Microsoft

18 Janvier 2010

Mr. Alex Türk

Président de la CNIL et du groupe de travail de « l'article 29 »

8 rue Vivienne

75002, Paris

France

Confidentiel

Monsieur le Président,

Nous vous remercions pour votre lettre du 23 Octobre 2009 par laquelle vous nous avez fournie des éléments de clarification complémentaires sur la vision du groupe de travail de « l'Article 29 » en matière d'anonymisation et de conservation des données de recherche (« *Search »*).

A ce titre, nous apprécions la reconnaissance par le groupe de l'Article 29 des avancées que nous avons faites dans ce domaine. Nous sommes heureux de pouvoir continuer ce dialogue sur le sujet de la protection de la vie privée des internautes tout en prenant en considération l'attente des consommateurs en termes d'expériences innovantes.

La nouvelle approche de Microsoft en matière d'amélioration de l'Anonymisation et de la Conservation des données de recherche.

Comme nous l'avions déjà indiqué, nos pratiques de collecte, utilisation et conservation des données font l'objet d'une évaluation constante. Ce faisant, nous nous assurons que nous mettons en place des pratiques et des procédures permettant de minimiser l'impact négatif des technologies sur la vie privée des consommateurs. Cette démarche doit pouvoir être conciliée avec notre souhait d'offrir aux internautes des produits et des services correspondant au mieux à leurs attentes et en ligne avec nos légitimes objectifs économiques.

A cet égard, nous avons donc le plaisir de vous annoncer aujourd'hui une amélioration majeure dans notre politique de conservation de données. En effet, **nous avons décidé de procéder à la destruction de *l'intégralité* des adresses IP provenant des requêtes de recherche *au bout de six mois*.**

Cette nouvelle étape sera intégrée à nos pratiques existantes telles que nous vous les avions précédemment décrites. Nous continuerons, notamment, à appliquer notre méthode de « désidentification » du Cookie ID afin que les requêtes de recherche soient associées, dès leur première collecte, à un cookie Identifiant anonymisé. [1]

En conséquence, nous détruirons donc l'adresse IP dans son intégralité au bout de six mois et nous détruirons le cookie ID ainsi dé-identifié, comme toutes les autres sessions croisées, au bout de 18 mois.

Cette approche prend en considération l'Opinion du groupe de travail de "l'Article 29" sur la protection des données liée à l'utilisation des moteurs de recherche (WP 148) qui semble admettre l'idée que plusieurs "niveaux" appropriés d'anonymisation ou pseudo-anonymisation puissent exister en fonction des circonstances.

Notre procédure de désidentification constitue un niveau d'anonymisation ou pseudo-anonymisation qui intervient dès le début sur les données de recherche donc au moment de la requête. Elle est ensuite renforcée par un nouveau niveau d'anonymisation, plus robuste, au bout de 6 mois. Enfin, à cela s'ajoute la méthode complète et irréversible que le groupe de travail de l'Article 29 a d'ailleurs reconnue comme étant la plus fiable parmi toutes celles utilisées par les autres principaux moteurs de recherche.

Bien sûr, afin de pouvoir commencer à supprimer les adresses IP au bout de 6 mois, nous avons besoin de procéder à des modifications importantes dans nos systèmes. Malgré tout, nous pensons pouvoir être en mesure de mettre cette nouvelle politique totalement en application dans une période de 12 à 18 mois.

Néanmoins, comme cela avait été précédemment dit, nous souhaitons clarifier le fait que, dans un nombre de cas très limités, certaines données devront être retenues au-delà de la période décrite dans cette politique.[2]

---

[1] Comme nous l'avions précisé dans notre courrier du 8 Décembre 2008, ce procédé de désidentification crée un cookie ID anonymisé en utilisant un algorithme de hachage non réversible. Ce procédé permet d'éviter la corrélation des données de recherche avec les informations d'un compte utilisateur pouvant personnellement et directement identifier l'utilisateur, tel que un nom, une adresse email ou un numéro de téléphone. La conséquence de ce procédé de dés-identification utilisé aussi bien pour nos systèmes de recherche et nos systèmes de publicité en ligne, les données de requêtes de recherché que pour d'autres données utilisées pour la publicité ciblée, est associé avec un identifiant anonymisé plutôt qu'un identifiant de compte pouvant être utilisé afin d'identifier directement et personnellement un utilisateur. Un livre blanc décrivant comment nous avons mis en place ce procédé de dés-identification afin de protéger la vie privée dans le cadre des recherches (Search) et de la publicité en ligne est disponible à l'adresse suivante http://go.microsoft.com/?linkid=9702232.

[2] Par exemple, nous pourrions, dans le futur, offrir un service de recherche personnalisé impliquant une durée de conservation plus longue, mais nous le ferions de façon transparente notamment dans une déclaration de confidentialité proéminente. De plus, il pourrait exister des procédures judiciaires ou d'autres obligations légales qui nécessiteraient que certaines données soient retenues plus longtemps. Un autre exemple d'exception

Dans l'ensemble, nous estimons que cette nouvelle approche est significativement plus performante et permettra de fournir un meilleur niveau de protection de la vie privée par rapport à celui proposé par les autres principaux moteurs de recherche.

Dans l'avenir

Nous nous sommes engagés à un processus permanent d'évaluation de nos procédures et politiques de « privacy » afin d'améliorer la protection de la vie privée des internautes utilisant nos services. Cependant, il importe d'être conscient que, dans l'environnement concurrentiel existant, la gestion d'un moteur de recherche représente un investissement financier important et génère une activité reposant, en particulier, sur les données. Or, aujourd'hui Microsoft est un acteur de petite taille dans le secteur des moteurs de recherche avec seulement 1,9% des requêtes de recherche européennes. Par contraste, Google est le leader de ce marché avec une position dominante en Europe (79,2% des requêtes de recherche). A cet égard, Google collecte et détient bien plus de données de recherche que n'importe quelle autre société et, par voie de conséquence, dispose d'un énorme avantage concurrentiel lui permettant d'analyser et de monétiser lesdites données.

Cependant, et malgré cette situation, nous avons pris la décision unilatérale de franchir une étape significative en supprimant l'intégralité des adresses IP au bout de six mois tout en préservant notre capacité à atteindre l'objectif pour lequel les données sont conservées. Pour une société comme Microsoft s'imposer d'autres restrictions réduisant ses possibilités d'être pleinement compétitif ne pourrait être possible que si l'acteur dominant du marché adoptait à son tour des pratiques équivalentes.

Nous espérons que le groupe de travail de "l'Article 29" sera satisfait de cette avancée et comprendra, au regard de la pression du marché, que notre capacité à faire plus encore est, aujourd'hui, contrainte tant que le moteur de recherche dominant ne s'aligne pas sur les recommandations de votre groupe de travail.

En dernier lieu, nous notons que votre lettre mentionne que vous avez toujours des interrogations sur la publicité ciblée associée aux recherches (*Search*). A cet égard, nous avions précisé dans notre dernier courrier du 8 décembre 2008 que les données de recherche ne constituent qu'un des éléments pouvant contribuer à la sélection de publicités personnalisées apparaissant aux utilisateurs. Cela étant nous avons pris un certain nombre de mesures afin de protéger la vie privée des utilisateurs, tels que la

---

prévisible serait que lors d'une menace de sécurité par des réseaux de robots « botnet » spécifiquement identifiée, nous continuerions à stocker de très petites quantités de données de recherche en association avec des adresses IP ou des identifiants de sessions croisées au-delà de la fenêtre de 6 mois. Ces requêtes sont la "signature" du réseau de robots « botnet » lui-même et associées à un plus petit groupe de requêtes typiques cela nous permettrait d'identifier facilement « un botnet » donné à l'avenir s'il devait réapparaître. Il y a actuellement 40 réseaux de robots « botnets » connus pour lesquels nous stockons ce type de données.

procédure de désidentification des données de recherche et des autres données utilisées pour la publicité ciblée.

Nous comprenons que le groupe de travail de « l'Article 29 » a entrepris une revue plus complète de ce sujet et nous souhaitons engager avec vous une discussion productive sur ce sujet.

Nous vous prions d'agréer, Monsieur le Président, l'expression de nos salutations respectueuses.

John Vassallo

Vice Président des Affaires Européenne et Associate General Counsel
Microsoft Europe