10 June 2007

Mr. Peter Schaar
Chairman, Article 29 Data Protection Working Party
B-1049 Brussels, Belgium
Office No LX-46 01/43
Sent via Email

Dear Mr Schaar,

Thank you for your letter of May 16, 2007. Google is committed to raising the bar on our own privacy practices for the benefit of Google users. We're also committed to engaging in a constructive dialogue with the Article 29 Working Party and other leading privacy stakeholders around the world in order to raise the bar on privacy practices across the Internet.

We appreciate that the Working Party views our recent announcement to anonymize server logs after 18-24 months as a positive step. We have engaged in long and serious reflection about the balance that Google, and companies like ours, must strike between competing principles: privacy, security, innovation, and various legal retention obligations. We welcome your questions about our recent decision and feel that these kinds of questions contribute to the broader debate about how long search and other Internet companies should retain data. These are very hard questions involving many different factors and implicating many different stakeholders. That's why we're publishing this letter on our blog.

Google provides one level of privacy protection for our users worldwide, irrespective of the country where they reside (although of course we do comply with applicable law). Moreover, it's extraordinarily difficult to operate a global Internet service according to different privacy standards in different countries. Thus, the discussion regarding the right retention period is in fact a global discussion. Google is a U.S. company and we respect U.S. laws -- but we are also a global company, doing business across Europe and across the world, and we recognize the need to respect the laws of the countries in which we do business. We are therefore committed to data protection principles that meet the expectations of our users in Europe and across the globe. This commitment includes clear privacy policies and absolute transparency about our data retention practices so that users are well-informed about the data we collect when they use our services. We provide information to our users about the data stored in our server logs here in our Privacy FAQ.
There is no single right answer to the question of how long server logs should be retained. In keeping with the principle of proportionality enshrined in Article 6 of the General Data Protection Directive, decisions about data retention are about balance. You have asked us to explain further the factors that guided our decision to anonymize our server logs after 18 to 24 months, and to justify our decision in terms of EU data protection laws. Neither the General Data Protection Directive nor the E-Privacy Directive has set forth any specific periods beyond which personal data may not be

retained. The lack of such set periods indicates that data retention must be determined based on the general data protection principles contained in data protection law. We believe that our decision to anonymize our server logs after 18 to 24 months complies with data protection law, and at the same time allows us to fulfill other critical interests, such as maintaining our ability to continue to improve the quality of our search services; protecting our systems and our users from fraud and abuse; and complying with possible data retention requirements.

In requesting information on the purposes for which it retains server log data, the Working Party cites Article 6(1)(e) of the General Data Protection Directive, which provides that personal data must be "kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed". This provision sets forth a principle of proportionality that judges the legality of retention periods based on the purposes for which the data are collected and processed.

By its very nature, this principle of proportionality cannot lead to a black-or-white answer that applies in all cases, since it is based on an evaluation of the purposes for which personal data are retained in a particular case. We believe that our retention of log data for 18-24 months is proportionate under Article 6(1)(e) in light of the purposes for which the data are retained.

Retention of logs data is critical to our ability to operate and improve our services, and to provide adequate security for our users, as follows:

> – Analyzing log data is an important tool to help our engineers refine search quality and build helpful new services. Take the example of Google Spell Checker. Our spell-checking software automatically looks at your query and checks to see if you are using the most common version of a word's spelling. If it calculates that you're likely to generate more relevant search results with an alternative spelling, it will ask "*Did you mean:* (more common spelling)?" We can offer this service by looking at spelling corrections that people do or do not click on. Similarly, with logs, we can improve our search results: if we know that people are clicking on the #1 result we're doing something right, and if they're hitting next page or reformulating their query, we're doing something wrong. The ability of a search company to continue to improve its services is essential, and represents a normal and expected use of such data.
>
> – Log data is also crucial in helping prevent fraud and abuse. It is standard among Internet companies to retain server logs with IP addresses as one of an array of tools to protect the system from security attacks. For example, our computers can analyze logging patterns in order to identify, investigate and defend against malicious access and exploitation attempts. A failure to retain log data for a sufficient period would make our systems more vulnerable to security attacks, putting the personal data of our users at greater risk. Historical logs information

can also be a useful tool to help us detect and prevent phishing, scripting attacks, and spam, including query click spam and ads click spam. Moreover, log data helps us protect our systems from web and index spam, which in turn supports healthy traffic flow to many web sites on the Internet.

To achieve these purposes, we need to have a sufficient amount of historical log server data. In fact, all search engine companies need sufficient data to evaluate and improve their services based on the needs of users, as online services evolve very rapidly. In addition, there is tremendous growth in fraud on the Internet, posing serious challenges for service providers to keep their services secure. In determining a retention period, we closely examined the evolution of search engine services, and the needs of our engineers to ensure the security of Google services. The period chosen, 18 to 24 months, represents a period lengthy enough to achieve these purposes without being excessive. We therefore believe that this is a proportionate period for the retention of log server data.

In addition to proportionality, data retention policies must also respect the principle of legality set forth in Article 6(1)(a) of the General Data Protection Directive. The Data Retention Directive requires all EU Member States to pass data retention laws by 2009 with retention for periods between 6 and 24 months. Google is therefore potentially subject (both inside and outside the EU) to legal requirements to retain data for a certain period. Since not many Member States have implemented the Directive thus far, it is too early to know the final retention time periods, the jurisdictional impact, and the scope of applicability. Because Google may be subject to the requirements of the Directive in some Member States, under the principle of legality, we have no choice but to be prepared to retain log server data for up to 24 months.

There are many unanswered questions regarding the EU Data Retention Directive. The Working Party has criticized its lack of clarity in many respects, particularly with regard to divergent implementations in each Member State. We would welcome a definitive debate across Europe to answer such basic questions as:

1) What is an "electronic communication service provider" subject to data retention obligations, and would it include Google services, such as Gmail, Google Talk, or Google Search, in light of different definitions in each Member State?

2) What is the binding retention period for a global Internet company doing business in each Member State, when retention periods range from 6 to 24 months?

3) Do data retention requirements apply to the storage of personal data outside the EU by service providers established in the EU?

4) Will EU Member States go beyond the Directive and implement more stringent retention requirements?

For example, the German Ministry of Justice has proposed that webmail providers should be required to verify the identity of their account holders. Would the German authorities attempt to apply that requirement to Google? Could we challenge its legality in court, either as an unconstitutional infringement of privacy, or as an example of jurisdictional over-reach?

In short, there is tremendous confusion in legal circles across Europe on these issues, and both individuals and companies would benefit from greater clarity from authorities responsible for the Data Retention Directive to answer these very fundamental questions. A public discussion is needed between officials working in data protection and law enforcement to resolve these issues.

It is also important to remember that in the U.S., the Department of Justice and others have similarly called for a 24-month data retention period. Thus, there seems to be an emerging international consensus on 24 months as the outer limit for data retention. This period makes sense for a global company like Google that must comply with the laws of all countries where it does business. Regardless of data retention requirements, logs are an important tool for law enforcement to investigate and prosecute many serious crimes, such as child exploitation. While we have resisted excessive requests from governments in the past, we believe that it is our responsibility to respect law enforcement requests for logs information when law enforcement follows valid legal process. Once again, a reasonable balance needs to be struck between the goals of privacy and the legitimate goals of law enforcement.

In addition, data protection laws, such as Article 17 of the General Directive and Article 4 of the E-Privacy Directive, require companies to ensure that adequate security measures are taken to protect user data. As explained above, our systems engineers require a sufficient historical sample of log server data in order to analyze security threats. A period of 18 to 24 months provides our engineers with sufficient data to analyze these threats without being excessive.

Of course, other laws also impose obligations on companies to retain information. In the U.S., for example, the Sarbanes-Oxley law requires us to retain business records sufficient to establish adequate financial and other controls. The same is true of tax and accounting requirements, especially for paid services, such as clicks on sponsored links, where we have a contractual and accounting obligation to retain data, at a minimum until invoices are paid and the period for legal disputes has expired. These legal obligations must also be considered in connection with our server log retention policies.

So clearly, some period of retention is necessary. A policy of immediate deletion would not serve the interests of our users and would breach many of our legal and ethical obligations to protect our users and their data, and our company records and our systems. A policy of indefinite retention would not respect the privacy expectations of our users, or the requirements of the Data Protection Directive, even though such indefinite retention is common in our industry. We think that a period of 18 to 24 months has a

sound legal and practical basis and strikes the right balance. We are committed to informing our users about our data retention practices so that they can use our services with confidence and full understanding. Finally, it's important to note that logs retention is common in the Internet industry. Indeed, to our knowledge, most Internet companies retain logs for far longer than Google -- in many cases, indefinitely.

We are putting significant resources into creating processes for reliably anonymizing data. Although we are still developing our precise technical methods and approach, we can confirm that we will delete some of the bits in logged IP addresses (i.e., the final octet) to make it less likely that an IP address can be associated with a specific computer or user. And while it is difficult to guarantee complete anonymization, the network prefixes of IP addresses do not identify individual users. Logs anonymization will not be reversible. We will intentionally erase, rather than simply encrypt, logs data so that no one (not even Google) can read it once it has been anonymized. Finally, logs anonymization will apply retroactively and will encompass all of Google's search logs worldwide.

Your letter also raises concerns with regards to cookie notice, purposes and lifetime. In our privacy policy we provide notice to users regarding our use of cookies in industry-standard language:

> "When you visit Google, we send one or more cookies -- a small file containing a string of characters -- to your computer that uniquely identifies your browser. We use cookies to improve the quality of our service by storing user preferences and tracking user trends, such as how people search. Most browsers are initially set up to accept cookies, but you can reset your browser to refuse all cookies or to indicate when a cookie is being sent. However, some Google features and services may not function properly if your cookies are disabled."

We believe that cookies data management in a user's browser is fundamentally a browser/client issue, not a service/server issue. Therefore, the lifetime of a cookie does not indicate or imply any enforcement of data retention. We also believe that cookie lifetimes should not be so short as to expire and force users to re-enter basic preferences (such as language preference). Nonetheless, we acknowledge that cookie lifetimes should be "proportionate" to the data processing being performed.

After considering the Working Party's concerns, we are announcing a new policy: to anonymize our search server logs after 18 months, rather than the previously-established period of 18 to 24 months. We believe that we can still address our legitimate interests in security, innovation and anti-fraud efforts with this shorter period. However, we must point out that future data retention laws may obligate us to raise the retention period to 24 months. We also firmly reject any suggestions that we could meet our legitimate interests in security, innovation and anti-fraud efforts with any retention period shorter than 18 months. We are considering the Working Party's concerns regarding cookie expiration periods, and we are exploring ways to redesign cookies and to reduce their expiration

without artificially forcing users to re-enter basic preferences such as language preference. We plan to make an announcement about privacy improvements for our cookies in the coming months.

We trust that this responds to the issues raised in your letter. We look forward to a continuing discussion with the Working Party as we pursue the common goal of improving privacy protections for everyone on the Internet. Ensuring privacy on the Internet requires sustained energy and engagement, and careful thought about difficult and serious issues. Google is committed to the long road of privacy, and to working with you and other privacy stakeholders around the world to continue to advance privacy protections for all Internet users.

Sincerely,

Peter Fleischer
Global Privacy Counsel
Google Inc.