



18 January 2010

Mr. Alex Türk  
CNIL President and Chairman of the Art. 29 Working Party  
8 rue Vivienne  
75002, Paris  
France

**CONFIDENTIAL**

Dear Mr. President,

Thank you for your letter of 23 October 2009 providing additional clarification of the Working Party's views on search data anonymization and retention. We appreciate the Working Party's recognition of the positive and industry-leading steps we have taken in this area, and as always welcome further dialogue on our shared goals of protecting user privacy while providing consumers with useful and innovative products and services.

**Microsoft's New and Enhanced Approach to Search Data Retention and Anonymization**

As we have noted in the past, we are continually evaluating our data collection, use and retention practices, in order to ensure that we are implementing practices and processes that minimize privacy impacts while offering products and services that our customers want and meeting legitimate business needs (including the marketplace realities of competition with other service providers). As a result of this ongoing evaluation, we are pleased to announce a very significant enhancement to our data retention policy. Specifically, **under our new policy, we will delete the *entire* IP address from search queries at six months.**

This new step will be incorporated into our existing practices that we have previously described. Thus, we will continue to employ our method of de-identifying the cookie ID so that search queries are associated with an anonymized cookie ID from the time they are first collected.<sup>1</sup> We will then delete the full IP address at six months. Finally, we will delete the de-identified cookie ID and any other cross session IDs at 18 months.

---

<sup>1</sup> As we described in our letter of 8 December 2008, this de-identification process creates an anonymized cookie ID by using a one-way cryptographic hash. This process helps prevent the correlation of search data with any account information we may have that personally and directly identifies the user, such as name, email address or phone number. As a result of this initial "de-identification" process used in both our search systems and our online advertising systems, search query data, as well as other data used for ad targeting, is associated with an anonymized identifier rather than an account identifier that could be used to personally and directly identify a consumer. A white paper describing how we have implemented this de-identification process to protect privacy in the context of search and online advertising is available at <http://go.microsoft.com/?linkid=9702232>.

This approach takes into account the elements of the Working Party’s Opinion on data protection issues related to search engines (WP 148) which appear to support the idea that there may be different “levels” of anonymization or pseudonymization that may be appropriate in different circumstances. Thus, during the timeframe in which search data is retained there may be uses of data for which it can be partially anonymized, or pseudonymized. Our process of de-identification constitutes a level of anonymization, or pseudonymization, that occurs with our search data at the outset, followed by an additional and more robust level of anonymization at six months, and ultimately with a complete and irreversible method that the Working Party has recognized as stronger than that employed by the other major search engines.

Of course, in order to begin deleting IP addresses as six months, we will need to make significant changes to our systems and business processes; but we expect we will be able to fully implement our new policy within 12 – 18 months. Further, in keeping with our prior statements, we want to be clear that there may be very limited cases in which some data may need to be retained beyond the time periods described in this policy.<sup>2</sup> Taken as a whole, we believe this new, enhanced approach is significantly stronger and will provide greater overall privacy protections than the policies of the other major search providers.

### **Looking Forward**

We are committed to an ongoing evaluation of our processes and policies to enhance privacy protections for our customers. In doing so, however, we must remain aware that the operation of a search engine is a highly capital intensive and data driven business, and we are facing a very challenging competitive environment. Today, Microsoft is a small player among search engine providers with only 1.7% of European search queries.<sup>3</sup> By contrast, Google is the market leader with an overwhelmingly dominant position in Europe (79.9% of search queries).<sup>4</sup> As a result, Google collects and retains much more search data than any other company, and therefore has an enormous competitive advantage in terms of its ability to analyze and monetize that data.

Thus, for a company like Microsoft, self-imposed restrictions that hinder our ability to compete are possible only if the dominant market leader also adopts equivalent practices. Nevertheless, we have determined that we can unilaterally take the very significant step of deleting the full IP address at six months while still maintaining our ability to achieve the essential purposes for which the data is retained. We hope that the Working Party will be pleased with this step, and will understand the

---

<sup>2</sup> For example, we may in the future offer a personalized search services involving a longer retention period, but would do so in a transparent way with prominent notice and consent. Additionally, there could be pending litigation or other similar legal obligations in which some relevant data may need to be temporarily retained longer. Another example of a foreseeable exception would be that for a specific identified botnet security threat, we will continue to store very small quantities of search terms in association with IP addresses or cross session IDs beyond the 6 month window. These queries are the “signature” of the botnet itself and a smaller control group of typical queries which together will allow us to easily identify a given botnet in the future should it reappear. There are currently about 40 known botnets for which we store such data.

<sup>3</sup> comScore qSearch, Europe region, November, 2009.

<sup>4</sup> *Id.*

market pressures that currently restrict our ability to do more until the dominant search engine fully meets the Working Party's preferred approach.

Finally, we note that your letter states you still have questions regarding behavioral targeting in combination with search. We described in our earlier letter that search data is one type of data that can contribute to the selection of personalized ads shown to a user, but that we have taken a number of steps to protect the privacy of users, such as the de-identification process applied to search data and other data used for ad personalization. We understand that the Working Party is undertaking a more comprehensive review of this topic and, as always, we look forward a productive dialogue on this issue as well.

Kind Regards,

A handwritten signature in blue ink, appearing to read "John Vassallo". The signature is fluid and cursive, with the first name "John" being particularly prominent.

John Vassallo

Vice President European Affairs and Associate General Counsel  
Microsoft Europe