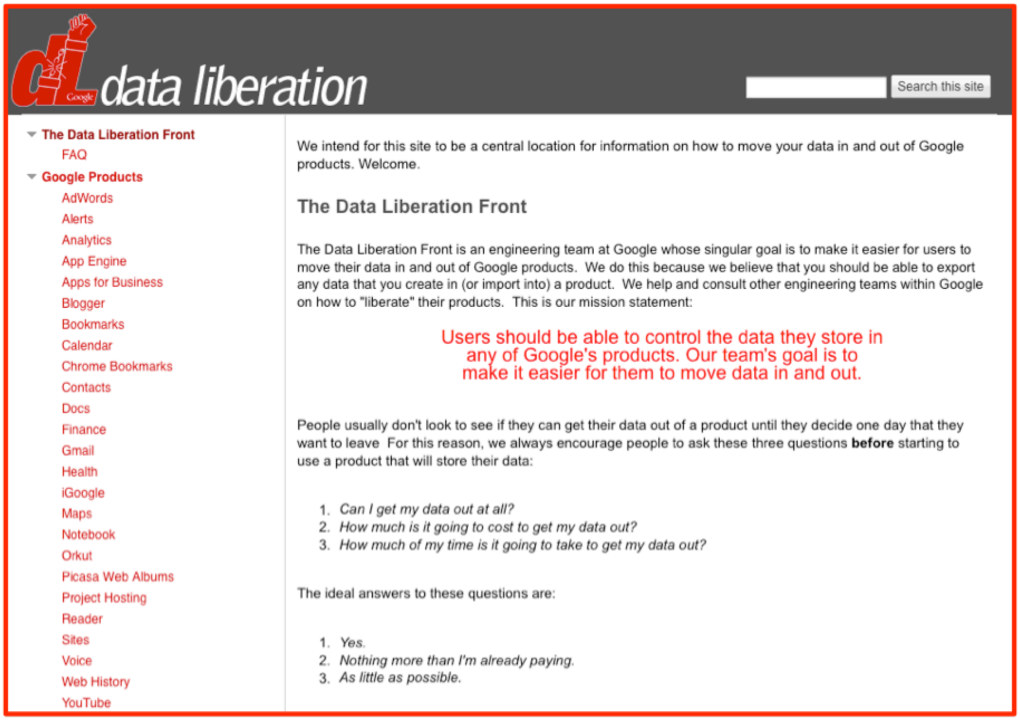
Fig. 2: Ads Preferences Manager

As an engineer, I like to evaluate things by looking at the data. In this case, we have begun to receive information about how users are interacting with the Ads Preferences Manager. While our data are preliminary, we have discovered that, for every user that has opted out, about four change their interest categories and remain opted in and about ten view their settings but do nothing. We take from this that online users appreciate transparency and control, and become more comfortable with data collection and use when we offer it on their terms and in full view.

Control through data portability

Providing our users with control over their personal information must also mean giving them the ability to easily take their data with them if they decide to leave. Starting with our Gmail service and now covering more than 25 Google products where users create and store personal information, a cadre of Google engineers – self-named the "Data Liberation Front" – has built tools to allow our users to "liberate" their data if they choose to switch providers or to stop using one of our services. The critical insight of these engineers was to recognize that users should never feel stuck using a service because they are unable to easily retrieve the content they created and transfer it to another service or provider at no additional cost.

Every user of Gmail, Picasa, Reader, YouTube, Calendar, Apps for Business, Docs, iGoogle, Maps, and many other products already have access to data portability tools, and the team continues to work on additional products. Detailed information for users is available at www.dataliberation.org.



We intend for this site to be a central location for information on how to move your data in and out of Google products. Welcome.

The Data Liberation Front

The Data Liberation Front is an engineering team at Google whose singular goal is to make it easier for users to move their data in and out of Google products. We do this because we believe that you should be able to export any data that you create in (or import into) a product. We help and consult other engineering teams within Google on how to "liberate" their products. This is our mission statement:

Users should be able to control the data they store in any of Google's products. Our team's goal is to make it easier for them to move data in and out.

People usually don't look to see if they can get their data out of a product until they decide one day that they want to leave For this reason, we always encourage people to ask these three questions **before** starting to use a product that will store their data:

1. *Can I get my data out at all?*
2. *How much is it going to cost to get my data out?*
3. *How much of my time is it going to take to get my data out?*

The ideal answers to these questions are:

1. *Yes.*
2. *Nothing more than I'm already paying.*
3. *As little as possible.*

Fig. 3: Data Liberation Front

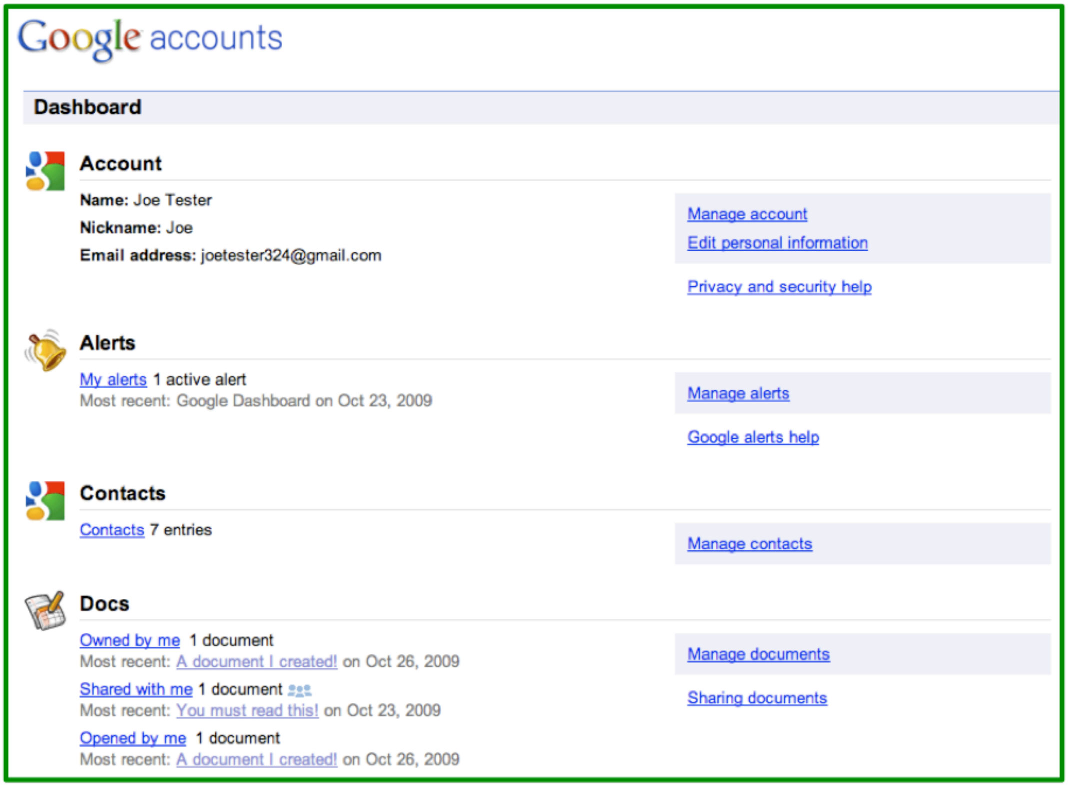
Data portability has benefits for our users and for Google. First, it keeps our product teams on their toes – they know just how easy it is for their users to move to a competitor's product, and understand that their success depends upon continuing to be responsive to privacy and product concerns and acting quickly to address them. Second, allowing our users the freedom to leave honors our commitment to put users in control.

In considering the testimony today and as the Committee develops its approach to consumer privacy, I urge you to consider the role that data portability can play in ensuring that consumer-facing businesses remain accountable for their privacy choices. Regulators should encourage this kind of "user empowerment by design" as an effective means of ensuring respect for user privacy without chilling innovation.

One-stop shop for transparency and control: the Google Dashboard

Google developed the Google Dashboard (www.google.com/dashboard) to provide users with a one-stop, easy-to-use control panel to manage the use and storage of personal information associated with their Google accounts and products – from Gmail to Picasa to Search.

With the Dashboard, a user can see and edit the personally identifiable data stored with her individual Google account. A user also can change her password or password recovery options using Dashboard, and click to manage various products' settings, contacts stored with the account, or documents created or stored through Google Docs. Dashboard also lets a user manage chat data, by choosing whether or not to save it in her Google account.



Google accounts

Dashboard

Account

Name: Joe Tester
Nickname: Joe
Email address: joetester324@gmail.com

Manage account
Edit personal information
Privacy and security help

Alerts

My alerts 1 active alert
Most recent: Google Dashboard on Oct 23, 2009

Manage alerts
Google alerts help

Contacts

Contacts 7 entries

Manage contacts

Docs

Owned by me 1 document
Most recent: A document I created! on Oct 26, 2009
Shared with me 1 document
Most recent: You must read this! on Oct 23, 2009
Opened by me 1 document
Most recent: A document I created! on Oct 26, 2009

Manage documents
Sharing documents

Fig. 4: Google Dashboard

Industry-leading security: encrypted Search and Gmail

Along with transparency and user control, good security is vital in maintaining user trust. Google faces complex security challenges while providing services to millions of people every day, and we have world-class engineers working at Google to help secure information. In fact, my own research background is in security. In a 1999 paper, "Why Johnny Can't Encrypt," I argued that security tools must be simple and usable to be effective. Unfortunately, it is sometimes the case that security technology is so complicated that it isn't usable, and thus ineffective. I have continued that theme at Google, working to build user-friendly, simple security features into our products.

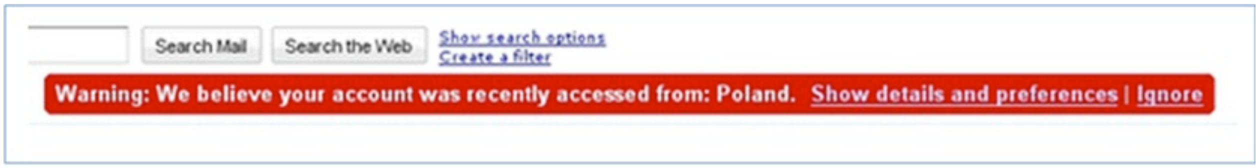
For example, Google recently became the first (and still only) major webmail provider to offer session-wide secure socket layer (SSL) encryption *by default*. Usually recognized by a web address starting with "https" or by a "lock" icon, SSL encryption is regularly used for online banking or transactions. As our Gmail lead engineer wrote:

> In 2008, we rolled out the option to always use https – encrypting your mail as it travels between your web browser and our servers. Using https helps protect data from being snooped by third parties . . . . We initially left the choice of using it up to you because there's a downside: https can make your mail slower since encrypted data doesn't travel across the web as quickly as unencrypted data. Over the last few months, we've been researching the security/latency tradeoff and decided that turning https on for everyone was the right thing to do.

We hope other companies will soon join our lead.

We also hope to see our competitors adopt another security tool we offer our users: encryption for search queries. Users can simply type in "encrypted.google.com" and encrypt their search queries and results. As we said in our blog post about encrypted search, "an encrypted connection is created between your browser and Google. This secured channel helps protect your search terms and your search results pages from being intercepted by a third party on your network."

And in March Google launched a system to notify users about suspicious activities associated with their accounts. By automatically matching a user's IP address to broad geographical locations, Google can help detect anomalous behavior, such as a log-in appearing to come from one continent only a few hours after the same account holder logged in from a different continent. Thus, someone whose Gmail account may have been compromised will be notified and given the opportunity to change her password, protecting her own account and her Gmail contacts.



Warning: We believe your account was recently accessed from: Poland. Show details and preferences | Ignore

Fig. 5: Recent Account Activity Warning

Similarly, we built Google Chrome with security in mind from the beginning, including features such as:

- Safe Browsing, which warns a user before he visits a site that it is suspected of phishing or containing malware;
- Sandboxing, which works automatically to help prevent web browser processes from harming one another or a user's computer, and
- Automatic updates that deliver security patches to users as quickly as possible.

Google also conducts extensive security research and provides free security resources to the broader Internet community. We make security tools available for free to webmasters to help them operate more secure sites, as well as to application developers to help them build more secure applications. For example, we recently released a tool called "skipfish" under an open source license to help identify web application vulnerabilities through fully automated, active security reconnaissance.

The challenges of designing for privacy and security

In addition to discussing Google's efforts to offer transparency, user control, and security, I want to also discuss just two of the many challenges I and others in similar roles face as we try to build privacy and security into innovative products. The first relates to data collection and use. The second involves how to best communicate to individuals how to manage their privacy.

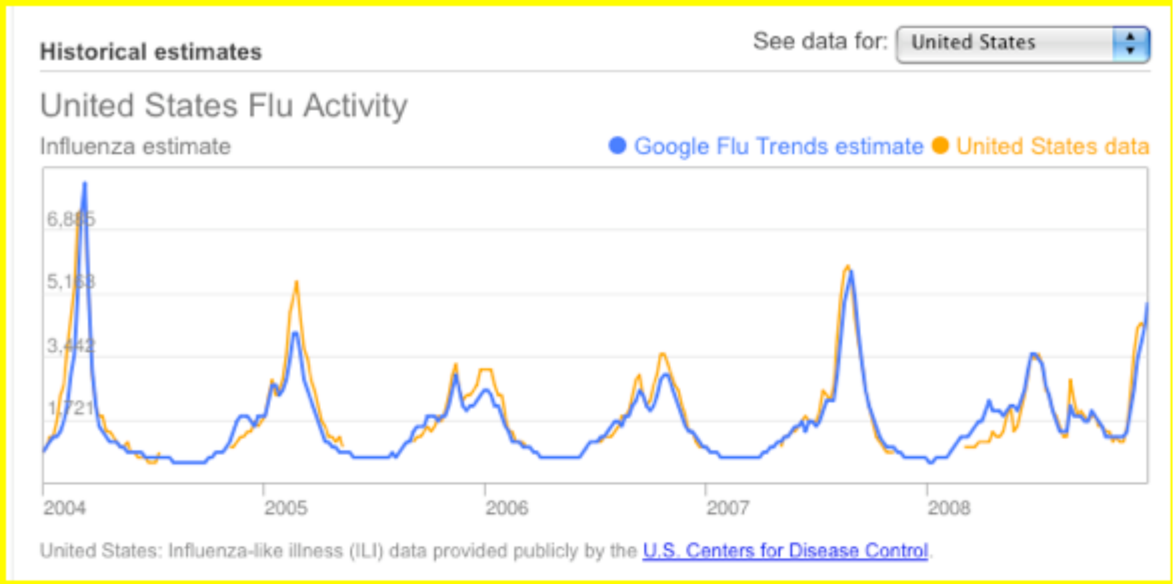
Every day we receive information from our users' interaction with our products and services. That information may be in the form of an email that we process, store, and protect in our Gmail product – or it could be generated by the interaction between a user's computer and our servers, such as a search query and the IP address associated with a specific computer or network of computers.

We are asked often why we retain this query and IP address data – which can be very sensitive even if it does not personally identify individuals. We certainly treat this data with strong security, and seek to build in transparency and user controls where appropriate – including tools like our Ads Preferences Manager. We also voluntarily anonymize IP addresses after nine months.

But this data is actually tremendously helpful to us in improving our products and protecting our networks from hackers, spammers, and fraudsters. For example, bad actors continually seek to manipulate our search ranking, launch denial-of-service attacks, and scam our users via email spam or malware. We use our log files to track, block, and keep ahead of the bad guys.

We also use information like IP addresses and search queries to develop products like Flu Trends (www.google.com/flutrends). A team of our engineers found that examining certain search terms on an aggregate basis can provide a good indicator of flu activity. Of course, not every person who searches for "flu" is actually sick, but a pattern emerges when many flu-related search queries are added together. By counting how often we see these search queries, we can estimate how much flu is circulating in different countries and regions around the world. Our results have been published in the journal *Nature*.

For epidemiologists, this is an exciting development, because early detection of a disease outbreak can reduce the number of people affected. If a new strain of influenza virus emerges under certain conditions, a pandemic could ensue with the potential to cause millions of deaths. Our up-to-date influenza estimates may enable public health officials and health professionals to better respond to seasonal epidemics and pandemics.



United States: Influenza-like illness (ILI) data provided publicly by the U.S. Centers for Disease Control

Fig. 6: Google Flu Trends

A second challenge is how to best communicate with our users about privacy.

At Google, we take great pride in our effort to provide our users with a better understanding of how we collect, use, and protect their data. For example, we have produced a series of short videos on privacy and made the videos available at Google.com and on YouTube. We also blog often about privacy in plain language aimed at educating our users. We believe that companies that interact and have relationships with consumers need to do more than simply provide and link to privacy policies;

we all need to offer consumer-friendly materials in a variety of media to help users better understand how their information is collected and used, and what choices they have to protect their privacy.

We also believe in "transparency in context" so that consumers can benefit from privacy information *when and where they're actually using a product or service*, in addition to through a privacy policy. The concept of transparency in context underlies our desire to provide in-ad notice for interest-based ads. With such notice, consumers have easy access to both information and choice tools at the point of interaction with the relevant product.

There are times, of course, where we do not get it right on the first try. When we launched Google Buzz, a social networking service for sharing updates, photos, videos, and more, we heard from some users that they were unhappy. So our engineers worked around the clock and within 48 hours we had made significant product changes. Now, instead of automatically creating a list of followers, we suggest people for Buzz users to follow. We also made it easier for users to block others from following them. And we added a tab to Gmail settings making it easier to hide Buzz or disable it completely. Soon after, we sent out a confirmation page to early Buzz users giving them another opportunity to understand and reconfirm their settings.

These are the kind of updates and improvements we are making to all our products all the time, from Gmail to search to mobile, because control is what our users want and deserve – and what we want to provide.

Understanding the WiFi incident

In those instances where mistakes occur, we try to understand and learn from our mistakes. I'd like to address the recent issue involving WiFi data in that context.

Several months ago, Google disclosed that we had mistakenly included code in the software on our Street View cars that collected samples of WiFi "payload data" – information sent over a WiFi network – from open (unencrypted) WiFi networks. Importantly, these samples of payload data have never been used in any Google product or service; nor do we intend to use them. If you would like more information about the facts and background of this incident, including the independent, third-party review of our software, my colleague Alan Eustace has described it on the Official Google Blog.

As Alan concluded, "We are profoundly sorry for this error and are determined to learn all the lessons we can from our mistake." While our legal team is still reviewing the matter, I can attest that it was not consistent with the value we place on the responsible handling of personal data. Google is taking the review of this matter very seriously and we will report back with the changes we'll make to prevent such a thing from happening in the future.

The incident also reaffirms to us the importance of transparency. Data collection and use practices should be disclosed, and in plain language. When mistakes occur, companies ought to continue providing that transparency – as Google did here even in the absence of any breach of personal data – by quickly and simply disclosing what occurred, any risk posed to users, and how users can mitigate that risk.

How Congress can encourage responsible privacy practices and build trust

Congress has a vital role to play in encouraging responsible privacy and security practices, both by bringing attention to these issues and through appropriate legislation. Google supports the development of comprehensive, baseline privacy legislation that can ensure broad-based user trust and that will support continued innovation and serve the privacy interests of consumers.

I am a scientist and engineer, not a lawyer, but I have some basic thoughts about what good policy needs to accomplish in this area.

- **Even-handed application.** A pro-innovation privacy framework must apply even-handedly to all personal data regardless of source or means of collection. Thus, offline data collection and processing should, where reasonable, involve similar data protection obligations.

- **Recognition of benefits and costs.** As with any regulatory policy, it is appropriate to examine the benefits and costs of legislating in this area, including explicit attention to actual harm and compliance costs.

- **Security requirements and breach notification.** We pride ourselves at Google for industry-leading security features, including the use of encryption for our search and Gmail services I discussed. A thorough privacy framework should promote uniform, reasonable security principles, including data breach notification procedures.

- **Clear process for compelled access.** The U.S. law governing government access to stored communications is outdated and out of step with what is reasonably expected by those who use cloud computing services. The problems in the law threaten the growth, adoption, and innovation of cloud technologies without a corresponding benefit. As part of the Digital Due Process coalition, we are working to address this issue. The Committee can play an important role in encouraging clear rules for compelled access to user data.

- **Consistency across jurisdictions**. Generally, Internet users neither expect nor want different baseline privacy rules based on the local jurisdiction in which they or the provider reside. Moreover, in many instances, strict compliance with differing state or national privacy protocols would actually diminish consumer privacy, since it would require Internet companies to know where consumers are located at any given time.

Any new privacy law must also offer baseline protections on which providers can innovate. A pro-innovation privacy framework offers providers the flexibility to both develop self-regulatory structures and individually innovate in privacy practices and tools. The advertising industry and online publisher efforts to develop self-regulatory rules for interest-based advertising, for example, are a strong example of the need for and utility of industry-driven efforts. As I have discussed, Google has been a leader in developing innovative privacy tools.

Continued innovation in the privacy space is vital for users. Unfortunately, compliance-based or overly complex rules can lock in a specific privacy model that may quickly become obsolete or insufficient due to the speed with which Internet services evolve. A principles-based model encourages innovation and competition in privacy tools.

A baseline framework needs to encourage the development of innovative tools like the ones I've described. We believe that stable, baseline principles set by law can permit flexible, adaptive structures to develop on top – much like the stable protocols and standards at the physical and network layers of the Internet allow flexible and innovative development at the content and application layers. With comprehensive, baseline privacy legislation establishing ground rules for all entities, self-regulatory standards and best practices of responsible industry actors will evolve over time. On top of that structure, individual companies will be free (and encouraged) to create innovative privacy tools and policies rather than stick with potentially outdated compliance structures.

Conclusion

Chairman Rockefeller, Ranking Member Hutchison, and members of the Committee, thank you for inviting me to testify today. We at Google appreciate the opportunity to discuss online privacy and how our company has helped lead in the effort to protect our users by providing them with transparency, user control, and security.

I look forward to answering any questions you might have about our efforts, and Google looks forward to working with members of the Committee and others in the development of better privacy protections.

Thank you.