

La vie privée n'existe pas sur Internet

Description

Qu'est-ce qu'une donnée personnelle, une adresse IP, un cookie ou encore « les traces » circulant sur Internet ? Quels sont les enjeux de l'interconnexion des fichiers comportant des données personnelles à l'heure du tout numérique ?

A l'occasion d'une interview à la BBC, Sir Tim Berners-Lee, l'un des fondateurs de l'hypertexte à l'origine du World Wide Web s'inquiétait : « *Je veux savoir, si je consulte (sur le Web) des livres sur certaines formes de cancer, que mon assurance ne va pas augmenter de 5 % parce qu'ils m'imaginent malade* »... Pourquoi pareille inquiétude ? Quel danger l'informatique peut-elle faire courir aux particuliers ?

Si l'on en croit le professeur Jean Morange, c'est le fichage des individus qui se révèle dangereux. En revanche, l'informatisation des fichiers s'avère d'une grande efficacité, à condition d'être correctement maîtrisée. La France figure parmi les premiers pays à avoir pris conscience que la vie privée pouvait être mise en danger par l'informatique : c'est le fondement même de la loi du 6 janvier 1978, modifiée par la loi du 6 août 2004, qui permet d'assurer cette maîtrise de l'informatique lorsqu'elle implique des données personnelles.

Pour comprendre les enjeux de l'informatisation des fichiers à l'heure du Web 2.0 qui place l'utilisateur au centre du système, il convient d'apporter quelques précisions sur le fonctionnement du Web, d'Internet ainsi que le régime juridique qui s'y rapporte.

Internet et le Web : fournisseur d'accès au réseau et éditeurs de services du Web

Internet ayant été popularisé par l'apparition du World Wide Web, les deux sont parfois confondus. En réalité, Internet est le réseau informatique mondial qui rend accessible au public des services dont le World Wide Web ou le courrier électronique, la messagerie instantanée ou encore les systèmes de partage de fichiers poste à poste (*peer-to-peer*). Le World Wide Web est un système hypertexte public utilisant le réseau Internet et qui permet de consulter, à l'aide d'un logiciel de navigation (Internet Explorer, Opera, Firefox etc.), des services hébergés sur des serveurs. Le droit distingue ainsi, parmi les acteurs privés, ceux dont l'objet est de fournir au public un accès au réseau Internet, comme Orange, Free, Neuf, Numéricable..., de ceux qui rendent accessibles sur le Web un service, comme Google, Facebook ou Fnac.com établissant ainsi la différence entre des prestataires techniques et des éditeurs de contenus.

Le droit impose aux fournisseurs d'accès au réseau de conserver les données de connexion de leurs clients et demande aux éditeurs de services du Web d'effacer les données personnelles de leurs utilisateurs. Pour

comprendre ce paradoxe, il faut distinguer les données de connexion des données personnelles et autres types de données informatiques.

Donnée de connexion, donnée personnelle, adresse IP et cookie

Pour accéder de chez soi à un site Web, chaque ordinateur connecté au réseau Internet est identifié par un numéro unique qui permet de le retrouver parmi l'ensemble des ordinateurs connectés. Ce numéro unique s'appelle l'adresse IP (Internet Protocol) et prend la forme d'une suite de chiffres : 212.85.150.134. Celle-ci est attribuée à l'internaute par son fournisseur d'accès. A chaque adresse IP correspond un client. L'adresse IP est donc un peu l'équivalent du numéro de téléphone pour les ordinateurs connectés à Internet.

L'adresse IP est-elle considérée ou non comme une donnée personnelle ? En France, une donnée personnelle est définie par loi du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés dite « Informatique et libertés » comme « *toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement, par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres* ». C'est par exemple le nom et le prénom d'une personne qui permet une identification directe de l'individu ou le numéro de téléphone qui permet l'identification indirecte de son propriétaire.

En France comme en Europe puisque l'adresse IP permet de remonter à l'ordinateur d'un internaute grâce à son fournisseur d'accès, il s'agit d'une donnée « indirectement nominative », permettant l'identification indirecte d'un individu.

Aux Etats-Unis, en revanche, où les risques d'atteinte à la vie privée peuvent être minorés au profit des enjeux économiques du commerce informatique, de la libre circulation des données et la liberté des échanges dont elles sont l'objet, les données informatiques dont les adresses IP ont été considérées comme les biens d'une économie de marché et ne sont pas considérées comme une donnée personnelle.

Pourquoi la loi impose-t-elle aux fournisseurs d'accès à Internet (qui ne le faisaient pas avant) de conserver les données de connexion ? Depuis septembre 2001 et les objectifs de sécurité qui en découlent, la conservation des données de connexion fait l'objet de nombreux débats. La loi du 21 juin 2004 pour la confiance dans l'économie numérique et la loi du 23 janvier 2006 relative à la lutte contre le terrorisme imposent aux fournisseurs d'accès la conservation des « données de connexion », pendant 12 mois, pour différentes finalités dont la recherche, la constatation et la poursuite des infractions pénales ainsi que la lutte contre le terrorisme.

Cette préoccupation sécuritaire est également relayée par la directive européenne 2006/25/CE du 15 mars 2006 qui prévoit une conservation des données générées ou traitées dans le cadre de la fourniture de services de communications électroniques accessibles au public ou de réseaux publics de communication pendant une durée allant de 6 à 24 mois et dont la finalité est « *la recherche, la détection et la poursuite d'infractions graves telles qu'elles sont définies par chaque Etat membre dans son droit interne* ». La loi

impose ainsi aux fournisseurs d'accès au réseau de conserver les données permettant d'identifier une personne à des fins de police préventive.

Pourquoi la CNIL demande-t-elle alors aux fournisseurs de services du Web (comme Google, Yahoo!, ou Facebook) de ne pas conserver indéfiniment les données de connexion ? Le World Wide Web est un service accessible sur le réseau Internet. Parmi les très nombreux services accessibles sur le Web, les moteurs de recherche, les réseaux sociaux ou encore les sites de commerce électronique sont parmi les plus utilisés.

La collecte des données effectuée sur les sites Web est rendue possible grâce aux cookies, ces petits fichiers dits « texte » placés sur le disque dur d'un ordinateur lorsqu'il est connecté pour la première fois à un site Web. Le cookie est installé par le site Web visité, via le navigateur Internet, sur le disque dur de l'ordinateur. Les cookies permettent au site Web qui les émet de « reconnaître » l'internaute en recueillant un certain nombre d'éléments d'identification : l'adresse IP, le système d'exploitation de son ordinateur (Windows, Mac, Linux...), le navigateur utilisé, ainsi éventuellement que des informations statistiques comme les pages consultées, le nombre de visites, les actions effectuées sur le site. Les habitudes de consultation de l'internaute concerné sont ainsi repérées.

En France, même si une entreprise offrant un service sur le Web n'a d'établissement ni en France ni en Europe, le fait d'enregistrer un cookie dans l'ordinateur de celui qui s'y connecte, rend le droit applicable puisque l'objet même d'un cookie est de traiter des données informatiques, dont l'adresse IP qui est une donnée « indirectement nominative ».

Les données collectées par les moteurs de recherche

Quand un internaute tape le mot « voiture » sur un moteur de recherche comme Google ou Yahoo!, la requête est envoyée de son ordinateur vers les serveurs du site Web qui recherche les résultats pertinents et les affiche ensuite sur l'ordinateur de l'internaute. Les moteurs de recherche conservent les traces de ces requêtes appelées « logs de connexion » et place également un cookie dans l'ordinateur de l'internaute.

Ces « logs de connexion » concernent un certain nombre de données dont l'adresse IP, la requête, le navigateur Internet, et d'autres informations relatives à la navigation. Jusqu'à récemment, Google conservait ainsi ces « logs de connexion » pendant 24 mois, après quoi les deux derniers chiffres de l'adresse IP étaient détruits et un nouveau numéro de cookie était attribué à l'internaute. Puis Google a volontairement ramené cette durée de conservation des données personnelles à 18 mois. Cependant le droit applicable aux moteurs de recherche étant celui de la directive européenne du 24 octobre 1995 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, cela signifie qu'ils ne sont pas tenus d'enregistrer les adresses IP des internautes qui utilisent leur service.

Cette directive est d'ailleurs applicable à tous les sites Web, soit du fait de leur établissement dans l'un des pays de l'Union européenne, soit parce qu'ils ont recours à des moyens de traitement des données personnelles comme les cookies. Or, comme les moteurs de recherche américains ne considèrent pas

l'adresse IP comme une donnée personnelle, ils estiment que la durée de conservation des « logs de connexion », dont l'adresse IP, ne rentre pas dans le champ de compétence de la protection des données personnelles, mais dans celui de la sécurité informatique.

En réponse, le G29, un groupe de travail regroupant les 27 « CNIL » européennes, a adopté à l'unanimité, le 4 avril 2008, un avis précisant les règles applicables aux moteurs de recherche. Cet avis précise notamment que les données personnelles enregistrées en Europe par les moteurs de recherche, doivent être effacées au plus tard au bout de 6 mois. L'avis précise également les conditions d'application des règles juridiques communautaires et formule des recommandations susceptibles d'améliorer la protection et le droit des utilisateurs des moteurs de recherche. Le G29 a rappelé à cet égard que les moteurs de recherche, étant des « services de la société de l'information », ne sont pas concernés par la directive 2006/24/CE du 15 mars 2006 relative à la conservation des données, contrairement aux fournisseurs d'accès au réseau Internet ou aux opérateurs de télécommunications.

Cela signifie que les moteurs de recherche ne sont pas légalement obligés de conserver des informations sur les connexions des utilisateurs. En pratique, un moteur de recherche comme Google ou Yahoo ne devrait donc pas conserver indéfiniment l'historique des requêtes effectuées et des sites consultés par un internaute. Cet historique peut contenir des informations très intimes, comme des problèmes conjugaux, une opinion politique, à partir desquelles il est possible de déduire des habitudes de vie supposées ou un certain comportement.

Si les moteurs de recherche invoquent un certain nombre d'arguments pour conserver ces données, arguments liés à la meilleure efficacité du service pour l'utilisateur ou à la sécurité informatique, leur finalité est avant tout de pouvoir dresser un profil très précis des utilisateurs du service, essentiellement à des fins commerciales.

Les données collectées par les réseaux sociaux

Un site Web dit « de réseau social » est un service du Web dont l'objet est de permettre à ses membres de publier des informations personnelles et de se « connecter » à ses amis ou relations professionnelles. Rien de plus utile pour retrouver un ancien camarade de classe, constituer un réseau professionnel, garder le contact avec des personnes disséminées aux quatre coins du monde. Sur le principe et dans le souci de sauvegarder les droits individuels, la loi du 6 janvier 1978 interdit de mettre ou de conserver en mémoire informatisée, sauf accord exprès de l'intéressé, des données nominatives qui, directement ou indirectement, font apparaître les origines raciales, les opinions politiques, philosophiques ou religieuses, ou les appartenances syndicales d'une personne.

Or, l'objet des réseaux sociaux est de s'appuyer sur cet accord exprès des internautes pour collecter et conserver en mémoire ce type de données. En se basant sur le côté ludique, social, voire utile du service, les réseaux sociaux permettent aux internautes de se « ficher eux-mêmes ». Le modèle d'affaire de ce type de site est fondé sur la gratuité du service pour l'internaute en échange de l'utilisation à des fins publicitaires de

ses données personnelles et ses interactions sociales.

Mais combien, parmi les 70 millions d'inscrits sur Facebook, sont informés qu'ils ne peuvent plus supprimer leur compte mais simplement le « désactiver », ce qui permet à Facebook de tracer les connexions entre les différents membres, y compris avec ceux ne souhaitant plus utiliser le service ? En s'inscrivant sans vraiment lire les conditions générales d'utilisation du site Web, les utilisateurs de Facebook ont accepté le transfert et le traitement de leurs données personnelles aux Etats-Unis en leur conférant un droit perpétuel et irrévocable d'utiliser, de copier ou de distribuer toutes les données communiquées.

De plus, Facebook se réserve le droit de collecter et de compiler autour d'un internaute toutes les informations publiques que ce dernier pourrait avoir publié sur d'autres services du Web comme les blogs, les forums, les sites participatifs ou les plates-formes d'échange pour enrichir son profil, toujours à des fins publicitaires. Il y a donc une certaine contradiction entre l'objet du service, constituer un réseau de relations sur Internet, et l'encadrement juridique des pratiques commerciales qui en découlent : dresser le profil publicitaire des membres de la manière la plus détaillée possible à partir des tous les éléments disponibles sur le Web pour offrir à un annonceur un ciblage précis.

Le groupe de travail G29 publiera avant l'été 2008 un avis et des recommandations à l'attention des sites Web exploitant un réseau social comme celui adressé le 4 avril 2008 aux moteurs de recherche.

Risques d'usurpation d'identité

L'usurpation d'identité (ou vol d'identité) consiste à prendre délibérément l'identité d'une autre personne, généralement dans le but de réaliser une action frauduleuse, comme d'accéder aux comptes bancaires de la personne usurpée, ou de commettre un délit ou un crime de manière anonyme. Un sondage, effectué pour le compte de la campagne britannique Get Safe Online de sensibilisation à la sécurité informatique, avance que 15 % des utilisateurs des sites de réseaux sociaux « *n'utilisent aucune des possibilités pour rendre confidentielles leurs informations sur ces sites* » et 24 % des internautes utilisent le même mot de passe pour tous les sites », que 27 % des 18-24 ans ont posté des photos de tiers sans leur consentement et que 34 % des 18-24 ans, et 30 % des 25-34 ans, « *révèlent des informations susceptibles d'être utilisées à des fins criminelles* ».

Pour éviter ces risques, l'application des principes préconisés par la CNIL (et plus largement en Europe) conduit à recommander d'effacer les données personnelles le plus vite possible, à rendre anonymes les traces et à refuser l'interconnexion des fichiers qui ne s'appuie pas sur le consentement exprès de l'utilisateur.

Les acteurs de la publicité en ligne, la publicité comportementale, le consentement de l'utilisateur, l'interconnexion de fichiers

Mais les fournisseurs d'accès, les sites Web et, d'une manière générale, les entreprises de l'informatique et

des télécommunications n'appartiennent jamais qu'à une seule catégorie d'acteurs. La société Google propose de très nombreux services sur le Web, moteur de recherche, plate-forme d'échange de vidéos, de photos, système de messagerie etc. et elle a participé début 2008 aux enchères organisées par la Federal Communications Commission (FCC) pour l'attribution de fréquences, afin de devenir opérateur de télécommunications aux Etats-Unis. Microsoft est à la fois un éditeur de logiciels pour la micro-informatique, la téléphonie mobile et propose également de nombreux services sur le Web. La CNIL demande que ces entreprises ne puissent pas systématiquement interconnecter les données personnelles d'un service à l'autre, à la suite d'un rachat ou d'une prise de participation, sans demander l'accord explicite de l'intéressé.

L'élément le plus important que les Anglo-Saxons semblent systématiquement oublier, est le consentement de la personne. En Grande-Bretagne, trois opérateurs rassemblant les deux tiers des abonnés à Internet, British Telecom, Carphone Warehouse et Virgin Media, ont utilisé, entre septembre et octobre 2007, les services et le logiciel espion (*spyware*) de Phorm, alias 121Media, pour épier les agissements de 18 000 internautes clients de leur prestation d'accès à Internet. Une surveillance qui s'est effectuée sans l'accord des personnes concernées.

Ce qui est confidentiel, ce n'est pas tant notre identité, déjà accessible par bien d'autres moyens que le Web, mais notre comportement et les liens qui rattachent ces « traces » à notre identité personnelle. Si hier, la publicité s'adressait à l'audience la plus large possible en s'affichant dans la rue, les magazines et d'autres supports fixes, le réseau Internet offre la possibilité de personnaliser les annonces publicitaires en les adressant à des publics ciblés et aisément identifiables.

Mais qu'elle soit « contextuelle » ou « comportementale », cette publicité personnalisée devrait reposer sur le consentement exprès de l'utilisateur. Les grands acteurs du Net profitent actuellement du vide juridique international pour collecter un maximum de données à l'insu de l'utilisateur, ne sachant d'ailleurs pas toujours quoi faire de cette masse d'informations.

Une étude sur le sujet menée par ComScore pour le *New York Times* révèle que Yahoo! récolte 811 informations sur chaque internaute qui visite l'un de ses sites alors qu'il fait une recherche, un achat ou regarde une vidéo. Menée en décembre 2007 sur le trafic des sites des quinze géants américains du Web (Yahoo!, Google, Microsoft, etc.), l'étude montre jusqu'à quel point ces sociétés collectent les données générées par les internautes. A eux seuls, Yahoo!, Google, Microsoft, AOL et MySpace ont enregistré pas moins de 336 milliards de transmissions de données, le but étant d'affilier chaque internaute à un profil publicitaire.

Des chiffres qui expliqueraient, entre autres, pourquoi Microsoft était encore prêt, fin avril 2008, à déboursier 47 milliards de dollars pour mettre la main sur Yahoo!. « *Beaucoup de contrats visent en fait ces données* », explique David Verklin, directeur général de l'agence publicitaire Carat Americas. Croiser et compiler les données récoltées sur le réseau et les sites Web permet de dresser une « identité numérique » très précise de l'internaute. Or, si ce croisement de ces données est fait sans l'accord de l'intéressé, il s'agit

bien là d'une atteinte au respect de la vie privée, puisque ces pratiques rendent possibles l'identification des utilisateurs à leur insu pour des finalités autres que celles pour lesquelles ils ont, d'une part, signé un contrat avec un fournisseur d'accès et, d'autre part, accepté les conditions générales des sites Web qu'ils utilisent.

Le caractère mondial du réseau Internet rend le contexte juridique particulièrement complexe car l'une des principales caractéristiques d'une règle de droit est d'avoir un champ d'application précis, limité à un territoire et suivant un champ de compétence ordonné. Or puisque les échanges de données sur le réseau Internet sont par nature transfrontaliers et que, comme le rappelle Gustave Le Bon, « *le droit ne commence à dater que du moment où l'on détient la force nécessaire pour le faire respecter* », toutes les pratiques sont actuellement possibles.

Pourtant, le succès des services du Web repose à la fois sur la confiance de l'utilisateur et sur une personnalisation des services toujours plus réussie. A cette personnalisation doit correspondre une réflexion de fond quant au respect de sa vie privée que le doyen Carbonnier définissait comme étant « *la sphère secrète d'où l'individu aura le pouvoir d'écarter les tiers [...], le droit d'être laissé tranquille [...]* ».

Aujourd'hui, les identités de chacun sont accessibles de plus en plus facilement sur le réseau public et de manière de plus en plus détaillées. Frédéric Cavazza et Raphaël Lefear, deux blogueurs ayant analysé les différents types de données personnelles laissées sur le Web par les internautes en dénombrant douze : ce que je dis, ce que je partage, ce que j'apprécie, ce qui me passionne, ce que je sais, ce qui me représente, ce que je fais, ce qui se dit sur moi, ce que j'achète, qui je connais, comment et où me joindre et qui atteste de mon identité... Ce sont tout à la fois, l'identité personnelle et professionnelle, les habitudes de consommation, de jeux, le contenu généré ou publié par l'utilisateur, ses rencontres et ses réseaux dont il est question. Autant d'informations privées qui, regroupées, constitueraient notre identité numérique que certains appellent aussi l'empreinte numérique.

Si aujourd'hui, la majeure partie des connexions à Internet se fait encore à partir d'un ordinateur, fixe ou portable, il faut anticiper une société où le téléphone portable, à la fois objet intime et fenêtre sur le monde, permettra également de s'y connecter. Demain, la majorité des objets de la vie courante seront également reliés au réseau Internet par fréquence radio via des puces invisibles (*Radio Frequency Identification*). Nous laisserons alors infiniment plus de traces sur Internet et rares seront nos déplacements, nos habitudes, nos préférences qui pourront échapper à un système informatique où les identités circulent de manière aussi fluide et transparente qu'aujourd'hui.

Plus la collecte massive de données sera rendue possible et même encouragée par certains, plus la protection des libertés individuelles, le droit au respect de la vie privée et de l'intimité sera difficile à protéger. C'est la raison pour laquelle Sir Tim Berners Lee s'inquiétait déjà en 1997 de la finalité pour laquelle nos données personnelles peuvent être collectées, compilées et analysées par des acteurs privés. A la même époque, un banquier américain du Maryland avait obtenu une liste de personnes atteintes d'un cancer qui, croisée avec celle de ses clients, lui permettait de rejeter systématiquement les malades candidats à un emprunt.

Et c'est bien ce genre de pratiques qui rendent le travail du G29 et de la CNIL déterminant pour l'avenir des services du Web et d'Internet dont le succès repose avant tout sur la confiance.

Sources :

- www.cnil.fr
- « Informatique, servitude ou libertés ? », colloque organisé au Sénat par la Commission nationale de l'informatique et des libertés, CNIL, 7-8 novembre 2005.
- « Le système français de protection des données personnelles », Annie Gruber, Les Petites Affiches, LPA, n° 90, 4 mai 2007.
- « La publicité en ligne mise sur l'intime », Cécile Ducourtieux et Laurence Girard, *Le Monde*, 11 novembre 2007.
- « Réseaux sociaux : quand les utilisateurs s'en fichent » Arnaud Belleil et Hubert Guillaud, Internet Actu, 20 novembre 2007.
- « Mais où va donc le Net », Gueric Poncet, *Le Point*, 14 février 2008.
- « Facebook, la CNIL appelle à la vigilance », Antoine Gendreau, Anne- Laure Villedieu, *Les Echos*, 27 février 2008.
- « Phorm and ISP Business Models », Ian Fogg, Jupiter Resarch, avril 2008.
- « Recherches en ligne : délai maximal de 6 mois en UE pour garder données », AFP, 11 avril 2008.
- « Les secrets de l'ogre du Net », Quentin Domart, Jean-Baptiste Su, *L'Expansion*, mai 2008.
- « Pas de confidentialité sur Facebook, selon la BBC », *Le Nouvel Observateur*, 5 mai 2008.

Categorie

1. Articles & chroniques

date créée

20 mars 2008

Auteur

jacquesandrefines