Vie privée et Internet

Document mis à jour le : lundi 29 janvier 2024

Vraiment, rien à cacher sur Internet?
Que s'y passe-t'il?
Pourquoi?
Peut-on se protéger?

Licence Creative Commons CC BY-NC-SA 4.0



Vous êtes autorisés à :

- Partager copier, distribuer et communiquer ce document par tous moyens et sous tous formats.
- Adapter remixer, transformer et recréer à partir de ce document pour toute utilisation, sauf commerciale.

https://creativecommons.org/licenses/by-nc-sa/4.0/deed.fr

Télécharger le document : https://chapellut.fr/Internet-VPr/

Par Jean-Michel Chapellut cidex 737 Fr 38090 Villefontaine https://contact.chapellut.fr/

Sommaire

1- Presentation	1
1. Au sujet de ce document	
2. Webographie d'introduction	
3. Pourquoi serais-je suivi, pisté, infecté ?	3
II- La réglementation	4
1. Données personnelles	
2. Le RGPD	
3. DSA et DMA	
4. La CNIL	
III- Bases d'Internet	
1. Adresses IP	
2. Le chemin des données	
Le système de nom de domaine (Domain Name System DNS) Pour les curieu.ses.x	
IV- Naviguer sur le web	
1. Adresser le serveur.	
Échanges entre le navigateur et le serveur Les cookies et le stockage du navigateur	
4. Même là où on ne s'y attend (peut-être) pas	
5. Alors, pourquoi est-on tracé en navigant sur le web ?	
6. Gérer l'utilisation et la configuration du navigateur	
7. Quelques extensions utiles	
8. Naviguer incognito	15
V- Utiliser la messagerie électronique	17
1. Risques auxquels la messagerie nous expose	
2. Tenter de se protéger	18
3. Services collecteurs d'adresses	
4. Logiciels clients de messagerie	19
VI- Diffusion « volontaire »	20
1. Services nécessitant une inscription	
2. Diffusion d'informations personnelles	
3. Les « Clouds »	
VII- Les logiciels (applications ou programmes)	21
1. Le système d'exploitation	
2. Les applications installées	
VIII- Applications et services libres ou propriétaires	
2. Fidélisation / dépendance des utilisateurs	
3. Diversifier les prestataires	
IX- Objets connectés - Internet des objets	
2. Les risques	24
X- Un exemple : Google	25
1. Un moteur de recherche	
2. Un navigateur web	
3. Une messagerie électronique	
4. Un système d'exploitation pour mobile	
XI- En conclusion	27
XII- Webographie sommaire	28
1. Vie privée	28
2. Logiciel libre	
3. Et aussi	30

I- Présentation

1. Au sujet de ce document

Ce document cherche à présenter l'autre versant d'Internet, celui que l'on ne voit pas ; mais aussi celui auquel on ne pense pas toujours...

Il essaye de montrer ce qui se passe du côté des serveurs et des prestataires de services informatiques et donne quelques pistes afin d'éviter ce qui peut vous sembler abusif.

Il ne s'agit ni de préconisations d'usage ni de conseils, chacun est libre de faire ce qu'il veut avec ses appareils (dans la limite de la réglementation), mais il paraît utile de connaître « l'envers du décor».

2. Webographie d'introduction

Je n'ai rien à cacher, mais...

Rien à cacher

Vidéo Médiapart et La Parisienne Libérée

https://tube.fede.re/w/3cw5w3oQQsLLRZdUjN2KT8

https://www.dailymotion.com/video/x1bchsv

https://www.youtube.com/watch?v=rEwf4sDgxHo

Profilage: vous reprendrez bien un petit cookie?

Audio France-culture (2023)

https://www.radiofrance.fr/franceculture/podcasts/la-science-cqfd/profilage-vous-reprendrez-bien-un-petit-cookie-7634890

Avons-nous livré toute notre vie privée à internet?

Audio France-culture (2018)

https://www.radiofrance.fr/franceculture/avons-nous-livre-toute-notre-vie-privee-a-internet-5922509

Si c'est gratuit, vous êtes le produit.

Vidéo Adesias agence de communication pour les entreprises (2014) https://www.youtube.com/watch?v=8vLSf1i4E7A

Souvent, mais pas toujours...

Scandale Facebook-Cambridge Analytica

Texte Wikipédia (2014...)

https://fr.wikipedia.org/wiki/Scandale Facebook-Cambridge Analytica

Texte Le Monde (2018)

https://www.lemonde.fr/pixels/article/2018/03/22/ce-qu-il-faut-savoir-sur-cambridge-analytica-la-societe-au-c-ur-du-scandale-facebook 5274804 4408996.html

Opération Doppelgänger

Texte Wikipédia (2022...)

https://fr.wikipedia.org/wiki/Op%C3%A9ration_Doppelg%C3%A4nger

Texte The Conversation (2023)

https://theconversation.com/operation-doppelganger-quand-la-desinformation-russe-vise-la-france-et-dautres-pays-europeens-208071

Publicité ciblée en ligne : quels enjeux pour la protection des données personnelles ?

Texte CNIL (2020)

https://www.cnil.fr/fr/cookies-et-autres-traceurs/regles/cookie-walls/publicite-ciblee-en-ligne-quels-enjeux-pour-la-protection-des-donnees-personnelles

Spyware, ramsonware, malware, virus, etc.

Texte Malekal (2020)

https://www.malekal.com/virus-malwares-dossier-complet/

Page: 2

3. Pourquoi serais-je suivi, pisté, infecté?

L'informatique et Internet en particulier sont très utiles et pratiques, mais comme il y a plus de 5 milliards d'internautes dans le monde₁, cela tente certains d'en tirer un profit plus ou moins honnête.

Les statisticiens

Il s'agit de connaître soit au niveau d'une organisation, soit au niveau plus global : pays, continent, monde. La fréquentation d'Internet, les équipements utilisés, leur configuration, etc. Cela n'est pas bien méchant et peut être utile.

Les publicitaires

Convaincre, tenter par le biais de la publicité, n'est pas nouveau. Mais Internet touche un grand nombre de personnes à moindres frais.

Comme il est préférable de s'adresser à des personnes qui seraient, à priori, intéressées par ce que l'on leur propose, il est utile de connaître leur centres d'intérêt et les suivre.

Ainsi l'internaute en visitant un site web anodin ou de réseau social pourra remarquer une publicité qui a une bonne probabilité de l'intéresser. Il peut être aussi sollicité par sa messagerie.

Les prosélytes

Là il s'agit de convaincre. Outre les bien connus influenceurs/influenceuses, certains cherchent à propager des idées, des rumeurs avec des buts pas toujours avouables.

Certains sites web sont spécialisés dans la diffusion de telles informations. Celles-ci sont aussi véhiculées par les réseaux sociaux et les messageries (cf. Opération Doppelgänger).

Amélioration de l'expérience utilisateur

Parfois utile; mais...

A partir des données obtenues concernant le profil de l'internaute, certains prestataires en déduisent ses opinions, tendances... L'internaute sera ainsi plus informé de ce qui les renforce (cf. scandale Cambridge Analytica).

Ces techniques appelées « Bulles de filtres » et « Chambre d'écho » tendent à isoler culturellement l'internaute₂.

Les escrocs

Eux aussi essayent de convaincre mais d'une information trompeuse; convaincre que son matériel a été piraté ; convaincre qu'un proche est en grande difficulté ; convaincre que son compte en banque, sa messagerie, son colis en livraison ont des soucis...

Le résultat attendu est le plus souvent pécuniaire. Cependant on peut aller bien plus loin, une fois que l'on a trouvé le « pigeon »...

Les méchants

Ils vont infecter (en vrai cette fois) l'ordinateur voire l'ensemble du réseau local.

Et cela leur permettra d'obtenir des informations qui y sont enregistrées (informations de compte, mots de passe, informations confidentielles...) soit pour en profiter directement, soit pour les vendre, soit pour menacer de les publier.

Une autre méthode consiste à chiffrer toutes les données (<u>rançongiciels</u>). Pour obtenir la clef de déchiffrement il faudra payer !

Dans tous les cas

Un autre problème peut aussi se poser si ces informations plus ou moins confidentielles enregistrées chez un prestataire viennent à être « piratées » puis utilisées frauduleusement ou diffusées.

Remarquons aussi que toutes ces intrusions dont la plupart sont indésirables et inutiles consomment de la bande passante Internet et de l'énergie.

¹ https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx

² https://www.fondationdescartes.org/2020/07/bulles-de-filtre-et-chambres-decho/

II- La réglementation

Depuis longtemps, le législateur s'est intéressé à la protection des données personnelles. L'arrivée d'Internet a permis une plus grande méfiance quant à l'acquisition et au traitement de ces données.

Dès 1978, la loi « Informatique et libertés » réglemente le traitement des données personnelles, puis « La loi pour la confiance dans l'économie numérique de 2004 », suivie par la loi « Pour une République numérique » de 2014, enfin en 2016 le « Règlement général sur la protection des données » disposent d'articles relatifs à la protection des données personnelles.

La vidéo « *La CNIL*, *40 ans et toujours dans l'air du temps !* » (2018) présente l'histoire de la réglementation et ses raisons d'être. https://www.youtube.com/watch?v=i k8ozkY2l4

1. Données personnelles

Une donnée à caractère personnel (ou « donnée personnelle ») défini toute information se rapportant à une personne identifiée ou identifiable. Il existe deux types d'identification :

- identification directe (nom, prénom, etc.);
- identification indirecte (identifiant, numéro, etc.).

Certains dont le métier est la capture de données personnelles (régies publicitaires par exemple) utilisent les informations recueillies depuis différents prestataires (sites web, application) afin d'obtenir des profils plus détaillés.

Notons aussi que des données personnelles sont collectées hors d'Internet via, par exemple, les cartes de fidélité.

2. Le RGPD

Depuis 2016, le **règlement général sur la protection des données** (RGPD) est en vigueur. C'est un texte réglementaire issu de l'Union Européenne, il établit des règles sur la collecte et l'utilisation des données. Il a été transposé en France par la loi du 20 juin 2018 relative à la protection des données personnelles.

Voici les grands principes du RGPD vu de l'utilisateur d'un service utilisant des données personnelles.

- Le consentement « explicite » et « positif » : l'utilisateur du service doit accepter la collecte des données et savoir l'utilisation qui en sera faite.
- Le droit d'accès : l'utilisateur doit savoir quelles données sont collectées.
- Les droits de rectification et d'effacement : l'utilisateur peut demander à corriger certaines données le concernant ainsi qu'un effacement complet.
- Ainsi que d'autres...

Notons que le RGPD s'applique à toute entreprise qui traite des données relatives aux citoyens de l'Union Européenne même si l'entreprise est établie à l'extérieur₃.

Voir « Le règlement général sur la protection des données » sur le site du Conseil de l'Union européenne : https://www.consilium.europa.eu/fr/policies/data-protection/

³ https://commission.europa.eu/law/law-topic/data-protection/reform/rules-business-and-organisations/application-regulation/who-does-data-protection-law-apply_fr

3. DSA et DMA

Les nouvelles législations sur les services numériques (DSA : Digital Services Act) et sur le marché numérique (DMA : Digital Market Act) forment un ensemble unique de règles qui s'appliquent à l'ensemble de l'UE₄.

leurs objectifs principaux sont la création d'un espace numérique sûr dans lequel les droits fondamentaux de tous les utilisateurs de services numériques sont protégés et la mise en place de conditions de concurrence équitables pour favoriser l'innovation, la croissance et la compétitivité, tant au sein du marché unique européen qu'à l'échelle mondiale.

Les obligations pour les fournisseurs de services numérique sont mise en place progressivement de 2022 à 2024 selon le nombre d'utilisateurs du fournisseur

4. La CNIL

En France, la Commission Nationale de l'Informatique et des Libertés (CNIL) est une autorité administrative indépendante française. Elle a été créée par la loi Informatique et Libertés du 6 janvier 1978.

La CNIL est chargée de veiller à ce que l'informatique soit au service du citoyen et qu'elle ne porte atteinte ni à l'identité humaine, ni aux droits de l'homme, ni à la vie privée, ni aux libertés individuelles ou publiques.

En particulier elle s'assure que le RGPD est correctement appliqué par les acteurs du numérique. Elle informe ces acteurs de leurs obligations et les particuliers de leurs droits : https://www.cnil.fr/fr/

4 https://digital-strategy.ec.europa.eu/fr/policies/digital-services-act-package

_

III- Bases d'Internet

Internet est un réseau informatique mondial d'ordinateurs accessible au public

Quelques notions techniques sont nécessaires pour comprendre comment un équipement peut communiquer avec un autre par le réseau Internet.

Voyons cela en s'appuyant sur le schéma montrant les connexions.

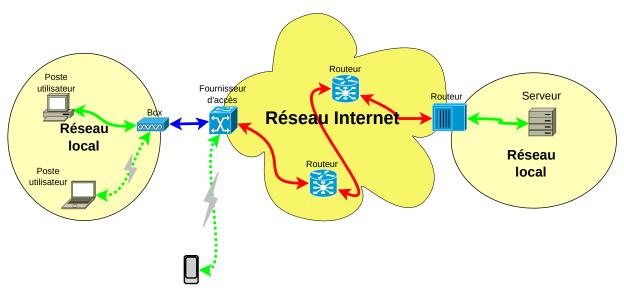


Figure 1: D'un point à un autre en passant par le réseau Internet

1. Adresses IP

Chaque équipement d'un réseau dispose d'une adresse IP unique dans le réseau, c'est ce qui leur permet de se trouver et de communiquer. Il y en a de deux types : IPv4 et IPv6.

L'adresse IP d'un équipement du réseau Internet est associée à la localisation géographique plus ou moins précise de cet équipement.

IPv4

A l'origine d'Internet seules les adresses IPv4 existaient. C'est un nombre de 32 bits, donc 2³² adresses sont possibles soit environ 4 milliards. En fait il y en a moins car certaines sont réservées à des usages spécifiques. Finalement depuis 2011, il n'y a plus d'adresse IPv4 disponible pour le réseau Internet.

L'adresse IPv4 est notée par 4 nombres, compris entre 0 et 255, séparés par des points comme 192.168.1.102

Du fait de la rareté de ces adresses, l'espace d'adressage des réseaux locaux est distinct de celui d'Internet.

Mais ces adresses sont encore (en 2023) très utilisées afin d'assurer la compatibilité avec les anciens équipements.

IPv₆

Afin de faire face à la pénurie d'adresses IPv4. Un nouveau système d'adressage a été conçu au cours des années 1990 : IPv6 standardisé en 2017.

L'adresse IPv6 est sur 128 bits donc 2^{128} adresses disponibles soit plus d'un milliard de milliards de milliards de milliards !

L'adresse IPv6 est notée sur 8 groupes de nombres en 16 bits écrits en hexadécimal et séparés par le caractère ':' comme 2001:0db8:0000:85a3:0000:0000:ac1f:8001 (il existe de notations abrégées).

Comme ces deux types d'adressage sont utilisés aujourd'hui, la plupart des équipements disposent de deux adresses, une en IPv4, une en IPv6.

2. Le chemin des données

Si votre appareil (poste utilisateur) est relié à une box / point d'accès / routeur, par câble ou en wifi, vous êtes dans un réseau local. Votre appareil dispose d'adresses IP locales v4 et v65 et la box également.

Dans la plupart de ces cas, l'adresse IP de votre appareil a été attribuée automatiquement, lors de son démarrage, par la box. Celle-ci est reliée au réseau Internet, donc elle a pour ce réseau une adresse IPv4 (sans rapport avec les précédentes). Dans le cas d'une connexion en IPv6 c'est un peu différent mais le principe est le même.

Si vous utilisez un mobile sur un réseau xG, celui-ci se connecte à Internet par le point d'accès mobile auquel votre opérateur vous a relié qui lui affecte une adresse IP.

Votre connexion doit aboutir finalement à un équipement relié au réseau Internet, généralement un serveur. Comme celui-ci doit vous répondre, il doit connaître votre adresse IP ou celle de votre box.

Bien entendu, tous les équipements intermédiaires ont connaissance de cette adresse IP, de celle du serveur et des données qui passent.

3. Le système de nom de domaine (Domain Name System DNS)

Comme il serait pénible, lorsque l'on veut joindre un serveur de l'adresser par son adresse IP, d'autant que beaucoup de serveurs hébergent plusieurs services (sites web par exemple). Le système de nom de domaine (DNS) joue le rôle d'annuaire.

Par exemple, si vous voulez consulter le site web « fr.wikipedia.org », votre appareil va demander à un serveur DNS l'adresse IP du serveur de ce site, ainsi il pourra s'y connecter. Souvent l'acronyme DNS désigne un serveur DNS.

Votre machine dispose des adresses IP de deux DNS (par sécurité). Dans la plupart des cas ils vous sont attribués par votre fournisseur d'accès Internet (FAI) via votre box, mais il vous est possible de les fixer arbitrairement.

Page: 7

⁵ Les adresses v6 ne sont pas locales, une de ces adresse (« temporaire ») sera transmise sur le réseau.

4. Pour les curieu.ses.x

On peut trouver certaines informations réseau depuis son poste.

Dans la plupart des cas, il faudra passer une commande (tapée au clavier) depuis une invite de commande ou un terminal.

Sous Windows on peut ouvrir l'invite de commande depuis le bouton de recherche (barre des tâches) en tapant « cmd »

Connaître la configuration réseau de sa machine

Sous Windows, commande « ipconfig » ou « ipconfig /all » pour en savoir beaucoup plus. Sous Linux, c'est la commande « ip r »

Trouver l'adresse IP d'un nom de domaine

Taper « nslookup » suivi du nom de domaine.

Exemple: nslookup fr.wikipedia.org

Obtenir des informations sur une adresse IP ou un domaine

On utilise un site web de « whois », par exemple : https://dnslytics.com/whois-lookup

Localiser géographiquement une machine reliée à Internet

- https://www.geodatatool.com/fr/ par l'adresse IPv4 ou le domaine ;
- https://vps.chapellut.fr/outils/geo/?t= par l'adresse IPv4/v6 ou le domaine.

Suivre un chemin

• https://geotraceroute.com/ (à partir de l'adresse IPv4 ou du nom de domaine).

Connaître le trafic réseau de sa machine

En utilisant un logiciel analyseur de trafic comme le bien connu « Wireshark » (libre et gratuit). Cela nécessite toutefois quelques compétences techniques.

IV- Naviguer sur le web

Le « web », en développé le « World Wide Web » : la toile d'araignée mondiale est le service d'Internet le plus connu à tel point qu'il est souvent confondu avec Internet lui-même.

Le web permet à un utilisateur relié à Internet et équipé d'un logiciel « navigateur » de se connecter à des serveurs web afin d'obtenir des informations parfois d'interagir avec celui-ci.

Voici quelques navigateurs parmi les plus utilisés, ils disposent d'une version « ordinateur » et d'une version « mobile ». Certains sont propriétaires, d'autres libres₆

Navigateur	Licence, maintenu par	Notes
Google Chrome	Propriétaire : Google	Le plus utilisé au monde
Safari	Propriétaire : Apple	Pour appareils Apple
Microsoft Edge	Propriétaire : Microsoft	Pour appareils sous Windows
Opera	Propriétaire : Opera Software	
Mozilla Firefox	Libre: Mozilla foundation	
Brave	Libre: Brave Software, Inc.	
Iridium browser	Libre : Iridium	

Les navigateurs de cette liste utilisent tous, à l'exception de Mozilla Firefox et Safari, le même moteur interne « Blink » / « Chromium » ¬(libre).

La configuration du navigateur est modifiable par l'utilisateur.

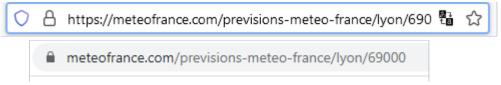
Il est aussi possible d'ajouter des fonctions par la mise en place de modules d'extension.

1. Adresser le serveur

La barre d'adresse, affichée en haut de la fenêtre de navigation, indique le site où est connecté le navigateur ainsi que parfois le chemin de la page concernée.

Cette information est très importante car elle indique le site où l'on est connecté.

Voici deux exemples de barres d'adresses, en haut celle de Firefox, en bas celle de Chrome.



Examinons ces contenus:

panneau de danger.

- Le cadenas indique que la connexion utilise le protocole sécurisé « https » (Firefox indique explicitement ce protocole).
 Si ce n'est pas le cas on aura, selon les navigateurs, un cadenas barré (Firefox) ou un
- 2. « meteofrance.com » est le nom du domaine où l'on est relié. Parfois, on reçoit un message d'un prestataire bien connu, contenant un lien vers le prétendu site du prestataire. En examinant ce lien on voit qu'il n'est est rien, c'est un autre site! Bien vérifier l'orthographe et le nom complet. En effet certains indélicats utilisent le
 - « typosquattage » où le nom de domaine, mal orthographié, ressemble au véritable.
- 3. La suite (pas toujours présente) indique le chemin sur le serveur de la page affichée, avec parfois quelques informations complémentaires.

L'utilisateur peut taper directement dans la barre d'adresse le nom du site qu'il veut consulter, attention alors à l'orthographe.

⁶ Un logiciel propriétaire est créé et diffusé par une entreprise, son fonctionnement est secret. Un logiciel libre est créé et diffusé par un groupement (association, fondation...), son fonctionnement est ouvert à tous. Voir le chapitre « VIII-Applications et services libres ou propriétaires », page :22.

⁷ https://fr.wikipedia.org/wiki/Chromium

Protocoles

A l'origine les sites web utilisaient le protocole « HTTP » (HyperText Transfer Protocol) assez simple dans sa mise en œuvre.

Mais comme certaines malversations ont été réalisées sur les noms de domaine, des sécurités supplémentaires ont été mises en place avec le protocole « HTTPS », en particulier :

- le site web doit produire un certificat d'authentification généré par une autorité tierce habilitée, ainsi il est vérifié que le domaine indiqué est authentique ;
- les échanges entre le navigateur et le serveur sont chiffrés, donc secrets à tout intermédiaire entre le navigateur et le serveur. Cependant la requête préalable du DNS n'est pas nécessairement chiffrée, votre opérateur (et votre DNS) connaît ainsi les sites visités.

HTTPS est très important car se fier au nom du domaine affiché, bien que nécessaire, n'est pas suffisant. En effet certains postes utilisateurs peuvent avoir été infectés et utilisent des DNS « menteurs ». D'autre part, comme on peut avoir dans l'adresse tout types de caractères, il se peut qu'un domaine utilise dans son nom le caractère 'a' de l'alphabet cyrillique (codé 1072) donc différent du 'a' latin (codé 97) mais très ressemblant (entre autres exemples).

Aujourd'hui le protocole HTTPS est utilisé par presque tous les sites web.

2. Échanges entre le navigateur et le serveur

La requête faite au serveur

Le navigateur demande au serveur de lui envoyer un document.

Il lui communique nécessairement son adresse IP pour recevoir la réponse, et dans sa requête il lui donne d'autres informations dont les principales sont :

- le chemin serveur du document demandé;
- le nom de domaine du site car le serveur peut en gérer plusieurs (host);
- la liste des langues souhaitées et les types de données attendues ;
- l'indication du système de l'appareil et celle du navigateur (user-agent);
- l'adresse complète de la page où se trouve le lien menant à cette requête (referer) s'il existe ;
- la demande de non pistage (DNT) si présente dans la configuration du navigateur₈;
- les cookies éventuels.

Dans le cas de formulaires, les informations saisies sont généralement transmises au serveur.

La réponse

Une fois, la requête reçue et traitée, le serveur envoie au navigateur le document demandé (précédé de quelques informations techniques).

Le document peut être de différents types (texte HTML, image, script, etc.). C'est au navigateur de gérer cela, mais ce sera souvent une page de texte en langage HTML.

La page texte HTML

Elle est en texte clair lisible et compréhensible pour qui connaît l'HTML (HyperText Markup Language).

Elle informe le navigateur sur la structure de la page (mise en forme, présentation), elle contient les textes qui sont à afficher, d'éventuels « scripts » (éléments de programme pouvant être exécutés) et des liens vers d'autres éléments comme les images à afficher ou d'autres scripts.

En effet de nombreux éléments ne font pas partie de la page mais seront demandés par celle-ci. Ces éléments peuvent être sur d'autres sites.

Page: 10

⁸ Cette demande est proposée dans la configuration du navigateur, il n'est pas certain que le site la prenne en compte.

3. Les cookies et le stockage du navigateur

Les pages sont indépendantes et ne peuvent pas communiquer facilement entre-elles. Pour remédier à cette difficulté, les cookies (témoins) ont été créé dès 1994.

Un cookie est une information textuelle enregistrée dans le navigateur, elle est de la forme nom=valeur Il y a d'autres informations associées telles que :

- une date d'expiration au-delà de laquelle le cookie sera effacé₉ ;
- le domaine sur lequel le cookie est disponible.

Dans la plupart des cas la valeur du cookie est juste un identifiant indexant un groupe d'informations résidant sur le serveur.

Le cookie peut être placé directement par le serveur (dans les données d'en-tête du protocole) ou par un script de la page.

De même le cookie peut être lu par le serveur (lors de la requête) ou par un script de la page.

Certains cookies sont nécessaires aux sites où un suivi de page en page de l'utilisateur est indispensable. C'est le cas des sites nécessitant une authentification de l'utilisateur ou les sites de vente en ligne (panier). Ce sont des cookies de session qui seront détruit automatiquement à la fermeture du navigateur.

Les cookies « persistants » peuvent aussi être utilisés pour reconnaître l'internaute de visites en visites, à condition qu'il utilise toujours le même navigateur sur le même appareil.

Cookies tierce partie

On a vu qu'une page web peut contenir des liens vers des éléments externes qui seront chargés par la page. Ces éléments peuvent être sur un autre domaine et contenir des cookies, ils connaissent aussi le « referer » (page appelante). Ainsi cet autre domaine (invisible) sait ce que regarde l'internaute.

Autres stockages du navigateur

L'historique de navigation est accessible à l'utilisateur du navigateur. Il n'est pas lisible par les scripts₁₀ ni par le serveur.

Il convient cependant d'être prudent si la machine est utilisée par plusieurs personnes.

L'évolution du web fait que les navigateurs d'aujourd'hui disposent de mécanismes de stockage supplémentaires plus puissants que les cookies dont une base de données locale. Mais cela n'est accessible qu'aux scripts du site (qui peuvent aussi communiquer avec un serveur).

FingerPrinting

Il s'agit d'une technique plus élaborée et difficilement dé-jouable. En français « empreinte (digitale) ».

Il s'agit d'obtenir, notamment via les scripts, un maximum d'informations concernant le navigateur et son environnement voire le comportement de l'internaute. Ce qui fait que l'on peut être suivi même si on efface ou refuse les cookies.

Certains navigateurs et extensions visent à minimiser ce risque.

On peut tester par https://amiunique.org/fr/fingerprint (Université de Lille)

Voir: https://www.cnil.fr/fr/cookies-et-autres-traceurs/regles/cookies-solutions-pour-les-outils-de-mesure-daudience

10 Mais un script peut avancer ou reculer dans l'historique, voire y ajouter des éléments.

⁹ En France la CNIL demande à ce que cette date n'excède pas 13 mois,.

4. Même là où on ne s'y attend (peut-être) pas

CAPTCHA et reCAPTCHA

Certains sites web souhaitent s'assurer que l'internaute est bien un être humain et pas un robot. Pour cela ils invitent l'internaute à répondre à un test que (théoriquement) seul un humain peut (facilement) résoudre, c'est le CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart). Il peut s'agir de demander à recopier un texte présenté distordu ; de répondre à une question mathématique ou de culture générale ; identifier des éléments sur une image.



Figure 2: Source: : Wikimedia Commons

Comme de tels tests sont compliqués à produire certains se tournent vers la sous-traitance et incorporent dans la page un élément externe fourni par un prestataire, généralement il s'agit de Google.

Effectivement, celui-ci fourni gratuitement des tests très élaborés « reCAPTCHA ». Mais au-delà du test en lui-même ce dispositif permet à Google d'entraîner ses applications d'intelligence artificielle.

De plus, l'insertion de cet élément sur la page permet à Google de récupérer beaucoup d'informations concernant la navigation de l'internaute sans que celui-ci en soit informé₁₁.

Se connecter avec...

Sur certains sites web, on nous propose de se connecter (fidélisation?) et parfois cette proposition est associé à la possibilité de se faire authentifier par un autre prestataire : Google, Facebook, etc.

Même si l'on ne se connecte pas ainsi, dans beaucoup de cas l'encart le proposant est externe au site et est situé chez le prestataire concerné. Ainsi, il a suffit d'ouvrir la page du site d'origine pour que ce prestataire soit informé de ce que l'on voit.

5. Alors, pourquoi est-on tracé en navigant sur le web?

La plupart des sites web tiennent des statistiques sur les accès qui y sont faits.

Cela permet à leurs gestionnaires d'avoir des informations utiles sur les pages les plus/moins vues, sur la localisation des internautes, sur la configuration dont ils disposent (systèmes, navigateurs, versions) et sur l'origine des visites. Ainsi ces sites peuvent être optimisés et d'offrir une meilleure expérience utilisateur. De plus, comme l'obtention de ces statistiques peut être déléguée à des prestataires externes, ceux-ci on une vision globale des équipement et agissements des internautes.

Comme beaucoup de sites sont rémunérés (ou en partie) grâce à la publicité, ils incorporent beaucoup de cookies tiers (parfois plus d'une centaine). Ces cookies tiers sont produits par des régies publicitaires et sont implantés sur de nombreux sites. Ces régies ont ainsi une bonne connaissance des centres d'intérêt des internautes.

Par la suite, ces régies peuvent placer sur les sites consultés, la publicité adaptée à l'internaute qui, si elle est cliquée, rémunérera le site.

Certains sites, vous connaissant, pourront vous présenter des contenus plus en rapport avec vos centres d'intérêt. Ils vous placent ainsi dans une « Bulle de filtres ».

¹¹ Voir : « Bien sûr que je ne suis pas un robot ! Quoique... » sur : https://www.editions-legislatives.fr/actualite/bien-sur-que-je-ne-suis-pas-un-robot-quoique.../

6. Gérer l'utilisation et la configuration du navigateur

Si l'on veut minimiser les risques de suivi, la solution la plus simple est de passer en navigation privée, de préférence avec un navigateur libre.

Dans ce cas, l'historique de la session, les cookies et autres informations déposées seront effacés à la fermeture du navigateur.

Plus généralement on peut changer la configuration du navigateur. Voici quelques informations pour Firefox (version 120) PC, pour les autres navigateurs, l'équivalent (ou presque) existe.

A partir du « grille-pain » en haut à droite, cliquer sur « Paramètres », dans le menu de gauche sélectionner « Vie privée et sécurité ». La partie droite de la fenêtre permet la configuration de cette partie. Voici quelques rubriques utiles

Protection renforcée contre le pistage

Firefox contient des possibilités de blocage des traceurs. On peut utiliser la partie « Personnalisée » pour définir ce qui doit être bloqué.

Envoyer aux sites web un signal « Ne pas me pister »

C'est toujours utile mais parfois pas efficace.

Cookies et données de sites

On peut ici effacer tout ou partie de ce qui a été enregistré. Éventuellement provoquer l'effacement à chaque fermeture du navigateur.

Identifiants et mots de passe

Donne la possibilité au navigateur de conserver les identifiants et mots de passe des sites.

Dans ce cas, on peut connaître ce qui a été enregistré.

A utiliser avec circonspection, éventuellement établir un mot de passe principal.

Historique

Gère le fonctionnement de l'historique de navigation.

Permissions

Par défaut, lors de l'accès à un site qui demande une utilisation particulière (caméra, notification...), le navigateur demande cette autorisation à l'utilisateur.

On peut ici, par type d'autorisation, interdire ou autoriser systématiquement et par sites, cet accès.

Sécurité

DNS via HTTPS

Si activé, Firefox fera la recherche DNS via une connexion chiffrée. Il est aussi possible par là d'indiquer quels DNS doit utiliser Firefox.

7. Quelques extensions utiles

Bien que certains navigateurs disposent en interne de mécanismes de filtrage et de protection, il est possible d'aller plus loin en leur installant des modules d'extension à cet effet.

Cependant vérifiez la qualité de ces extensions car certaines peuvent faire plus de mal que de bien. Attention aussi, si votre machine est un peu juste en performance, vitesse, mémoire disponible, car les extensions ralentissent (un peu) la machine et occupent de la place en mémoire. Dans ce cas, il convient d'agir avec parcimonie.

En voici quelques-unes réputées sûres et assez connues. Notez que selon le navigateur, certaines peuvent être indisponibles.

Blocage des publicités

Les publicités nuisent à notre attention sur le web, elles consomment de l'énergie, de la bande passante et elles contiennent très souvent des « traqueurs/pisteurs » pour nous suivre.

L'extension la plus connue est certainement « Adblock », ce n'est probablement pas la meilleure. « uBlock Origin » est un bloqueur libre, efficace et il utilise peu de ressources.

Les bloqueurs de publicité peuvent être détectés par les sites web. Certains vous en avertissent, voire restreignent leurs contenus.

Blocage des traqueurs

Pour limiter notre suivi sur le web on peut se tourner vers « Privacy Badger » conçu par la « Electronic Frontier Foundation (EFF) ».

Blocage des scripts

Bien connue « NoScript » est très (trop ?) efficace mais nécessite une configuration assez fine car beaucoup de scripts sont indispensables au bon fonctionnement des pages web.

Et encore...

« **I don't care about cookies** » vous évite, sur la plupart des sites, à ne pas avoir à répondre au message relatif aux cookies qu'ils déposent.

Contrôle du « referer », ces extensions modifient ou cachent aux sites des liens internes, l'adresse de la page qui les contient.

« **Lightbeam** » cette extension ne bloque rien, c'est juste pour voir. Elle affiche, sous forme graphique, tous les liens entre les sites que vous avez parcourus et ceux qu'ils ont contactés.

8. Naviguer incognito

Bien que l'on ait vu quelques moyens de se protéger des intrusions il reste que lorsque l'on se connecte à un site, celui-ci connaît nécessairement notre adresse IP, il en est de même sur tout le chemin qui y conduit (notamment notre fournisseur d'accès) avec aussi la connaissance du site que l'on consulte et des informations qui y sont échangées si celles-ci ne sont pas chiffrées.

Certains voudraient une meilleure protection, en voici deux.

Utiliser un VPN

Un VPN (virtual private network) est un réseau privé virtuel. Il assure la confidentialité des échanges entre l'adresse de départ et la sortie du VPN (en cachant votre adresse IP et en chiffrant le trafic).

Dans sa version la plus simple, il s'agit de passer par un serveur intermédiaire. Ainsi, lorsque vous voulez atteindre le serveur S vous faites la demande au serveur du VPN V. Votre fournisseur l'accès et le site S ne verront que V. En général, les échanges entre vous et V seront chiffrés.

Plusieurs sociétés proposent ce type de service par abonnement payant.

L'utilisation d'un VPN n'est pas seulement restreinte au web mais concerne tous vos échanges sur le réseau.

Notons que dans ce cas c'est le VPN qui connaît tous nos échanges.

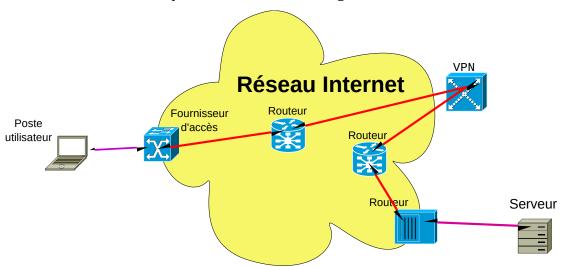


Figure 3: Chemin des données via VPN

Les VPN sont généralement payants, mais certains navigateurs (Opera) proposent l'accès au web par un VPN gratuit (et peu performant).

D'une façon plus générale et moins simple le VPN est un réseau virtuel transitant par Internet

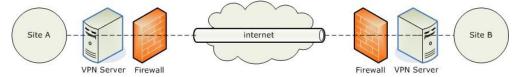


Figure 4: Source: : Wikimedia Commons

Cela est souvent utilisé par les entreprises afin que leurs collaborateurs distants puissent se connecter à leur réseau local depuis des lieux peu sûr.

Réseaux superposés

Il s'agit de réseaux fonctionnant « au dessus » du réseau Internet tout en s'appuyant sur celui-ci. Les plus connus sont Tor, I2P et Freenet.

Exemple « Tor »

C'est probablement le plus connu, Tor « The Onion Router » : routage en oignon, utilise plusieurs intermédiaires pour rejoindre le site demandé et par des communications chiffrées.

La figure ci-dessous montre le chemin des données pour joindre « Diti » depuis « edhelas ». En vert : chiffré, en rouge : clair.

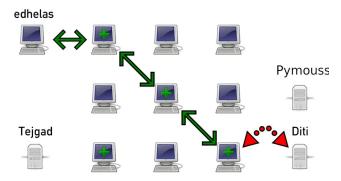


Figure 5: Source: : Wikimedia Commons

Ce réseau dispose aussi de sites web internes (domaines se terminant par « .onion » inaccessibles par le web standard.

Sans rien avoir à installer on peut accéder au web par Tor en utilisant le navigateur spécifique « Tor Browser » ou plus simplement avec le navigateur « Brave » qui dispose de cet accès.

Pour mieux visualiser

Le site https://www.eff.org/fr/pages/tor-and-https propose une infographie afin de visualiser les informations qui sont visibles ou cachées sur le chemin allant de son ordinateur au destinataire selon l'utilisation ou pas de HTTPS et TOR.

V- Utiliser la messagerie électronique

Comme on peut s'en douter, la messagerie électronique (e-mail ou messagerie instantanée) peut être la source d'intrusion par les messages reçus.

Comme les messageries électroniques utilisent Internet, ce qui a été dit précédemment est toujours vrai en particulier si on utilise la messagerie par un navigateur.

Lorsque les messages ne sont pas chiffrés (cas général), ils sont connus du prestataire de messagerie (lu par ses robots), donc sont potentiellement une atteinte à la vie privée.

1. Risques auxquels la messagerie nous expose

Du moins grave au plus grave

- Chaînes: bien que de plus en plus rares, inutiles, gênantes et consommatrices de ressources. Parfois la diffusion peut être problématique. On a vu des messages cherchant des donneurs pour soigner une personne malade dont l'affection n'est pas datée, bien que cela fut exact, il faut savoir que cette propagation ne s'arrêtera (peut-être) que bien longtemps après l'envoi du message originel et, qu'alors, elle n'a plus de sens.
- **Canulars** (hoax) : il s'agit de messages propageant des informations douteuses. Souvent envoyés par nos connaissances, ils nous incitent à les propager. Si on a des doutes, on peut tenter de vérifier l'information et selon le cas en informer l'expéditeur. Dans tous les cas il est préférable de ne pas les retransmettre.
- **Indésirables** (spam) : ce sont souvent des messages publicitaires, non sollicités. Bien qu'ils ne semblent pas poser de problèmes, ils sont pénibles s'ils sont nombreux. Certains peuvent contenir des éléments destinés à nous pister. Si un désabonnement est proposé, il n'est pas conseillé de le faire car cela indique à l'expéditeur que le message a bien été reçu, donc que l'adresse est bonne.
- **Hameçonnage** (phishing) : c'est plus ennuyeux, il y a plusieurs cas.
 - ✓ L'expéditeur usurpe l'identité d'un service que (peut-être) on utilise. Par ce moyen, il espère que nous lui donnerons des informations confidentielles.
 - ✓ L'expéditeur se fait passer pour une de nos connaissances nous demandant un service.
 - ✓ Le message nous informe que notre ordinateur est infecté. Dans ce cas il nous propose la réparation (souvent via un numéro de téléphone) ou il s'agit d'une tentative de chantage.
- **Malveillant** (malware) : le message contient un élément (tel que pièce jointe) exécutable qui contient du code malicieux. Si cet élément est ouvert, une modification du fonctionnement de l'ordinateur est tentée. Le plus souvent il s'agit de :
 - ✓ perturber le fonctionnement de l'ordinateur, éventuellement panne (virus traditionnel) ou faire afficher des annonces publicitaires ;
 - ✓ récupérer des données dans l'ordinateur (spyware) ;
 - ✓ permettre la prise de contrôle à distance de l'appareil (porte dérobée, cheval de Troie) ;
 - ✓ chiffrer les données de l'ordinateur puis proposer une solution payante de déchiffrement (rançongiciel / ransomware).

De plus, l'élément malveillant va essayer de se propager aux autres machines du réseau et à d'autres par l'envoi de messages du même type.

2. Tenter de se protéger

Tout d'abord, la plupart des prestataires de messagerie analysent les messages afin de détecter les « failles » potentielles. Parfois, ils en font un peu trop et catégorisent à tort comme indésirables certains messages (faux positifs), voire même ne les transmettent pas !

Cependant bien d'autres passent à travers ces filtres, il convient donc de rester vigilant.

Pour limiter les risques, il est utile de ne pas donner son adresse électronique inconsidérément, notamment à certains prestataires (il y a quand même des sérieux). Comme parfois c'est nécessaire, il est utile d'avoir une adresse de messagerie supplémentaire prévue pour cela.

Notons aussi que le prestataire à qui l'on donne son adresse (et quelques autres informations) peut se faire pirater, ces informations peuvent être publiées ou communiquées à des indélicats.

On peut tester son adresse sur: https://haveibeenpwned.com/

Une adresse électronique affichée en clair sur un site web sera vue par énormément de robots dont le but est la collecte de ce type d'information.

Réception

La première réaction à avoir lors de l'arrivée d'un message et avant de l'ouvrir est de regarder qui est l'expéditeur et quel est le sujet, la présence de fichier joint peut être aussi un signal. Cela donne déjà quelques indications sur la nature du message.

Si l'on a des doutes sérieux il est préférable de ne pas ouvrir le message, éventuellement le supprimer.

En ce qui concerne l'expéditeur, quelques clients de messagerie n'affichent que son nom et pas son adresse électronique, or c'est cette dernière qui est importante.

Bien que ce soit de plus en plus difficile, certains arrivent à usurper une adresse d'expéditeur.

Lecture

Beaucoup d'expéditeurs indélicats incorporent dans leurs messages un lien interne vers un site qui permettra de vérifier que le message a été lu, par exemple pour afficher une image, c'est une forme d'accusé de réception invisible. Voici un extrait du code HTML d'un message reçu :

```
<img
src=3D"https://e.customeriomail.com/e/o/eyJlbWFpbF9pZCI6ImRnVGg3d0VCQU=F
BQUFZdzFDUW91aVJ2cWxtN1gtUDZpLXc9PSJ9" style=3D"height: 1px !important;
ma=x-height: 1px !important; max-width: 1px !important; width: 1px !
important;= display: none !important;" alt=3D""
/>
On voit que l'adresse de l'image contient une suite de caractères étranges (indexant l'adresse
```

du destinataire), que la taille de l'image est de **1 pixel** et qu'elle ne doit pas être **affichée** !

Quelques clients de messagerie (Thunderbird) peuvent bloquer ces liens.

Pour protéger votre vie privée, Thunderbird a bloqué l'affichage du contenu distant dans ce message.



Bien sûr si le message propose de consulter un site web, c'est le même cas, notamment si l'adresse proposée se termine par une suite de caractères bizarres.

Plus généralement, dans les cas où le message propose des liens, il est très utile de voir, avant de cliquer, où mènent ces liens. En particulier vérifier le nom de domaine du site évite bien des hameçonnages. Si le lien est affiché dans le message, il se peut que le vrai lien (si cliqué) soit différent.

Hameçonnage

Attention au contenu, restez réaliste. Voici un message reçu d'une connaissance :

```
Salut à toi
J'ai besoin de solliciter ta faveur.
Comment vas-tu ? J'espère bien ?
Mon tel n'est pas en service.
```

Le mieux est tout de même d'essayer de lui téléphoner, au moins pour l'informer du problème...

Pièces jointes

La plus grande méfiance est de règle. Même si vous connaissez l'expéditeur, son appareil a pu être piraté.

Il est très utile de déterminer la nature de la pièce jointe (par l'extension du nom du fichier) afin de voir s'il s'agit d'un exécutable (programme), certains documents « texte » contiennent des macros qui en font des exécutables.

On peut vérifier la pièce jointe en la chargeant, sans l'ouvrir, et en la transmettant à un site de vérification tel VirusTotal (https://www.virustotal.com12), ce site fait analyser votre document par plus de 50 antivirus!

Signaler

Dans les cas de réception de messages frauduleux, il peut être utile d'en informer les sites spécialisés.

Si cela ne semble pas trop grave utiliser: https://www.signal-spam.fr/

Dans les cas les plus graves, il est possible de signaler à Pharos :

https://www.internet-signalement.gouv.fr

3. Services collecteurs d'adresses

Même si vous êtes très attentif à ne pas divulguer votre adresse électronique inconsidérément sachez que d'autres peuvent le faire pour vous. Voici deux exemples.

S'abonner à une lettre d'information électronique (newsletter)

Beaucoup de prestataires proposent ce type de service. Ainsi il vous informe régulièrement de ses activités et de ses nouvelles.

Même si ce prestataire vous paraît sérieux et fiable, il est possible qu'il sous-traite ces envois à un service spécialisé dans le « emailing », il en existe beaucoup (Mailchimp, Mailjet, Postmark, Amazon SES, ...) qui collectent ainsi une grande quantité d'adresses électroniques. Mais ceux-la sont-ils bien respectueux de la confidentialité des adresses collectées ?

Cartes virtuelles en ligne

Certains services Internet proposent l'envoi de « cartes postales » virtuelle par la messagerie électronique. Là aussi ces services ont accès aux adresses électroniques des expéditeurs et destinataires. En font-ils d'autres usages ?

Naturellement, ce sont des services à éviter. Bien que ce soit un peu plus compliqué, il est préférable d'envoyer ces cartes soi-même.

Malheureusement, certains de vos amis peuvent le faire, ainsi vous recevez une belle carte et votre adresse électronique a été capturée...

4. Logiciels clients de messagerie

Bien que la plupart des internautes utilisent le webmail (par le site web de leur prestataire), certains préfèrent utiliser une application cliente de messagerie.

En effet cela présente des avantages notamment pour ceux qui disposent de plusieurs comptes de messagerie chez différents prestataires.

De plus, certains clients de messagerie permettent un chiffrage des messages de bout en bout ainsi que l'authentification de l'expéditeur₁₃.

Cependant comme toute application, le client de messagerie peut provoquer des fuites de données₁₄ (voir « Les logiciels », page : 21). Il convient de bien le choisir, si possible comme logiciel libre tels Thunderbird (PC) ou K-9 Mail (Android).

En savoir plus sur la messagerie électroniques

Par le site web très détaillé « Arobase.org » : https://www.arobase.org/

¹² Virustotal appartient à Google.

¹³ Certains prestataires de messagerie le proposent aussi.

¹⁴ https://www.it-connect.fr/le-nouvel-outlook-copie-vos-donnees-vers-les-serveurs-microsoft-y-compris-pour-les-comptes-gmail/

VI- Diffusion « volontaire »

1. Services nécessitant une inscription

Dans la plupart des cas où l'on désire obtenir un service en ligne, il faut s'inscrire. C'est notamment le cas du commerce en ligne, des réseaux sociaux, des services payants, etc.

Souvent cette inscription est nécessaire afin d'accorder à l'inscrit, dès la connexion, les droits d'accès au service. Elle permet aussi de suivre l'inscrit.

Pour les européens, le fournisseur du service doit respecter le RGPD, par conséquent il est tenu d'informer l'inscrit de ce qui est enregistré et à quelles fins (voir le «Le RGPD » page : 4). Cela est décrit habituellement dans les Conditions Générales d'Utilisation du service (CGU).

Cependant la plupart des internautes ne consultent pas les CGU (souvent tout en bas de page), de plus il est difficile de savoir si elles sont respectées. Enfin, si l'on veut profiter du service, il faut bien accepter...

Quant au traçage, il est souvent difficile de savoir s'il est fait et d'en connaître précisément les finalités.

2. Diffusion d'informations personnelles

Le déploiement d'Internet et l'arrivée du web 2.0 (web interactif) au début du XXIème siècle a permis au plus grand nombre de produire des contenus dans les messageries électroniques et sur le web, par les blogs, réseaux sociaux et espaces de discussions en ligne.

Ainsi tout un chacun peut diffuser à une communauté, voire au monde entier, ce qu'il veut sans connaissances techniques particulières.

Il est important de prendre en compte que ce que l'on publie ainsi peut être utilisé par certains, plus ou moins bien intentionnés, voire republié autre part.

Même si cette publication est à destination d'un groupe restreint comme dans les cas des réseaux sociaux, le prestataire du service peut évidemment y accéder. Que peut-il en faire ? Ce prestataire peut aussi être piraté.

Attention surtout aux photographies, notamment dans le cas où elles présentent plusieurs personnes. Celles-ci sont-elles d'accord pour être publiées ? Certaines œuvres d'art sont aussi protégées. De plus, la plupart des images prises avec un appareil numérique contiennent des méta-données (EXIF) indiquant les conditions de prise de vue, parfois, la localisation.

Il en est de même des photographies d'enfant, ceux-ci/leurs parents sont rarement consultés pour autoriser ces publications. Ces photos peuvent être détournées et republiées sur des réseaux « douteux » où être utilisées à des fins malveillantes₁₅.

3. Les « Clouds »

Ou l'informatique dans les nuages. Essentiellement il s'agit de services informatiques distants, éclatés dans le nuage qu'est Internet. Bien qu'ils soient très utilisés par les entreprises, le particulier y voit essentiellement un espace de stockage de documents, un agenda, un carnet d'adresses voire la liste des applications installées sur son appareil. Il est aussi possible de synchroniser cet espace avec son ordinateur personnel, tablette et smartphone.

Cela est d'utilisation simple et a un aspect sécurisant, ces informations sont sauvegardées « à l'abri », si je ne peux plus utiliser mon appareil je peux les retrouver.

Mais où sont-elles vraiment? Chez Google (Google Drive), chez Microsoft (OneDrive), chez Apple (iCloud)? Ou ailleurs... Et ces prestataires ont accès à mes informations personnelles, qu'en fontils ? De plus pourrais-je vraiment récupérer ces données hors de ce service ?

Enfin certains ont des problèmes de fonctionnement, disparaissent₁₆, ou son piratés

Certes, il en existe des plus sûrs où parfois les informations sont chiffrées... mais c'est plus cher...

16 Affaire Megaupload: https://fr.wikipedia.org/wiki/Megaupload

¹⁵ https://www.francetvinfo.fr/internet/reseaux-sociaux/vous-postez-des-photos-de-vos-enfants-sur-internet-on-vous-explique-ce-qu-est-lesharenting-et-pourquoi-cette-pratique-inquiete_5976872.html

VII- Les logiciels (applications ou programmes)

1. Le système d'exploitation

Tout ordinateur, y compris mobile, dispose d'un ensemble logiciel destiné à faciliter son utilisation par les applications et permettre à l'utilisateur de le commander. C'est le système d'exploitation.

Les systèmes d'exploitation les plus répandus sont, pour ordinateur : Microsoft Windows (≈ 70 %), OS-X pour Macintosh ($\approx 21 \%$); pour mobile: Android ($\approx 70 \%$) et iOS ($\approx 30\%$)₁₇.

Ces systèmes sont très liés aux entreprises qui les développent : Microsoft pour Windows, Apple pour OS-X et iOS, Google pour Android (bien qu'il existe des versions Android indépendantes de Google) et souvent aussi liés au constructeur de l'appareil.

En pratique, on constate que beaucoup d'informations relatives au fonctionnement de l'appareil sont transmises à ces entreprises. Il est toutefois possible de limiter ces échanges par la configuration du système.

Les systèmes basés sur le noyau Linux (distributions) sont bien plus respectueux de la vie privée mais sont très minoritaires chez les utilisateurs non professionnels de l'informatique.

2. Les applications installées

Bien que le web soit un élément majeur d'Internet, il est de plus en plus supplanté par des applications communiquant directement avec ce réseau.

En effet les prestataires de nombreux services insistent pour que l'utilisateur installe leur application, parfois c'est même la seule possibilité pour accéder au service.

Sur les appareils mobiles, plusieurs millions d'applications sont disponibles et beaucoup nécessitent un accès à Internet.

Si une application peut se substituer à son éventuel équivalent web, cela peut présenter quelques avantages : meilleure adaptation au format de l'appareil, accès plus simple et plus rapide, notifications instantanées...

Mais il y a quelques inconvénients :

- nécessité d'installer l'application qui occupera alors de la place sur l'appareil;
- souvent, fonctionnement en arrière-plan d'où consommation électrique et charge réseau ;
- quasi impossibilité de connaître la nature des informations échangées avec le réseau ; les systèmes de filtrage que l'on peut avoir avec un navigateur sont inopérants ici.

Ce dernier point est important sur le plan de la protection de la vie privée car certaines informations personnelles peuvent être transmises au fournisseur du service tels que la géolocalisation, la liste des contacts, les photographies prises avec l'appareil, etc. Si dans certains cas cela est un avantage, voire le but du service, cela peut être intrusif.

Par exemple l'application Temu (vente en ligne) place sur Android (en 2023) 4 pisteurs et accède à au moins 13 permissions dont la géolocalisation.

Enfin, il faut indiquer que certaines applications ont des aspects malveillants : portes dérobées, chevaux de Troie, etc.

Il est donc nécessaire de les installer depuis des sites ou magasins d'applications réputés sûr. Ne pas utiliser de logiciels contrefaits.

Le site « Exodus » (https://reports.exodus-privacy.eu.org/fr/) permet de connaître les pisteurs utilisés par les applications Android.

VIII- Applications et services libres ou propriétaires

Certaines applications et services sont dits « libres ». Selon la Free Software Foundation, cela signifie que l'utilisateur dispose alors de quatre libertés :

- les utiliser quel qu'en soit l'usage ;
- en étudier le fonctionnement, éventuellement modifier / corriger ;
- redistribuer des copies ;
- distribuer les éventuelles modifications réalisées au bénéfice de la communauté.

Ce sont presque toujours applications gratuites.

Ce qui n'est pas le cas des « propriétaires ». En effet ceux-ci sont généralement associées à un aspect commercial (bien que certains aient une version gratuite) et protégées par les législations relatives à la propriété intellectuelle. Dans ce cas on ne sait pas vraiment si des informations sont transmises au prestataire ni lesquelles.

Logiciels libres

Aujourd'hui sur ordinateur quasiment toutes les possibilités d'applications logicielles existent en version libre, en voici quelques listes (on y retrouve souvent les mêmes logiciels),

ici: https://fr.wikipedia.org/wiki/Liste_de__libres ou là: https://www.jdbonjour.ch/logiciel-libre/

ou encore: https://framalibre.org/

et celles recommandés par l'état français pour ses

services: https://code.gouv.fr/sill/list

Quant aux utilisateurs d'Android, ils en trouveront ici : https://f-droid.org/packages/

Attention

Comme les logiciels libres sont modifiables par n'importe qui, on peut trouver des versions de ces logiciels qui ont été gravement corrompues.

Il ne faut donc pas les télécharger depuis n'importe où mais uniquement sur le site du concepteur (les listes ci-contre y conduisent).

Dans le doute utilisez :

https://www.virustotal.com/qui/home/upload

Services libres

A l'instar des applications, certains services accessibles sur Internet sont libres. Il s'agit de logiciels libres installés sur un serveur dont le gestionnaire donne un accès libre aux utilisateurs.

En voici certains (liste non exhaustive et à trier...) : https://doc.ubuntu-fr.org/liste de services web libres

De son côté, l'association d'éducation populaire Framasoft à l'occasion de sa campagne « Dégooglisons Internet » propose un grand nombre de services libres placés sur leurs serveurs ou ailleurs.

https://degooglisons-internet.org/fr/

Et quelques-uns bien connus :

- encyclopédie avec Wikipédia https://fr.wikipedia.org/
- cartographie avec uMap (Open Street Map) https://umap.openstreetmap.fr/fr/

Mais bien d'autres existent de par le monde, la difficulté est de les trouver...

2. Fidélisation / dépendance des utilisateurs

Les utilisateurs du service lui restent fidèles car non seulement ils l'apprécient mais aussi ils sont intéressés par l'accès au groupe des autres utilisateurs, tel est le cas des réseaux sociaux. De plus certains services devenus très populaires ont éliminés presque toute concurrence.

Donc il est très difficile de quitter « son » service.

Certains services auparavant gratuits deviennent payants une fois le nombre d'utilisateurs fidélisés rendu conséquent. Un exemple significatif est la plate-forme de covoiturage « BlaBlaCar » anciennement « Covoiturage.fr » gratuit avant 2012, qui prélève désormais une commission (progressivement revue à la hausse !).

3. Diversifier les prestataires

Si l'on ne trouve pas d'alternatives libres à ses besoins, il est utile de diversifier les prestataires des services que l'on utilise. En effet certains prestataires proposent un grand nombre de services, plus de 300 pour Google ; Meta avec Facebook, Instagram, WhatsApp, Messenger, Threads ... Ces services sont interconnectés et les informations que nous leur donnons sont croisées afin de mieux nous connaître.

Certains prestataires permettent à l'utilisateur de leurs services de savoir ce qui a été enregistré, voire l'effacer (mais c'est la loi !). Est-on vraiment certain que l'effacement est complet ? Par exemple Google le propose à ceux qui disposent d'un compte Google sur :

https://myactivity.google.com/activitycontrols

Pour Meta (Facebook etc.), il y a une notice ici :

https://www.phonandroid.com/facebook-comment-consulter-ou-telecharger-vos-donnees-personnelles.html

IX- Objets connectés - Internet des objets

Selon le Journal du Net il y aurait plus de 14 milliards d'objets connectés en 2022 dans le monde dont près de 250 millions en France. De plus le secteur est en forte croissance. Au delà des inquiétudes liées à la consommation de bande passante Internet et d'énergie, se pose la question de la conservation et de l'utilisation des données transmises par ces objets.

En effet, beaucoup de ces objets connectés sont notre quotidien tels les smartphones. D'autres ne sont guère plus rares : montres, voitures, appareils ménagers (réfrigérateurs, aspirateurs, ...), domotique (climatisation, volets, alarmes...) etc.

Intrusion dans la vie privée

Quels qu'ils soient, ces objets obtiennent des informations qui nous sont personnelles : l'aspirateur connaît la géographie de notre appartement ; certaines montres connaissent notre santé, nos performances physiques ; le système de géolocalisation (GPS) de la voiture est informé de nos déplacements, etc.

A partir du moment où l'objet est connecté, les informations seront souvent transmises à une entité centralisatrice,. Celle-ci, à part nous restituer ces informations en fait-elle d'autres usages ?

2. Les risques

Diffusion d'informations personnelles

Le fournisseur du service peut donc transmettre les informations recueillies aux régies publicitaires qui pourront ainsi compléter le profil de la personne concernée, mais elles peuvent aussi être transmises à des tiers concernés par l'activité suivie, par exemple les données de santé pourraient être envoyées à un assureur.

Il peut aussi y avoir vol de ces informations, soit chez le prestataire du service, soit directement au niveau de l'objet connecté. En effet, ceux-ci ne sont pas toujours bien protégés contre les intrusions.

Prise de contrôle non sollicitée

Il s'agit aussi de piratage, mais en sens inverse. L'attaquant peut prendre le contrôle de l'objet connecté, dans le cas d'un véhicule ce peut être dramatique.

En savoir plus

Les Français.es face aux objets connectés : quels équipements et perceptions des impacts ?

https://labo.societenumerique.gouv.fr/fr/articles/dossier-les-fran%C3%A7aises-face-aux-objets-connect %C3%A9s-quels-%C3%A9quipements-et-perceptions-des-impacts/

Objets connectés : quels risques pour la protection de la vie privée, et que peut-on y faire

https://theconversation.com/objets-connectes-quels-risques-pour-la-protection-de-la-vie-privee-et-que-peut-on-y-faire-208118

Page: 24

X- Un exemple : Google

La société états-unienne « Alphabet.Inc » est bien plus connue par les nombreux produits et services qu'elle développe et promeut dont les principaux portent le nom de Google.

Comme certains de ces services nécessitent une authentification, celle-ci est faite au travers d'un compte Google. Celui-ci est unique pour tous les services, ce qui montre leur interconnexion. Ce compte utilisé pour un premier service web posera un cookie (expiration 1an) qui sera utilisé pour les autres services web de Google.

Parmi les nombreux produits et services de Google, en voici quatre qui sont très utilisés dans le monde des technologies numériques.

1. Un moteur de recherche

Il est à l'origine de la société. Lancé en 1998, il a rapidement conquis les internautes : selon Statcounter plus de 90 % des recherches mondiales sur le web sont faites par Google, au second rang est Bing, le moteur de recherche de Microsoft avec 3.4 %! C'est grâce à son moteur de recherche, par les revenus publicitaires, que la société Google a pu s'étendre.

Du fait de sa prédominance dans la recherche les gestionnaires de sites web ont tout intérêt à être bien référencés par Google!

Suivi de l'internaute

A chaque caractère tapé dans la zone de recherche, une requête est envoyé à un serveur afin de prédire ce qui pourrait suivre.

La validation affiche la page des résultats de recherche, à ce moment il y a déjà eu plus d'une centaine de requêtes au serveur.

En cliquant sur un des liens affiché, la page commence par envoyer près d'une dizaine de requêtes aux serveurs de Google (dont Youtube) puis affiche la page demandée.

Toutes ces requêtes permettent à Google, puis au site concerné de savoir si ce dernier apparaît souvent dans les demandes de recherche et si son accès est fait par Google (clic).

Les administrateurs de sites web peuvent consulter cela par la « Google Search Console ».

Et les autres?

Bien qu'ils soient loin derrière, il se défendent en mettant en avant, pour certains, leur respect de la vie privée ou la mise à disposition d'une partie de leurs revenus à de nobles causes (ils ont souvent aussi des revenus publicitaires).

Cependant, afficher des résultats de recherche est une chose, se construire une base de données pour pouvoir les afficher avec pertinence en est une autre. La plupart des « petits » moteurs de recherche sous-traitent cette partie à d'autres plus grands, notamment à Bing (Microsoft).

Remarque utile

Aussi efficace qu'il soit, un moteur de recherche ne peut trouver que ce qu'il connaît. Bien que ce soit difficile de l'estimer précisément, on considère que moins de 20 % des pages web sont connues de moteurs de recherche₁₈!

2. Un navigateur web

Autre position prédominante, le navigateur « Google Chrome », créé en 2008, est utilisé en 2023 par près de 65 % des internautes! Devant « Safari » (Apple) qui en a moins de 20 %, les autres sont sous les 5 %.

Bien que ce navigateur soit basé sur un logiciel libre « Chromium » au développement duquel Google participe, Google Chrome n'est pas vraiment libre et ouvert.

Suivi de l'internaute

Google Chrome embarque une pléthore de mécanismes destinés à espionner l'utilisateur. En particulier la nouvelle « Topics API » qui suis le parcours de l'internaute (alternative aux cookies) afin de le catégoriser. Et beaucoup d'autres informations sont envoyées à Google 19.

3. Une messagerie électronique

« Gmail » la messagerie électronique de Google est lancée en 2004. En 2020 Gmail compterait 1,8 milliards d'utilisateurs₂₀ la mettant à la première place des services de messagerie.

Vie privée

Comme le font d'autres prestataires de messagerie, les robots de Gmail lisent les messages, cela serait fait pour filtrer les indésirables. Cependant, il semblerait que le contenu des messages ait une incidence sur les résultats de la recherche Google... Bulle de filtre, chambre d'écho ?

4. Un système d'exploitation pour mobile

Android est un système d'exploitation libre développé partiellement par Google. Cependant la plupart des appareils mobiles utilisent une version « adaptée » par Google et par le fabricant. Près de 70 % des appareils mobiles utilisent Android, \approx 30 % pour iOS (Apple), les rares autres se partagent moins de 1 %.

Vie privée

Les appareils Android adaptés par Google sont équipés de nombreuses applications de Google dont la désinstallation de la plupart est impossible : Google Play, Youtube, Gmail, etc. Celles-ci contiennent de nombreux pisteurs.

A l'installation du nouvel appareil, Google demande l'utilisation d'un compte Google. Bien que cela soit théoriquement facultatif, l'utilisateur est fortement incité.

Ensuite beaucoup de données seront transmises à Google telle la géolocalisation. Bien que la plupart de ces transmissions soient désactivables, cela n'est pas fait à l'installation₂₁.

¹⁹ Voir la bande dessinée : Contra Chrome https://contrachrome.com/comic/688/

²⁰ https://financesonline.com/number-of-active-gmail-users/

²¹ https://www.phonandroid.com/android-comment-piste-google.html

XI- En conclusion

On a vu que certains ont un grand intérêt pour nous connaître et que nous disposons de moyens, techniques et stratégies pour limiter leur appétit.

Cependant les services que nous utilisons présentent face à ces intrusions des avantages, il nous faut établir un équilibre entre ce que nous donnons et nous voulons.

Ce choix nous appartient.

Pour les personnes qui ont de grandes craintes concernant la surveillance et la censure

Dans ce cas, il faut abandonner Windows, MacOS, etc. et s'équiper avec le système d'exploitation libre « Tails » (The Amnesic Incognito Live System) https://tails.net/

Il s'agit d'un système basé sur Linux,/Debian, Internet passe par le réseau TOR et, bien sûr, il n'est équipé que de logiciels libres.

XII- Webographie sommaire

Si certains des liens présentés ici ou dans le reste du document ne sont pas corrects nous vous prions de m'en informer afin que ce document soit corrigé. https://contact.chapellut.fr/

1. Vie privée

Vidéo « Disparaître - Sous le radar des algorithmes »

Documentaire France24 / Arte (avril 2022 45mn).

Nous laissons sur Internet un grand nombre d'informations sur notre vie privée. Mais certains, ayant pris conscience des menaces qui pèsent sur leur liberté trouvent des moyens de passer plus inaperçus.

https://www.youtube.com/watch?v= mXyQQnZIMA

Vidéo « Tous tracés »

Reportage et interview, de « Geo Politis » de la Radio Télévision Suisse (mars 2020 27mn). https://pages.rts.ch/emissions/geopolitis/11047229-geopolitis.html

Article- ► vidéo « La vie privée est-elle morte, ou cherche-t-on à vous le faire croire ? »

Entretien Nextimpact / Arrêt sur images (mars 2014 45mn)

Bien qu'un peu ancien cet interview du sociologue Antonio Casilli reste d'actualité. https://www.nextinpact.com/article/11410/85644-14h42-la-vie-privee-est-elle-morte-ou-cherche-t-on-a-vous-faire-croire

Verbatim- ▶ vidéo « Cloud, vie privée et surveillance de masse »

Conférence de Tristan Nitot sur le site de l'APRIL (juin 2015 – 47mn) https://www.april.org/cloud-vie-privee-et-surveillance-de-masse-tristan-nitot

Bande dessinée - ▶ Contra Chrome

Au sujet de Google Chrome, les informations qu'il nous prend, son évolution, son modèle économique.

Lire en ligne ou témécharger (PDF). https://contrachrome.com/comic/681/

Article « Données personnelles : comment nous avons peu à peu accepté d'en perdre le contrôle »

Sur le site « The Conversation » article de Yoann Nabat enseignant-chercheur (déc. 2023) https://theconversation.com/donnees-personnelles-comment-nous-avons-peu-a-peu-accepte-den-perdre-le-controle-218290

Livre téléchargeable ► « Guide d'Autodéfense Numérique »

A l'usage de ceux qui veulent vraiment aller loin dans la protection de la vie privée

Le « Guide d'Autodéfense Numérique » fait découvrir les failles de l'utilisation de l'informatique et du réseau et les moyens de s'en protéger. Peut être lu en ligne ou téléchargé, il est très complet donc volumineux (434 pages dans la 6° édition de janvier 2023). https://quide.boum.org/

Ce que devrait faire les professionnels pour être en accord avec la loi

Présenté par la CNIL voici le « Guide de la sécurité des données personnelles ». https://www.cnil.fr/fr/principes-cles/guide-de-la-securite-des-donnees-personnelles

Protection des données et respect de la vie privée en ligne

Sur le site de l'Union Européenne, résumé des droits des utilisateurs agrémenté de courtes « histoires vécues ».

https://europa.eu/youreurope/citizens/consumers/internet-telecoms/data-protection-online-privacy/index fr.htm

La Quadrature du Net

La Quadrature du Net promeut et défend les libertés fondamentales dans l'environnement numérique. L'association lutte contre la censure et la surveillance, que celles-ci viennent des états ou des entreprises privées. Elle questionne la façon dont le numérique et la société s'influencent mutuellement. Elle œuvre pour un Internet libre, décentralisé et émancipateur.

https://www.laquadrature.net/

2. Logiciel libre

Qu'est-ce que le logiciel libre ?

Définition et détails des libertés requises. https://www.gnu.org/philosophy/free-sw.fr.html#beyond-software

Pourquoi les logiciels ne doivent pas avoir de propriétaire

Par Richard Stallman (traduction française). Richard Stallman a été l'initiateur du mouvement du logiciel libre en 1984 en lançant le projet GNU (GNU's Not Unix). https://www.gnu.org/philosophy/why-free.fr.html

Free Software Foundation (FSF) site Europe

Pour la promotion du logiciel libre, le site Europe de la FSF présente ses activités et les dernières actualités du logiciel libre (la plupart des articles sont en anglais). https://fsfe.org/

Livre (téléchargeable en PDF) « Utopie du logiciel libre, du bricolage informatique à la réinvention sociale »

Par Sébastien Broca, sociologue.

Né dans les années 1980 de la révolte de hackers contre la privatisation du code informatique, le mouvement du logiciel libre ne semblait pas destiné à renouveler nos imaginaires politiques. Les valeurs et les pratiques du Libre ont pourtant gagné d'autres domaines, dessinant peu à peu une véritable « utopie concrète ».

http://linux-ventoux.org/index.php?post/2018/01/11/Utopie-du-Logiciel-Libre (édition de 2013, 290 pages) Une édition de 2018 est disponible en librairie (ISBN : 978-2-36935-097-2).

Livre illustré (téléchargeable en PDF) « Ada & Zangemann»

Texte: Matthias Kirschner, dessins: Sandra Brandstätter (Novembre 2023).

Un conte sur l'informatique libre, la camaraderie et le rôle des filles pour une technique au service de l'autonomie.

Peut être téléchargé, commandé en ligne ou acheté en librairie (ISBN 978-2-37662-075-4). https://cfeditions.com/ada/

3. Et aussi...

CookieViz

CookieViz est un navigateur/application libre et gratuit, développé par la CNIL vous permettant, à partir d'un parcours du web, de voir sous forme graphique les sites sur lesquels vous avez été connectés à votre insu.

https://linc.cnil.fr/cookieviz-23-une-nouvelle-version-plus-securisee-plus-stable-et-une-mise-en-avant-du-role-des

La protection « au maximum » : Tails

Tails (The Amnesic Incognito Live System) est un système d'exploitation portable qui protège contre la surveillance et la censure.

https://tails.net/

Qu'est-ce qu'Internet ? - Cycle de conférences à Sciences Po

Trois conférences en vidéo (qualité moyenne) faites par Benjamin Bayart, ancien président de l'association « French Data Network » (FDN) en 2010 : approche technique de l'internet ; applications ; impacts politiques et sociétaux.

https://www.fdn.fr/actions/confs/qu-est-ce-qu-internet/

Les dangers d'Internet

Vidéo Sensibiliser les jeunes aux dangers liés à Internet Préfecture de police de Paris (2015) https://www.dailymotion.com/video/x2mrfoy

Formations

Formations libres et gratuites proposées par l'Institut national de recherche en sciences et technologies du numérique (INRIA). A installer sur son mobile. https://epoc.inria.fr/

Objets connectés, définitions, risques et protection

Sur le site de la Direction générale de la concurrence, de la consommation et de la répression des fraudes (DGCCRF).

https://www.economie.gouv.fr/dqccrf/Publications/Vie-pratique/Fiches-pratiques/objets-connectes

Wi-Fi publics

Le site de Numérama présente les risques liés à l'utilisation des Wi-Fi publics. https://www.numerama.com/tech/1053366-wi-fi-public-quels-sont-les-risques-pour-vos-donnees.html

Caméras de surveillance

Site de cartographie des emplacements de caméras https://sunders.uber.space/