UCLouvain
FUCaM MONS

LOUVAIN
School of Management

Louvain School of Management

Analyse et comparaison des différents outils permettant de limiter ses traces numériques

Auteur : Arnaud Caldow
Promoteur : François Fouss
Année académique 2018-2019
Master [120] Ingénieur de gestion, à finalité spécialisée

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management
Chaussée de Binche 151, 7000 Mons, Belgique | www.uclouvain.be/lsm

Résumé

Ce mémoire avait pour objectif d'établir une analyse et une comparaison des différents outils permettant de limiter ses traces numériques.

La revue de la littérature a engendré plusieurs hypothèses à tester. Dans un premier temps, le questionnement porte sur l'obligation des utilisateurs à partager des informations personnelles. L'hypothèse suivante est directement liée à la première et tente de savoir si un individu à la possibilité d'avoir un contrôle sur les données qu'il laisse filtrer en ligne. De nombreux spécialistes différencient les divers outils mis à disposition des utilisateurs pour garantir la protection de leurs données. C'est sur cette base qu'il est tenté de déterminer si n'importe quel outil apporte une protection équivalente et s'il faut en privilégier certains.

Afin d'obtenir des éléments de réponses aux différentes hypothèses posées, il a été choisi d'établir une analyse en 3 étapes. Dans un premier temps, les systèmes d'exploitation les plus populaires ont été étudiés pour comparer leurs différentes options de confidentialité. Il s'est avéré que Linux soit la meilleure option, même si Windows et MacOS offre des fonctionnalités de protection d'information non négligeables.

Cette première étape terminée, nous sommes passés à l'étude des navigateurs les plus utilisés sur le Web. L'approche était, cette fois-ci, un peu différente. Une comparaison des performances de sécurité des navigateurs a été établie lorsqu'ils étaient configurés par défaut. Nous les avons ensuite comparées lors de leur configuration la plus stricte. On en a déduit que les navigateurs configurés par défaut n'offrent pas de protection aux utilisateurs. Tandis que lorsqu'ils sont paramétrés, Firefox et Chrome sont les options les plus sûres.

La dernière partie de l'analyse s'est concentrée sur les applications proposées aux internautes pour bloquer les différentes publicités et les traqueurs qui peuvent être présent lors de la navigation. Un tableau comparatif contenant les caractéristiques principales des extensions a été créé pour permettre d'en établir un classement. Ublock Origin et Privacy Badger se sont avérés être les meilleures options de protection pour n'importe quel utilisateur. Même si Ublock proposait notamment plus d'options pour des individus possédant des connaissances plus avancées

Il ressort des résultats qu'un utilisateur peut choisir de ne partager aucune donnée, pour autant qu'il utilise l'outil approprié. Même s'il reste toujours identifiable via l'empreinte unique de son navigateur. On remarque aussi que l'utilisateur peut obtenir le contrôle sur certaines des données qu'il laisse filtrer lors de l'utilisation de sa machine. Mais il semble que cela doive parfois se faire au dépend de certaines fonctionnalités proposées. Enfin, il en découle qu'il existe une hiérarchie parmi les outils, et que tous n'offrent pas le même niveau de protection à l'utilisateur.

Remerciements

Il me parait important d'adresser mes plus sincères remerciements à toutes les personnes dont la présence proche ou lointaine a contribué aux différentes étapes qui m'ont mené vers la concrétisation de ce mémoire.

Mes premiers remerciements sont, me semble-t-il, prioritaires, puisqu'adressés au Professeur François Fouss. En tant que promoteur, il m'a été d'une aide précieuse du fait de sa disponibilité ainsi qu'à l'élaboration d'une plus juste formulation de ma question de recherche.

Ses conseils judicieux, tout comme les différents questionnements qu'il suggérait, m'ont permis d'avancer dans mon travail de façon ciblée et de traiter mon sujet de manière objective afin qu'il soit le plus riche et informatif possible.

Je tiens également à remercier l'ensemble des professeurs que j'ai eu la chance de côtoyer durant ces cinq années d'étude. Ceux-ci ont tous contribué, d'une manière ou d'une autre, à ma formation et à la finalité de celle-ci que représente mon mémoire.

Je remercie aussi, Quentin Sellier, pour son aide à la relecture et ses indications contribuant à l'amélioration de ce travail.

Enfin, j'aimerais adresser un merci particulier à mes parents, Michel Caldow et Anne Vandenheede, qui m'ont permis d'entreprendre les études de mon choix et m'ont apporté un soutien permanent en me témoignant leur confiance dans la gestion de celles-ci.

Je terminerais par un petit « clin d'œil » complice à mon grand-père, Jacques Vandenheede, ingénieur, avec qui j'ai souvent échangé, m'amusant de constater la passion qu'il a eu et a encore pour son métier.

Table des matières

Introduction

1. Question de recherche

Ces dernières années, il est de plus en plus fréquent d'entendre parler de traqueurs, de Cookies, et de publicités, présents dans nos navigateurs. En parallèle, des systèmes se sont développés pour contrer ces programmes. Ce sont sur ces divers phénomènes que les analyses de ce travail vont se concentrer. La problématique se formule de la manière suivante : « Analyse et comparaison des différents outils qui permettent de limiter ses traces numériques ». De plus en plus de données personnelles sont récoltées par certaines entreprises lorsque nous naviguons sur le Web. La personnalisation de l'expérience proposée sur internet pourrait représenter une atteinte à la vie privée de certains (Pras, 2012). Effectivement, de plus en plus d'options de confidentialité et de bloqueurs prônant la protection des données privées apparaissent sur nos ordinateurs.

Ce travail analyse et compare les divers outils employés par les utilisateurs pour limiter leurs traces numériques et évalue leurs impacts sur la confidentialité en se basant sur la littérature et diverses analyses. Pour pouvoir établir une classification correcte de ceux-ci, il est nécessaire de connaître les positions de la communauté scientifique sur le sujet et de savoir comment différents experts les caractérisent. Il est primordial d'établir quels critères déterminent l'efficacité d'un outil par rapport à un autre et quels types d'applications existent pour aider l'utilisateur à protéger ses données.

En outre, il est essentiel de définir quelles données sont protégées et quels réglages permettent de maximiser au mieux la préservation des informations. Ce travail de fin d'étude apporte des éléments de réponse à plusieurs questions telles que « existe-t-il des moyens de protection de données accessibles aux utilisateurs ? », « existe-t-il un outil meilleur qu'un autre pour protéger sa vie privée ? », « peut-on réellement empêcher la collecte de nos données en ligne ? », etc.

Il existait divers moyens d'aborder cette problématique sur les outils permettant de limiter les traces numériques. Dans le cadre de ce travail, nous nous limitons aux données récoltées de manière indirecte lors de l'utilisation d'un navigateur et par les systèmes d'exploitation des utilisateurs.

C’est pour cela que nous nous concentrons sur l’analyse de différentes options permettant de bloquer la récolte de données et non de directement anonymiser l’utilisateur. Le but de ce mémoire n’est pas de constituer une liste d’alternatives aux applications utilisées par la grande majorité des internautes mais bien de montrer qu’il existe des moyens de se protéger sans pour autant changer ses habitudes de navigation.

Il est certain qu’il aurait été possible de considérer la question en partant des données communiquées par les utilisateurs, comme par exemple le contenu de mails, messages instantanés, etc. Néanmoins, ces sujets méritent autant d’attention que celui que j’ai décidé de traiter et il n’est donc pas réalisable de tous les étudier en même temps. Cela signifie que l’analyse se limitera aux données récoltées lors de l’utilisation d’un ordinateur, les smartphones pouvant faire l’objet d’une analyse séparée. Ce choix cible essentiellement les données récoltées par des entreprises pour ensuite être traitées et utilisées ou non à des fins commerciales.

2. **Contexte**

De nos jours, internet fait partie intégrante de la vie de nombreuses personnes. S’afficher sur les réseaux sociaux ou partager des informations personnelles ne représente plus un problème pour personne, ce qui peut donner cours à certains débats concernant la limite existant entre ce qui constitue la vie privée et ce qui est publique. Les gens naviguent sur le Web sans nécessairement se douter qu’une myriade de données est collectée via la moindre interaction. Pour comprendre ces informations, il est primordial de prendre conscience de la place prise par internet dans la société actuelle. Avec un nombre d’utilisateurs de 4,38 milliards au niveau mondial et un taux de pénétration de 94% en Belgique en Janvier 2019 (Kemp, 2019) la croissance constante d’utilisation du Web n’est plus à prouver. Il est maintenant possible de faire ou de trouver à peu près tout sur internet. Mais cela ne veut pas dire que la plupart des gens sont conscients de ce qui est fait de leurs informations.

Il n’est pas rare de rencontrer lors de la navigation des publicités dîtes « ciblées » qui peuvent correspondre à ce qu’un internaute aurait mentionné au détour d’une conversation ou bien d’une recherche. Il n’est pas rare non plus d’obtenir des offres pour des activités se trouvant dans

notre région ou correspondant à nos intérêts. En effet, la publicité sur internet représente une source de revenu non-négligeable pour de nombreuses entreprises. Par exemple, lors du premier trimestre de 2018, 85,95% des revenus de Google venait de la publicité (Duffez, 2018), ce qui représentait 26,64 milliards de dollars. Il n'est donc pas étonnant que la quantité de données laissées par les internautes en ligne soit devenue un atout pour de multiples entreprises partout dans le monde.

Dans ce cadre, la possibilité de protéger ses données et de choisir les informations que nous voulons mettre à disposition des entreprises est devenue une des priorités de nombreux organismes comme la Commission Européenne avec l'arrivée de la loi sur la protection des données. Mais selon l'agence internationale We Are Social (Kemp, 2019), seulement 47% de la population mondiale (dont 36% de la Belgique) utilise actuellement un bloqueur de publicité lors de sa navigation.

Chapitre 1 : Revue de littérature

Avant de rentrer pleinement dans le vif du sujet, il est utile d'analyser les différentes sources se rapprochant de la limitation et de l'étude des traces numériques. Nous allons donc faire une revue des différents auteurs ayant abordé ce sujet tout en résumant et présentant le contenu de leurs recherches afin d'obtenir une meilleure compréhension de notre objectif.

Cette revue de littérature se divisera en 3 grands axes. Premièrement, nous aborderons le concept de « Traces numériques » en le définissant et en approfondissant les diverses interrogations que cette nouvelle notion va amener. Nous enchaînerons ensuite sur la sécurité liée à l'utilisation d'internet quant aux données que nous pouvons communiquer. Nous terminerons en abordant la notion d'outil à proprement parler lorsqu'il s'agit de ne pas laisser de traces numériques.

1. Cadre théorique

1.1 Notion de trace

La numérisation de la trace

Pour pouvoir définir des critères de sélections précis et utiliser les outils appropriés à la limitation de nos traces numériques, il convient de bien comprendre ce que la notion de « trace » contient et son origine. Les différents auteurs traitant du sujet s'accordent sur la définition de ces deux concepts. Il semblerait que l'apparition du numérique ait bouleversé la relation que l'Homme entretient avec ses propres traces, et ce bouleversement est perceptible dans les points de vue des experts que nous allons analyser lors de cette première partie.

Pour commencer, il est intéressant de donner une première définition du concept de « trace » avant d'exposer les opinions des différents experts sur le sujet plus spécifique des traces numériques. Un bon point de départ semble être l'introduction au débat proposée dans le cadre d'une conférence sur l'Intelligence Artificielle sur lequel se base la réflexion d'Alain Mille (2013).

La notion de trace est, ici, exprimée comme telle : « En première analyse, nous considérons qu'une trace est constituée à partir d'empreintes laissées volontairement ou non dans l'environnement à l'occasion d'un processus. La trace ainsi construite est inscrite (ou non) dans l'environnement utilisé comme support à la mémoire (en tant que processus). » (Mille, 2013, p. 1).

Sans même se pencher sur l'aspect numérique de la trace, nous retrouvons ici des concepts relativement importants. Tout processus peut avoir comme conséquence l'apparition de traces, que cela soit voulu ou non. Celles-ci pourront ensuite être retrouvées dans l'environnement dans lequel elles ont été laissées.

Mille (2013) nous rappelle ensuite que les empreintes peuvent avoir une nature et une persistance variable dépendant de l'environnement et qu'elles restent identifiables par certains observateurs. Les exemples les plus parlants sont, sans doute, ceux des empreintes digitales ou encore de traces de pas qui permettent d'identifier ou d'obtenir des informations concernant certaines personnes.

Nous pouvons maintenant partir de cette définition pour arriver à la dimension numérique de la trace et tout ce que cela implique. L'avènement du numérique a modifié la manière dont les traces de l'être humain sont traitées et conservées. Mille (2013) insiste sur les caractéristiques propres à l'aspect numérique des traces. Il nous explique que chaque empreinte laissée est codée dans l'environnement informatique afin de lui donner un caractère numérique. Cela implique la possibilité de traiter et modifier ces données tout en pouvant les replacer dans le temps. Cette numérisation des traces et ses propriétés vont créer toute une série de préoccupations qui n'existaient presque pas en dehors du monde informatique.

L'enregistrement systématique des traces

La première préoccupation des experts à propos des traces numériques concerne la connaissance des utilisateurs par rapport aux traces qu'ils peuvent laisser dans l'environnement informatique et l'enregistrement systématique de celles-ci.

Merzeau (2009) nous explique que la trace va créer une signature propre à chacun de nos comportements informationnels, alors que nous ne croyons pas nécessairement donner d'information nous concernant. Elle continue en expliquant que : « L'empreinte numérique, elle, est automatiquement produite à l'occasion d'un calcul, d'un codage ou d'une connexion,

le plus souvent sans que le sujet en soit conscient » (Merzau, 2009, p. 24). En conséquence, elle argumente que de simples actions telle que ; cliquer sur un lien ou encore ; commander un produit en ligne, génèrent une trace qui peut être utilisée ou retravaillée. Elle affirme qu'il n'est plus possible de ne laisser aucune trace car chacun de nos agissements sociaux va être traduit en données. En effet, selon elle, l'hypothèse d'un comportement zéro n'existe pas. Le Web, comme nous le connaissons, pousse à la spontanéité du mode conversationnel pour ensuite enregistrer la moindre de nos paroles.

D'autres auteurs ayant mis en avant l'enregistrement systématique de nos moindres interactions avec le monde informatique, sont Galinon-Mélénec et Zlitni (2013). Dans la même lignée que Merzeau (2009), ces auteurs nous expliquent que nos données numérisées sont soumises à toute une série de traitements, calculs, et transformations ignorés par l'usager. Pour représenter ce concept, ils utilisent la notion de « double numérisé » de l'écriture de l'usager. Ce double permet, selon eux, de nombreuses manipulations grâce à son côté numérique, il pourra donc se trouver entièrement métamorphosé grâce aux divers dispositifs de traitement de l'information existant.

Serres (2012) nous démontre que la trace numérique se démarque de son homologue non-numérique via différentes caractéristiques intrinsèques à son fonctionnement. Selon l'auteur, l'une des principales caractéristiques de cette trace est son intentionnalité. Il nous explique que la trace numérique est « provoquée ». Tous les processus informatiques sont construits de manière à recueillir la trace, grâce à différents systèmes conçus spécialement pour cette tâche. D'après lui, il n'existe pas de trace numérique spontanée, elle est toujours « préconstruite » dans le but d'un usage ultérieur envisagé par l'instance qui va la recueillir. Ensuite, il insiste sur la mémorisation accentuée par sa caractéristique numérique. Tout est sauvegardé, classifié. Enfin, il conclut en argumentant que nous sommes passés de la rareté à l'abondance de traces dû à la traçabilité généralisée provoquée par l'environnement du Web.

Ceci est confirmé par une panoplie d'auteurs dont font partie Rossi et Bigot (2018). En effet, ils affirment que :

> Il est commun de reconnaître que l'utilisation des médias informatisés s'accompagne toujours de la production et du stockage de deux types d'informations : d'une part les informations produites par les utilisateurs […] et d'autre part les informations résultant d'un enregistrement automatique des actions effectuées par les utilisateurs. (Rossi & Bigot, 2018, p. 164).

Ils complètent cela en soulignant le caractère non-intentionnel de la production de données émanant des actions des utilisateurs, celles-ci comprenant des informations telles que ; le temps passé sur un site ou encore les types de contenus consultés.

Privacy

L'une des autres préoccupations qui émane des articles des experts tourne autour du concept de « Privacy », et plus particulièrement celle du monde numérique. L'étude de l'avis de différents auteurs sur cette notion va nous permettre de comprendre pourquoi certains groupes de personnes ont décidé de créer des outils qui permettent de limiter les traces numériques. L'apparition de la loi GDPR, qui vise à protéger les données des consommateurs est un bon exemple de l'intérêt croissant porté à la notion de vie privée, qui semble être menacée par le Web et l'ère des réseaux sociaux.

Ce sujet provoque, un peu plus, la polémique parmi les auteurs causée par ce qu'on appelle le « Privacy Paradox ». Pras (2012) le décrit comme le partage des individus entre l'envie de protéger leurs données personnelles de l'exploitation et le souhait de bénéficier des nouvelles technologies et du style de vie qu'elles permettent d'obtenir. Il existe donc une division quant aux démarches qu'il faudrait prendre pour réguler ou contrôler la collecte et l'utilisation des données selon les spécialistes. L'auteur appuie ce paradoxe en affichant certains chiffres obtenus par *Consumers & Convergence* en 2010 qui précise que 79% des personnes interrogées se souciaient de l'utilisation non-autorisée de leurs données par des entreprises alors que 58% se disaient prêts à autoriser un accès s'ils pouvaient obtenir une réduction de leur coût d'internet mobile.

La polémique provoquée par ce « Privacy Paradox » se caractérise surtout par la différence d'opinion concernant les actions à prendre pour protéger et/ou réguler l'accès et la collecte des données des usagers. En effet, comme le dit Arnaud (2009) : « […] comment considérer que les données sont encore personnelles quand elles sont ainsi exposées ? ». Il appuie ses dires en argumentant que l'utilisateur possède une liberté de choix quant à l'affichage ou non de ses propres données, même s'il faut se demander si les internautes sont conscients du danger relatif au partage de celles-ci. Belanger et Crossler (2018) soutiennent que les individus sont souvent submergés face au contrôle de l'accès à leurs données personnelles. D'après leur enquête menée auprès d'utilisateurs d'iPhone, jusqu'à 28 % des gens interrogés ne prennent même pas la peine de consulter leurs options de confidentialité sur les réseaux sociaux. Certaines de ces personnes

vont jusqu'à avouer ne pas changer leurs options de confidentialité après avoir appris un accès non-autorisé à leurs informations. C'est pour cela qu'un auteur comme Türk (2011, cité dans Pras, 2012) pose la question de la sensibilisation et formation des jeunes par rapport aux dangers que les outils d'information et de communication peuvent représenter pour leur vie privée. Selon Rey (2014), il s'est opéré un basculement de perspective. Cela a pour conséquence le fait que les individus sont, en quelque sorte, devenus des acteurs pour produire, diffuser et consommer des données.

Au-delà de ce qui peut être fait au niveau de l'usager, il existe toute une dimension légale qui a animé de nombreuses discussions ces dernières années avec l'apparition de la loi GDPR. Perriault (2009) aborde le problème en précisant que la difficulté se situe dans la conception que nous nous faisons de l'activité numérique. En effet, si la problématique se résume à un contrôle par l'état, il suffirait de protéger les données personnelles de sociétés et d'intrusions considérées comme malfaisantes. Mais Perriault rejoint Arnaud (2009) en se questionnant sur le caractère délibéré de la production de traces par les utilisateurs. De plus, il nous informe sur les comportements très contrastés des différents Etats concernant les politiques de protection des données personnelles. En effet, il n'est pas rare de voir une comparaison des Etats-Unis avec l'Union Européenne pour ce qui est des lois mises en place à l'égard de l'utilisation des données du consommateur, les Etats-Unis ne possédant pas nécessairement de loi protégeant le citoyen.

Il semble qu'il y ait tout de même une chose sur laquelle tous les auteurs s'accordent. Il n'existe pratiquement plus de délimitation entre la vie privée et la vie publique depuis la création des réseaux sociaux sur internet. Merzeau (2009) l'exprime comme une sphère unique regroupant public, privé et intime où tout va être mis en commun. C'est le Web, selon elle, qui pousse à une spontanéité de la parole qui va ensuite être enregistrée. Perriault (2009) apporte une certaine nuance à ces propos. Selon lui, le rôle d'Internet dans l'augmentation d'expositions détaillées de soi reste une question ouverte. Il insiste en expliquant qu'Internet ne possède pas le monopole du « Tout est permis » et que ce type de manifestations se produisait déjà dans les sociétés anciennes telle que la Grèce Antique. Pras (2012) appuie lui aussi les propos de Merzeau (2009) en mettant en évidence le caractère de plus en plus poreux des frontières délimitant le travail et les loisirs d'un grand nombre de salariés. Ceci est dû, d'après lui, au fait que le salarié lambda devient joignable à toute heure et en tout temps grâce aux nouvelles technologies. Il complète sa pensée en expliquant qu'un même outil peut parfois être utilisé à des fins privées comme professionnelles tout en étant couvert en partie par l'entreprise.

1.2 La sécurité informatique

Avant d'analyser les outils informatiques et leurs utilisations à proprement parler, il est utile d'approfondir le concept de sécurité informatique et son développement, conséquence directe de la vulgarisation des technologies de l'information et de la communication. Cela sera utile afin de comprendre le « pourquoi » et le « comment » de l'apparition des logiciels permettant de limiter ses traces numériques. Il est aussi important de discerner les différentes techniques mises en place afin de proposer à l'utilisateur des logiciels et/ou paramètres qui pourraient lui permettre de protéger ses propres données.

Les défis à venir

Depuis la création des objets connectés et de l'Internet of Things, le domaine informatique voit apparaitre de plus en plus de défis concernant la sécurité des infrastructures et des données. Sfar et al. (2018) définissent l'Internet of Things comme suit : « Le concept d'IoT vise à tout connecter, avec tout le monde, partout et tout le temps ». Différentes technologies ont été développées dans le but de connecter le monde physique au monde virtuel (RFID, capteurs, etc.). Ils continuent en expliquant que la conception de tous ces objets connectés amène toute une série de nouveaux challenges sécuritaires ne faisant pas partie des problèmes rencontrés et de leurs solutions de manière générale. Les auteurs expliquent que la majeure partie des anciens modèles de sécurité semblent être adaptables à l'Internet of Things pour des services basiques comme l'authentification ou la confidentialité. Mais le nombre d'objets connectés entre eux et au réseau pose un réel problème. En effet, un grand nombre d'appareil vont interagir entre eux et il sera très difficile de développer un modèle de sécurité qui pourra convenir à chacun d'eux.

Lopez et al. (2017) appuient cet avis et le besoin de développer des systèmes de sécurité suffisamment efficaces en expliquant que, chaque individu devient la cible de la collecte de données sans même s'en rendre compte vu le nombre d'appareils connectés dans notre environnement proche de façon journalière. Effectivement, selon eux, l'Internet des Objets constitue un scénario différent de celui de l'utilisation quotidienne d'internet où les utilisateurs se doivent de mettre en place des paramètres pour protéger leur vie privée. Vu le contexte relativement récent des objets intelligents et les différents problèmes de sécurité que ceux-ci apportent, il semble important pour les utilisateurs d'être capable de protéger eux-mêmes leurs données personnelles afin de pallier les manquements de ces nombreux appareils connectés.

Le Cloud Computing, qui tisse des liens relativement étroits avec l'Internet of Things, présente aussi toute une série de problèmes liés à la sécurité et à la confidentialité des données. Aujourd'hui, de nombreux utilisateurs stockent leurs données sur le Cloud ou en utilisent les ressources. Il est donc cohérent que certains experts se préoccupent de la fiabilité et de la sécurité de ce moyen de stockage devenu extrêmement populaire. C'est ce qu'ont fait Deyan et Hong (2012) dans un article concernant les problèmes de protection de la vie privée et de sécurité des données dans le Cloud Computing. D'après eux, le principal problème vient du fait que les moyens utilisés pour sécuriser l'environnement du Cloud sont identiques à ceux d'un environnement informatique plus classique alors que le Cloud représente de nouveaux défis.

Dans la suite de leur article, Deyan et Hong (2012) décrivent chaque phase du cycle de vie des données, et les divers dangers pouvant être rencontrés lors de ces différentes étapes dans le Cloud. Voici les différentes étapes et les préoccupations que celles-ci amènent :

1. La génération de données : Il est important pour chaque individu de savoir quels types d'informations sont collectées et de réagir lorsque c'est nécessaire ;
2. Le transfert : Si un transfert de données est effectué entre plusieurs instances ou entre ces instances et le Cloud, il faut absolument s'assurer que la confidentialité et l'intégralité des données soient maintenues ;
3. L'utilisation : Dû à sa propriété d'utilisation multi-entité, les données stockées et utilisées dans le Cloud ne sont généralement pas cryptées, ce qui pose de sérieux problèmes de sécurité ;
4. Le partage : Il faut faire très attention lorsqu'on partage des données avec un utilisateur, car celui-ci peut les partager à son tour sans pour autant demander le consentement du propriétaire ;
5. Le stockage : Les auteurs résument les soucis de stockage sur le Cloud en trois termes. Les données doivent rester confidentielles, garder leur intégrité et être accessibles en tout temps ;
6. L'archive : Il est nécessaire d'avoir un endroit où archiver les données pour qu'elles soient toujours accessibles et intègres sur une longue période ;
7. La destruction : Il faut s'assurer que les données détruites ne soient plus récupérables et ne puissent plus être restaurées par la suite.

Les experts semblent tous s'accorder sur les différents défis et problèmes qu'il reste à relever. C'est pour cela qu'il est nécessaire que l'utilisateur soit capable de protéger ses données à son échelle via divers outils.

Data Privacy

Il existe différentes manières de protéger et/ou cacher ses données. Pour pouvoir comprendre le fonctionnement des différentes applications que nous allons analyser lors de ce mémoire, il faut savoir quels algorithmes et techniques sont utilisés afin de rendre nos données non-identifiables. Lopez et al. (2017) ont divisé les résultats de nombreuses recherches sur les techniques pouvant améliorer la confidentialité de nos données en 6 catégories :

- Les solutions basées sur l'anonymisation : Ces algorithmes ont pour but de préserver l'aspect privé des données pour éviter de pouvoir différencier une personne en particulier d'un groupe ;
- Les Blocks Ciphers : Ce groupe représente des algorithmes de cryptage. On va, par exemple, transformer un texte en un cryptage d'une même longueur en utilisant une clé symétrique (qui restera la même pour l'encryptage et le décryptage) ;
- Les Stream Ciphers : Ce sont des algorithmes de cryptage assez similaire aux Block Ciphers. La différence étant que la clé de cryptage générée de manière aléatoire est de la même longueur de texte que ce qui a été encodé. Les auteurs nous précisent que son utilisation est limitée, notamment à cause du temps requis pour générer la clé ;
- Les fonctions de hachage : L'article précise que ces fonctions ont pour but principal de vérifier l'intégrité des messages, des signatures digitales et des empreintes. L'algorithme va transformer des données d'une taille arbitraire en des données de taille fixe que nous appellerons valeurs de hachage ;
- Les algorithmes de clé publique : Ces algorithmes comportent une clé publique et une clé privée et permettent à n'importe quelle personne d'encrypter des données via la clé publique, alors que seule la clé privée permettra d'ensuite décrypter ces données. Cela signifie que cela ne nécessitera pas d'échange entre les différentes parties et augmentera la sécurité. Seule la clé privée doit rester inconnue des autres utilisateurs ;
- Les générateurs de nombres pseudo-aléatoire : Ces algorithmes sont utilisés pour créer une séquence output imprévisible afin d'encrypter les données.

Kouicem et al. (2018) insistent sur l'importance de la préservation de la vie privée dans l'Internet of Things en argumentant que les données qui sont distribuées par les objets intelligents, sont généralement privées et très sensibles car liées à la vie privée de l'individu possédant l'appareil. Selon eux, les politiques de confidentialité servent à remplir trois objectifs principaux.

Il faut tout d'abord pouvoir garantir l'anonymat des individus pour que leur identité ne puisse pas être reconnaissable par une entité extérieure parmi d'autres individus. La deuxième propriété est liée à l'association d'informations produite à une personne en particulier. Il faut que les politiques de confidentialités ne permettent en aucun cas de relier du contenu à la personne qui l'a produit. La dernière caractéristique est l'absence de traçabilité. Il doit être difficile de pouvoir tracer des actions ou des informations émanant du comportement d'une des entités dans le système.

1.3 Data Collection

Nous sommes, depuis peu, entré dans l'ère du Big Data et il existe désormais de nombreuses techniques pour pouvoir récolter les données des utilisateurs d'internet et en retirer des informations importantes. Nous allons analyser la littérature de différents experts concernant les diverses techniques de récolte et de « tracking » utilisées sur le Web et grâce au Big Data.

Pour pouvoir choisir les outils appropriés afin de protéger et de limiter la collecte de nos données, il est important de savoir comment celles-ci sont récoltées et quelles informations nous donnons lorsque nous sommes connectés. Ce chapitre est important car nous allons essayer de déterminer quels types de données sont bloquées par les outils que nous allons utiliser. Il n'existe pas vraiment de controverse sur le sujet des traqueurs à proprement parler, les auteurs étant assez unanimes sur la question de l'existence de collecte de données.

Le nombre de techniques pour traquer les utilisateurs en ligne étant relativement élevé, nous allons d'abord les classer en différents groupes afin de pouvoir les distinguer selon le type de données et la manière dont ils les collectent. Bujlow et al. (2015) donnent trois catégories principales pour définir les différents traqueurs présents sur le Web :

1. Les traqueurs utilisant les données stockées : Ce sont les plus courants, ils utilisent des données stockées directement sur l'ordinateur de l'utilisateur ;
2. Les traqueurs utilisant les données du cache : Ces traqueurs se servent des données qui vont être stockées temporairement dans la mémoire cache de l'ordinateur de l'utilisateur ;
3. Le « fingerprinting » : Ce terme contient plusieurs méthodes qui se basent sur toute une série de technologies pour pouvoir identifier les utilisateurs, son nom vient de

l'empreinte unique qui peut être constituée de plusieurs valeurs que peut laisser un appareil, un OS, ou un navigateur.

Nous allons tenter de mentionner toutes les techniques pertinentes présentes dans la littérature du sujet, même si celles-ci ne rentrent dans aucune des catégories mentionnées.

Les premières données collectées

Il est important de mentionner la première information que nous rendons disponible lorsque nous surfons sur le Web : l'adresse IP. Hassan et Hijazi (2017) présentent différentes informations permettant de suivre ce qu'un utilisateur effectue comme action en ligne grâce aux données collectées par un navigateur ou par le fournisseur d'accès à internet. D'après eux, notre adresse IP constitue la première donnée qui est analysée par les sites auxquels nous nous connectons. Ils expliquent qu'énormément de sites Web associent systématiquement notre IP à la date et l'heure à laquelle une page est consultée. Cette technique permet de savoir combien de temps un utilisateur a passé sur telle ou telle page.

En effet, lorsque nous surfons sur internet, nos données sont transférées grâce au protocole TCP/IP (Transmission Control Protocol et Internet Protocol). Comme nous l'expliquent Zhou et al. (2018), les réseaux étant basés sur ces protocoles utilisent généralement le transfert de données par paquets. Nos informations importantes sont donc collectées puis séparées en paquets pour être plus facilement transmises. Les auteurs expliquent que ces paquets sont généralement constitués de deux parties distinctes : le titre qui contient les données permettant de guider le paquet et le contenu à proprement parler. C'est grâce à ce titre et aux informations qu'il contient que n'importe quelle entité va pouvoir relier une IP à ce qu'elle a effectué en ligne, pour ensuite relier cette IP à un utilisateur en particulier.

Il existe encore une autre manière de collecter les données basées sur ce qu'on appelle les « Log Files » ou historique en français. Lorsque nous effectuons n'importe quelle action sur notre ordinateur, celle-ci est enregistrée dans un fichier appelé journal qui contient chaque événement avec sa date d'occurrence. Ce fichier donne la possibilité de retracer les différents processus effectués par un utilisateur. Lazar et al. (2017) parlent plus précisément d'un type de journal en particulier, les « Web Logs ». Tous les serveurs internet, email et de base de données possèdent un historique d'événements qui contient des informations sur l'utilisation du serveur en question. Il est plutôt commun que les entrées de l'historique soient analysées pour savoir où un utilisateur s'est rendu et à quel moment.

Bien que contenant toute une série d'informations utile, Zhou et al. (2018) expliquent tout de même qu'il est généralement difficile d'exploiter un historique d'événements de manière manuelle. C'est pour cela que différents processus ont été créés pour filtrer les informations intéressantes du journal.

Le « Cookie »

Le premier type de traqueur largement utilisé est « le Cookie ». Hassan et Hijazi (2017) le définissent comme suit : « Les « Cookies » sont de petits fichiers textes habituellement stockés dans le navigateur du client de l'ordinateur. Ils sont souvent cryptés et contiennent des informations qui distinguent le client de l'ordinateur. ». Les « Cookies » peuvent être divisés en deux catégories principales. Les « Cookies de session » qui ne collectent pas d'information sur l'utilisateur mais permettent, par exemple, d'avoir un panier d'achats. Ces « Cookies sont stockés de manière temporaire et sont effacés dès que l'utilisateur se déconnecte ou ferme son navigateur. Le second type de « Cookies » qui est celui qui nous intéresse ici, est quant à lui permanent.

Estrada-Jiménez et al. (2017) nous donnent plus d'informations concernant la manière dont les « Persistent Cookies » sont utilisés pour personnaliser l'expérience d'un individu. En effet, à chaque connexion, un « Cookie » va stocker des données sur les actions d'un utilisateur effectuées sur une page et va permettre de l'identifier lors de sa prochaine visite. Une partie du contenu de ce « Cookie permanent » va appartenir à une entité externe comme par exemple un service de publicité travaillant avec le site consulté. Ensuite, par le mécanisme de « Cookie Matching », ces services de publicités vont pouvoir relier les différents « Cookies » d'un individu afin de partager ses informations, ce qui va permettre de faire de la publicité ciblée sur certains sites. Ruiz-Martinez (2012) nous informe que la technique des « Cookies » est énormément utilisée pour suivre toutes les activités de navigation d'un individu et nous rappelle que ceux-ci peuvent aussi être volés ou bien manipulés.

Généralement, les « Cookies » sont composés au moins : du nom du site Web correspondant, de l'ID de l'utilisateur ainsi que leur date d'expiration. Ce type de « Cookie » est le plus connu des internautes et est plus communément appelé « HTML Cookie ». Cependant, plusieurs auteurs, dont notamment Bujlow et al. (2015), mentionnent un autre type de « Cookies » sollicité par les sites utilisant Adobe Flash qui permettent de stocker plus d'informations et de rester plus longtemps sur notre navigateur.

Ces « Flash Cookies » ont la capacité de nous suivre sur plusieurs navigateurs et n'ont pas nécessairement de date d'expiration selon Estrada-Jiménez et al. (2017). Cette caractéristique les rend plus efficace que les « Cookies » plus « traditionnels » pour traquer le comportement d'un utilisateur.

Les « ETags » et l'HTML5

Il existe un autre terme qui ressort dans la littérature scientifique concernant la collecte de données, à savoir « ETags » ou tag entité. Nous avons regroupé ici « ETags » et l'HTML5 car cette dernière version du langage HTML offre la possibilité de stocker localement certaines informations de manière permanente, dont les « ETags ». D'après Bujlow et al. (2015), le tag entité donne la possibilité d'identifier l'utilisateur grâce aux informations qu'il contient. Ce tag va ensuite être stocké dans le Web cache et être réutilisé lorsque l'utilisateur se reconnectera à la page. Selon Hassan et Hijazi (2017), les tags entités peuvent servir à suivre un utilisateur en particulier de la même manière que le ferait un « Cookie », leur particularité réside dans la possibilité d'utilisation de ces tags pour faire « réapparaître » les « Cookies » associés à certains sites internet qui auraient été supprimés.

Le « fingerprinting »

Lorsque nous recherchons des informations sur la collecte de données, il existe un terme qui revient souvent parmi les experts car utilisé depuis l'apparition du Big Data. C'est celui de « Fingerprinting » ou empreinte digitale. L'expression d'empreinte digitale est utilisée de manière générique pour répertorier toute une série de méthodes, qui mises ensemble, permettent de nous identifier de manière unique. Estrada-Jiménez et al. (2017) présente cette empreinte comme une « String » contenant des informations sur les paramètres de notre navigateur, par exemple. Lorsque nous nous connectons à une page internet, notre navigateur envoie différentes données concernant le type de police, le navigateur ou encore les extensions ajoutées à celui-ci. Toute cette série d'informations pourrait permettre de nous identifier, même sans avoir à utiliser les « Cookies » d'un site.

Mayer et Mitchell (2012) confirment qu'un site Web est capable d'engranger toute une série de propriétés concernant un navigateur qui permet de l'identifier de manière unique. Pour appuyer cette affirmation, ils présentent les résultats d'un échantillon de 500 000 navigateurs présentés

par Eckersley (2010). 83,6 % des navigateurs ont pu être identifiés de manière unique grâce aux procédés de « Fingerprinting ». Les procédés d'empreinte sont d'ailleurs souvent utilisés par les agences de publicités ou dans le Big Data pour pouvoir associer un utilisateur à un navigateur ou encore à un des appareils qu'il possède.

1.4 **Outils**

L'anonymité et la vie privée sont des sujets que nous ne pouvons plus ignorer dans le paysage des données. Il existe divers outils présentés par différents experts pour pouvoir protéger la confidentialité de ses informations en ligne dont nous allons discuter afin de pouvoir choisir au mieux ceux que nous allons étudier. Au lieu de présenter une liste d'outils précise, nous allons les classifier selon les différents types présents dans la littérature scientifique du sujet. Certains outils que nous allons utiliser ne rentrent pas spécialement dans une catégorie en particulier car ils constituent simplement des versions plus « privées » de leurs homologues normalement utilisés, comme un système d'exploitation par exemple. D'ailleurs, de nombreux outils vont avoir recours à d'autres outils des classes dont nous allons discuter ici.

Les « Proxy »

Lorsqu'il est question de rester anonyme en ligne, l'approche la plus connue et qui semble être la plus simple est d'utiliser un serveur que l'on appelle « Proxy ». D'après Li et al. (2013), un « Proxy » reste la manière la plus facile de naviguer sur internet de manière anonyme. Le serveur « Proxy » auquel un individu va se connecter fonctionne comme une sorte de porte. En effet, au lieu d'utiliser sa propre adresse IP afin de se connecter à un réseau, le client se sert de l'identité du « Proxy ».

Lorsqu'un individu effectue une requête, celle-ci est transmise au serveur « Proxy » qui va ensuite relier l'information au réseau pour faire parvenir le résultat de la requête au client en sens inverse, d'où la notion de porte.

VPN

Un VPN (ou « Virtual Private Network) est un autre type de réseau qui est largement utilisé lorsqu'il s'agit de crypter ses données sur le Web. Selon Longworth (2018), les VPN ont connu une vague de vulgarisation, étant à l'origine, en majeure partie, utilisés par des entreprises. Il explique qu'un VPN est utilisé pour protéger l'accès à des données lorsqu'elles voyagent sur un réseau. Plus précisément, Karuna Jyothï et Indira Reddy (2018) expliquent qu'un réseau privé virtuel va étendre un réseau privé sur un réseau publique. Cette action va permettre d'effectuer des requêtes et obtenir des données comme si nous effectuions les manœuvres directement depuis le réseau auquel le VPN est connecté. Cela présente un avantage car toutes les fonctionnalités et les moyens de sécurité mis en place sur le réseau privé restent d'application lorsque nous utilisons le VPN. Sylvestre (2017) nous permet de compléter nos explications en mentionnant que le VPN peut nous connecter à internet via un serveur distant grâce au chiffrage de la communication. Cela permet de rendre nos données indéchiffrables en cas d'un éventuel piratage.

Bloqueurs

Depuis un certain temps, un nouveau type de protection des utilisateurs a fait son apparition sur le Web. Accessible gratuitement et sans connaissance préalable, les bloqueurs permettent d'empêcher l'affichage de publicités ciblées et proposent de stopper les traqueurs malveillants utilisés par certaines entreprises (Ruiz-Martinez, 2012). Ce type d'outil n'offre pas à l'utilisateur une anonymité complète, mais permet plutôt de limiter les informations qu'il laisse filtrer lors de sa navigation.

Conclusion

Nous avons pu constater que les outils permettant de limiter les traces numériques sont en lien avec de multiples sujets. En premier lieu, le terme de trace est utilisé depuis longtemps déjà et a toujours intéressé l'être humain. C'est pour cela que des auteurs transmettent leurs préoccupations quant à la connaissance de celles-ci pour la plupart des utilisateurs. La naissance du « Privacy paradox » représente bien les inquiétudes que certains experts émettent face à la constante collecte de ces traces.

Par la suite, la sécurité informatique actuelle et ses futurs défis ont été abordés. La notion de « Data Privacy » et les difficultés auxquelles il va falloir faire face ont été étudiés. De ce fait, la majorité des experts s'entendent sur les nouvelles possibilités offertes par le Cloud et sur les problèmes de sécurité qu'elles impliquent. Ce sujet nous a ensuite amené aux multiples techniques utilisées pour collecter les données des utilisateurs. La grande partie des auteurs s'accorde à dire que des informations sont récoltées à divers niveaux et via diverses pratiques.

Pour finir, nous avons constaté qu'il existait différents types d'outils pour protéger ses données en ligne. Certains experts mettent plutôt l'accent sur des logiciels permettant de rendre l'utilisateur anonyme tandis que d'autres préfèrent garder une approche basée sur la protection directe des informations. En effet, certains auteurs pensent qu'il est préférable de naviguer de manière sécurisée plutôt qu'anonyme. Il y a donc une différenciation faite entre vie privée et anonymité.

2. Hypothèses à tester

Grâce à la revue de littérature, nous pouvons dégager certains éléments qui vont nous aider à répondre à la question de recherche. Il existe différentes hypothèses que nous avons la possibilité de tester dans ce travail. Il faut, bien entendu, garder en mémoire que les hypothèses ont un lien direct avec différents outils que nous allons étudier.

Les utilisateurs ne sont pas obligés de transmettre des données

La toute première hypothèse que nous pouvons tester dans ce travail concerne les données qui sont récoltées sur les utilisateurs. Il faut vérifier que le consommateur n'ait aucune obligation de procurer des informations le concernant pour pouvoir bénéficier des fonctionnalités d'un outil ou surfer en ligne. Il va falloir tester les différentes configurations par défaut des outils pour identifier si l'utilisateur est tenu de partager ses données. D'après certains auteurs, les risques de sécurité liés à l'Internet des Choses ne font qu'augmenter (Hou, Qu, & Shi, 2018). Il est donc nécessaire de vérifier si les internautes possèdent des moyens de protection.

Les utilisateurs peuvent avoir le contrôle des données qu'ils laissent filtrer en ligne

Cette hypothèse est directement liée à la première mentionnée. En effet, outre le fait de savoir si un utilisateur peut totalement empêcher la récolte de données le concernant, il est utile d'analyser s'il peut aussi choisir quelles informations il veut fournir. Il faut tester si les outils utilisés donnent la possibilité aux internautes d'effectuer des choix par rapport aux données collectées. Tous les experts tendent à s'accorder pour affirmer qu'il n'est plus possible de surfer sur le Web sans être traqué, nous allons voir s'il existe un moyen de contrôler cela.

N'importe quel outil n'offre pas une protection des données équivalente à l'utilisateur

Notre dernière hypothèse concerne directement les outils que nous allons tester et analyser tout au long de ce travail. Il existe toute une panoplie d'applications et d'extensions disponible sur internet pour se protéger contre les traqueurs et autres publicités ciblées. Notre but est d'analyser si ces protections se valent toutes et si l'utilisateur ne doit pas se soucier des choix qu'il fait lors de leurs installations. Est-ce qu'il existe un outil qui se démarque des autres ? Faut-il privilégier certaines solutions plutôt que d'autres ? C'est à ce type de questions que nous allons tenter de répondre.

Chapitre 2 : Méthodologie

Les lois qui protègent les données des utilisateurs apparaissent seulement. Cette nouveauté implique que de nombreuses personnes ignorent le type d'informations récolté sur elles et ce qu'on en fait. Le nombre d'individus généralement renseignés sur ce sujet ne permet pas d'obtenir un échantillon suffisamment représentatif de la majeure partie de la population. C'est pour cela qu'il semblait plus important de se concentrer sur l'analyse des différents outils et leur comparaison.

Cependant, le terme « outil limitant les traces numériques » reprend toute une série de logiciels et d'applications, possédant tous des spectres très différents. C'est pour cela qu'il a été décidé de se concentrer principalement sur des outils utilisés chaque jour par la grande majorité des utilisateurs.

1. Démarche de séparation des outils en catégories

Le terme « outil », et plus particulièrement informatique dans notre cas, est utilisé de manière relativement étendue. Pour pouvoir effectuer une analyse complète et efficace donnant lieu à une possible comparaison, il faut diviser ceux-ci selon des catégories. En effet, 2 outils peuvent ne pas du tout avoir le même but ou proposer les mêmes fonctionnalités. C'est pour cela qu'il faut déterminer des classes globales à étudier pour permettre une comparaison sur des critères objectifs définissant ces classes.

C'est dans cette optique que la rédaction de ce mémoire va diviser son analyse en trois parties. En premier lieu, une analyse de système d'exploitation est réalisée, suivie d'une analyse de navigateurs, pour se terminer par une analyse des applications mis à disposition des utilisateurs. Cette division s'est effectuée d'elle-même étant donné les différents types d'information récoltées par ces outils et les différents paramètres qu'ils proposent. En effet, un système d'exploitation donne accès à beaucoup plus de fonctionnalités qu'un navigateur ou une application.

1.1 Catégorie des OS

Cette première catégorie reprend les systèmes d'exploitation utilisés par la majorité des utilisateurs. Lorsqu'il est question d'OS en informatique, 3 types d'infrastructures sont généralement mentionnées :

- Windows
- MacOS
- Linux

Pour garantir que chacune de ces infrastructures soit représentées dans ce travail, nous avons décidé d'analyser un système d'exploitation pour chacune d'entre elles. Par soucis d'objectivité, il a été décidé de prendre la dernière version logicielle proposée pour chaque OS, les mises à jour de sécurité occupant une place primordiale dans chaque nouvelle version.

Le choix de cette catégorie a été effectué car c'est la première interaction qu'un utilisateur a avec sa machine, c'est ce qui lui permet d'accéder à toutes les autres fonctionnalités, et c'est ici qu'il peut, pour la première fois, configurer ses paramètres de confidentialité. De ce fait, l'OS va récolter toute une série d'informations sur son utilisateur et sur sa machine pour les transmettre à l'entreprise l'ayant développée.

1.2 Catégorie des navigateurs

Créer une classe contenant exclusivement les navigateurs semblait être une évidence, car ils représentent le portail permettant à l'utilisateur de se connecter au Web. Cela signifie qu'il existe tout un tas de données collectées sur les internautes, avant même que ceux-ci puissent accéder à une page.

Afin de suivre la logique d'impartialité et d'objectivité, le choix des navigateurs s'est effectué de la même manière que pour les systèmes d'exploitation. Ils ont été choisis pour être le plus représentatif possible de la majeure partie des utilisateurs. De ces choix sont ressortis 4 navigateurs différents, chacun préféré par une partie des internautes pour leur fonctionnalité ou parce qu'ils sont proposés avec les systèmes d'exploitation étudié.

1.3 Catégorie des applications

Cette catégorie fut la plus difficile à déterminer. En effet, le terme « application » reprend énormément d'outils ayant des fonctionnalités très différentes. C'est pour cela que dans cette situation, il reprend les extensions disponibles pour les navigateurs qui ont été choisis. L'analyse se concentre donc sur les applications permettant à l'utilisateur de protéger ses données de manière plus optimale lors de la navigation sur le Web.

Les applications les plus utilisées qui rentrent dans cette catégorie sont les nombreux bloqueurs qui peuvent être téléchargés sur les Web store des navigateurs étudiés. Une nouvelle fois, par soucis d'objectivité, le choix s'est porté sur les options les plus populaires pour obtenir des résultats concernant la majorité des utilisateurs.

2. Démarche et méthode d'analyse

En raison de la particularité du sujet et des outils qui doivent être analysés, la méthode d'analyse utilisée doit permettre d'établir une comparaison suivant des critères objectifs pour permettre d'obtenir des résultats concluants. De plus, la division de l'analyse en 3 parties distinctes donne aussi la possibilité d'effectuer des tests plus appropriés en fonction des différentes catégories étudiées. Bien que différentes, l'analyse et la comparaison de celles-ci repose sur l'utilisation de grilles de comparaison. Mais détaillons d'abord ce qui a été fait spécifiquement pour chaque catégorie.

2.1 Classe des OS

Pour pouvoir établir le tableau de comparaison de cette classe d'outils, il a fallu diviser notre analyse en 2 parties principales. La première aborde les paramètres de confidentialité et de sécurité proposés à chaque utilisateur utilisant les différents systèmes d'exploitation étudiés. La deuxième partie, quant à elle, se concentre sur ce qu'il advient des données une fois récoltées. Après ces deux parties analysées, la comparaison pourra être faite via la grille constituée.

2.2 Classe des navigateurs

Comme cela a été dit plus haut, les navigateurs sont les portails qui permettent aux internautes de se connecter à internet. De ce fait, ils sont enclins à d'autres types de collectes de données que les OS. Par leur but et leur installation simple et rapide, beaucoup d'utilisateurs se contentent simplement de télécharger un navigateur pour ensuite l'utiliser sans aucune configuration. C'est sur ce postulat que l'analyse va se baser.

Dans un premier temps, les paramètres offerts par défaut lors de l'installation vont être analysé pour constituer une première grille de comparaison. Une fois cette étape terminée, chaque navigateur va être testé via le test Panopticlick[1] proposé par l'Electronic Frontier Foundation qui se définit elle-même comme : « La toute première association caritative défenseuse des libertés civiles dans le monde digital. » (EFF, 2019). Ce qui va nous permettre de comparer les performances de chacun de nos navigateurs.

Lorsque cette première batterie de tests sera terminée, chacun des navigateurs choisis sera configuré pour respecter au mieux les informations des utilisateurs et garantir la confidentialité maximale lors de la navigation. La toute première grille contenant tous les paramètres présents par défaut sera remplie une nouvelle fois, avec les nouvelles configurations. De la même manière que lors de la première partie de cette analyse, les navigateurs seront confrontés au test Panopticlick pour évaluer leur performance et établir une comparaison.

2.3 Classe des applications

Dans le cadre de cette troisième et dernière partie d'analyse, une approche différente a été choisie. Effectivement, les extensions proposées via les navigateurs sont des applications prêtes à l'emploi, qui requièrent un simple clic pour être installées. La configuration de celles-ci n'entre donc pas en ligne de compte pour une grande partie des utilisateurs. C'est pour cette raison qu'il n'a été question que d'une grille de comparaison comprenant toutes les informations nécessaires pour pouvoir analyser les applications étudiées.

[1] Electronic Frontier Foundation, (2019). *Is your browser safe against tracking ?*. En ligne https://panopticlick.eff.org, consulté le 20 mars 2019 (cf. Figure 23 Annexe 2)

Certains critères présents dans ce tableau ont été collectés via le test de Panopticlick, ce qui permet aussi de mesurer la performance de certaines applications.

2.4 **Grille de comparaison**

Afin de pouvoir établir une comparaison des différents outils choisis dans le cadre de ce travail, le choix s'est porté sur l'élaboration de grilles de comparaison. En effet, la sélection de divers critères à partir desquels comparer nos outils de manière objective va permettre d'effectuer une meilleure analyse et d'obtenir des résultats représentatifs.

Pour créer ces grilles, il a fallu parcourir la littérature scientifique pour établir quels types de caractéristiques étaient généralement utilisés parmi les articles proposant une comparaison d'outils comme Ruiz-Martinez (2012). Cela a été utile pour constituer une base de critères sur laquelle démarrer pour pouvoir ensuite y ajouter notre propre analyse. Vu que les différentes classes d'outils traitent différents types d'informations récoltées, il était aussi nécessaire de construire des grilles de comparaison qui soient appropriées à chaque type d'outil.

Chapitre 3 : Analyse des outils

Lors de ce chapitre, nous traitons des différentes informations et réglages récoltés sur les outils que nous avons analysés par rapport aux hypothèses posées lors de la revue de la littérature. Comme mentionné dans la méthodologie, les outils analysés sont divisés selon 3 catégories distinctes afin d'effectuer différents tests. Premièrement, nous analysons les systèmes d'exploitation et leurs paramètres. En deuxième lieu, nous traitons des navigateurs et des différentes options proposées inhérentes à la confidentialité et à la sécurité des internautes. Enfin, nous jugeons de l'utilité de différentes applications permettant de protéger ses données en ligne.

1. Systèmes d'exploitation

1.1 Choix des systèmes d'exploitation :

Pour que cette étude obtienne des résultats concrets, il était nécessaire de choisir des systèmes d'exploitation utilisés par la majeure partie de la population. A cette fin, nous avons choisi de collecter des données sur les 3 OS les plus répandus sur le marché. Avant de rentrer pleinement dans les résultats obtenus et les différents paramètres proposer par chaque OS, nous allons présenter brièvement chacun de ceux-ci et établir leur carte d'identité. Il existe d'autres systèmes d'exploitation sur le marché mais nous avons décidé d'utiliser les dernières versions de Windows, MacOS, et du build Ubuntu de Linux pour vérifier les différents paramètres proposés par ces OS.

Pour comparer au mieux ces systèmes et nous permettre de les classer en fonction de leur niveau de confidentialité, nous allons procéder en identifiant les similitudes existantes entre les configurations possibles de ces outils avant de présenter quelles sont leurs différences majeures. Grâce à ces informations, nous pouvons établir un tableau avec les points forts et les points faibles de chaque OS du point de vue de la confidentialité et de la sécurité des données des utilisateurs.

Depuis l'instigation de la loi RGPD concernant la protection des données, il est devenu primordial que les entreprises donnent un droit de regard et de modification aux utilisateurs concernant leurs données. Les systèmes d'exploitation étudiés dans le cadre de ce travail ne doivent donc pas déroger à ces règles et devront, de ce fait, présenter certaines similarités dans la manière dont les données des internautes sont traitées. Mais chaque entreprise est libre d'interpréter la loi comme elle l'entend, tant qu'elle répond aux demandes exprimées par celle-ci.

Windows

Avec une part de marché de 88,22% en Avril 2019 selon NetMarketShare (2019), Windows est le système d'exploitation le plus présent sur le marché et fait partie de la famille Windows NT qui caractérise les systèmes d'exploitation multi-utilisateurs créé par Microsoft. Sa dernière version, Windows 10, est sortie en Juillet 2015 de manière gratuite pour ensuite passer à une version payante. L'OS créée par Microsoft peut fonctionner sur différents hardwares et est disponible sous différentes versions telles que la version Familiale ou Professionnelle.

MacOS

Le second système d'exploitation le plus utilisé au niveau mondial n'est autre que l'OS développé et commercialisé par Apple présente sur tous les ordinateurs du constructeur. Ce système appartient à la famille Unix comme les systèmes tournant sur Linux. MacOS représente une part de marché de 9,38%, toujours selon NetMarketShare (2019), et est donc la deuxième OS la plus utilisée au niveau mondial. Elle se démarque de Linux et Windows car ne fonctionne que sur des hardwares vendus par Apple.

Linux

Le troisième système d'exploitation que nous avons choisi d'étudier est un peu particulier car il se décline sous toute une série de builds différents qu'il est possible de télécharger. Nous avons opté pour une des options les plus populaires, dénommées Ubuntu.

1.2 Paramètres de confidentialité

Pour savoir quelles données sont récoltées par les systèmes d'exploitation que nous avons choisis d'utiliser, il faut consulter les paramètres de confidentialité proposés par ceux-ci dans les outils systèmes. Nous allons vérifier si Windows, MacOS, et Linux proposent aux utilisateurs les mêmes paramètres ou s'il existe des différences notoires. Cela nous permet de comparer ces OS et d'en établir un classement.

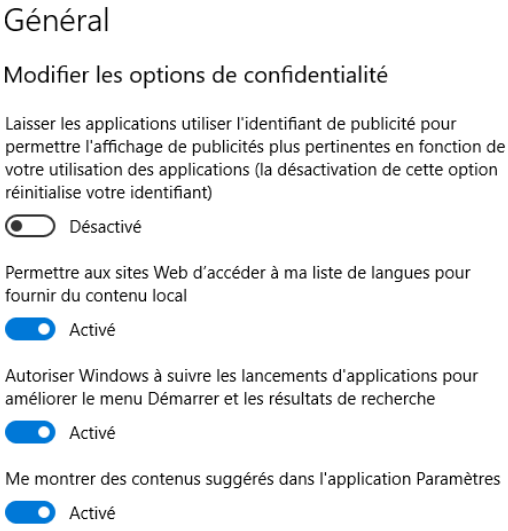
Windows

Les paramètres de confidentialité proposés par Windows 10 se divisent selon deux parties principales. La première concerne les différentes autorisations concernant Windows tandis que la deuxième permet de gérer la communication de certaines informations aux applications utilisées sur notre ordinateur. Il est demandé à l'utilisateur de configurer chacune de ces options durant l'installation de Windows, mais il peut, bien entendu, modifier n'importe laquelle de celles-ci lors de l'utilisation du système.

Paramètres de Windows

Les premières options configurables des autorisations de Windows concerne la personnalisation du contenu proposé à l'utilisateur (cf. Figure 1).

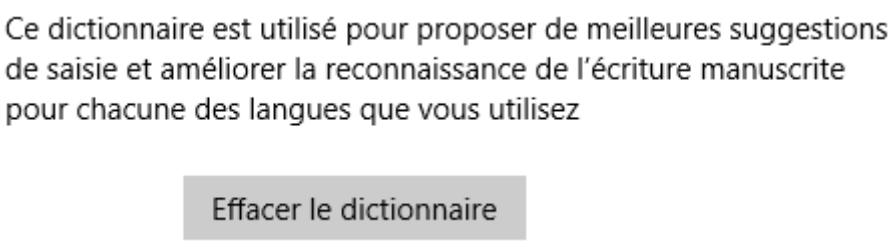
Figure 1: Autorisations générales de Windows



Ces options représentent les premières autorisations que nous pouvons donner à Windows pour la récolte de nos données. Microsoft peut donc récolter des informations concernant les diverses applications utilisées sur la machine ainsi qu'assigner un identifiant unique de publicité afin de proposer un contenu plus « personnalisé » pour l'utilisateur.

La deuxième partie des autorisations demandées par Windows 10 concerne la reconnaissance vocale et manuscrite récoltée par Cortana, l'assistant personnel intelligent de Microsoft (cf. Figure 2). Si l'utilisateur choisit de désactiver cette option, il ne peut pas communiquer avec Cortana et n'a pas accès aux suggestions de saisie de Microsoft. Toutefois, si l'utilisateur choisit d'activer cette option, il peut à tout moment demander d'effacer le dictionnaire de saisie comme nous pouvons le voir sur la figure ci-dessous.

Figure 2: Effacement du dictionnaire de l'utilisateur



Le troisième onglet des paramètres de confidentialité de Windows 10 constitue l'un des plus importants lorsque l'on veut configurer ses options de confidentialité. En effet, il traite des données de diagnostics que le système d'exploitation va envoyer à Microsoft. Ces informations sont d'une grande importance car elles donnent la possibilité à Windows de résoudre d'éventuels problèmes et de rester à jour. Il existe 2 options de configuration prédéfinies que l'utilisateur peut choisir :

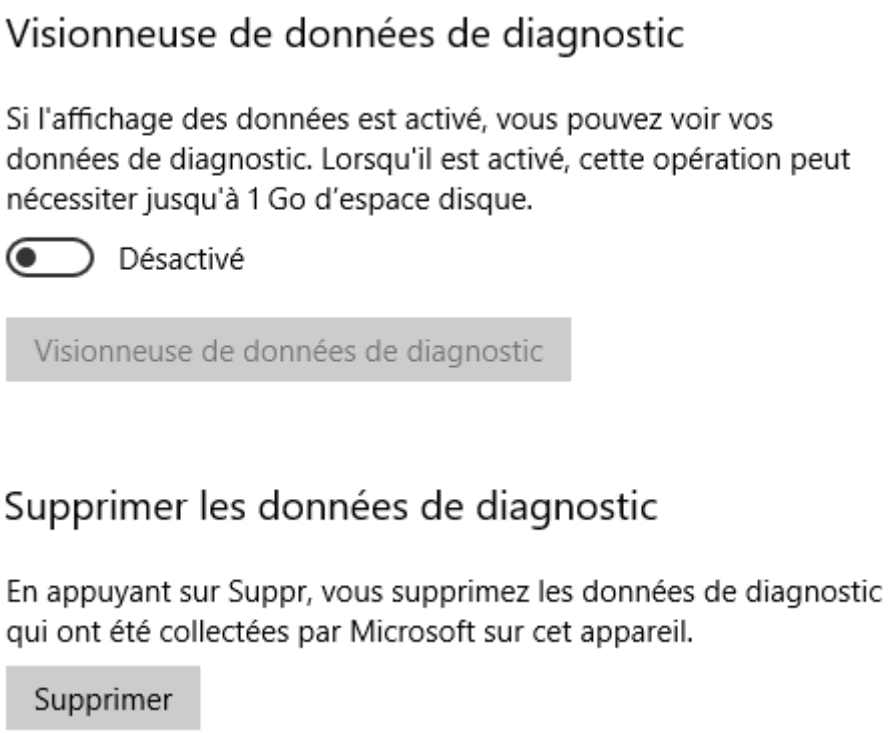
- La configuration de base : Seules des informations concernant l'appareil, ses paramètres, ses fonctionnalités et ses performances sont récoltées ;
- La configuration complète : En plus d'envoyer les données collectées de la configuration de base, Windows va aussi récolter des données sur les divers sites Web qui ont pu être consultés et les applications utilisées.

Encore une fois, le système d'exploitation laisse le choix à l'utilisateur concernant la configuration qu'il souhaite adopter pour la protection de ses informations. Les options précisent notamment que l'appareil bénéficie de la même sécurité avec la configuration de base

ou la configuration complète. A ces paramètres viennent s'ajouter la possibilité d'envoyer davantage d'informations pour personnaliser toujours plus l'expérience.

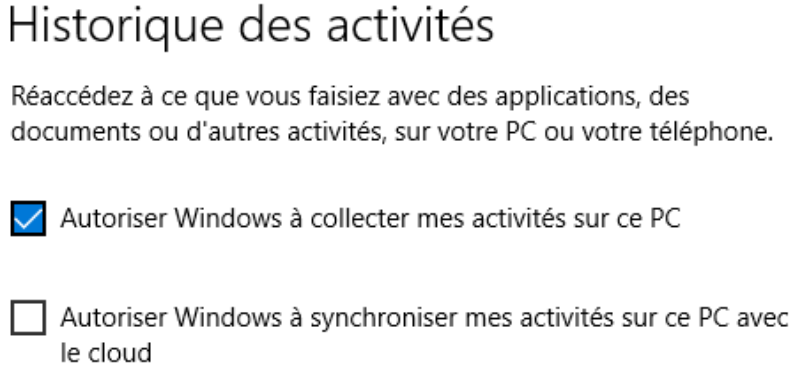
Encore une fois, il est possible, pour l'utilisateur, d'obtenir un droit de regard et de suppression sur les diagnostics établis par Windows concernant sa machine (cf. Figure 3).

Figure 3: Vision et suppression des données de diagnostic Windows



Le dernier paramètre des options de personnalisation de la confidentialité de Windows concerne l'historique des activités enregistrées sur l'ordinateur de l'utilisateur et sur le cloud. L'utilisateur a le choix de donner l'autorisation à Windows de récolter des informations à propos des différentes activités de sa machine et d'ensuite synchroniser ces données avec le cloud (cf. Figure 4). Il est utile de noter que l'historique des activités d'une machine est lié au compte que l'utilisateur a choisi lors de l'installation de son système d'exploitation. Nous

Figure 4: Historique des activités Windows



verrons dans la partie suivante, comment Microsoft gère les données des comptes d'utilisateurs et quelles sont les différentes actions proposées par la firme.

Paramètres des applications

Lorsque l'utilisateur a fini de configurer les options de confidentialité ayant trait à Windows, il peut ensuite configurer quelles informations vont être disponibles aux applications installées sur son ordinateur à travers 20 onglets séparant le type de données accessibles :

- Emplacement
- Caméra
- Microphone
- Notifications
- Informations sur le compte
- Contacts
- Calendrier
- Historique des appels
- Courrier électronique
- Tâches
- Messages
- Radios
- Autres appareils
- Applications en arrière-plan
- Diagnostics de l'application
- Téléchargements automatiques de fichiers
- Documents
- Images
- Vidéos
- Système de fichiers

Pour chacune de ces options, il est possible de choisir quelles applications obtiennent l'accès aux informations disponibles sur la machine. Il est aussi possible de désactiver leur accès à certaines données de manière totale.

MacOS

En ce qui concerne le système d'exploitation proposé sur les machines d'Apple, les paramètres de confidentialité sont présentés de manière plus simplifiée. En effet, Apple veut offrir à ses utilisateurs une expérience personnalisée selon tout un écosystème de produit. La firme propose donc un guide permettant à l'utilisateur de protéger ses données et de configurer directement tous les appareils qu'il possède[2].

Les options de confidentialité mises à disposition de l'utilisateur de MacOS sont beaucoup plus concises que celles présentes dans Windows. Elles se divisent en 4 parties distinctes permettant chacune de configurer un aspect du système. Mais nous pouvons les diviser selon 2 axes principaux : la sécurité et la confidentialité.

[2] Apple, (2019). Gérer votre confidentialité. En ligne https://www.apple.com/befr/privacy/manage-your-privacy/, consulté le 5 avril 2019

Sécurité

La plupart des options proposées parmi les 3 premiers onglets des paramètres de sécurité et confidentialité concerne les fichiers présents sur l'ordinateur et leur sécurité. Les paramètres suivants sont proposés à l'utilisateur :

- Général : Configuration des mots de passe et autorisation de téléchargement d'applications hors AppStore ;
- FileVault : Logiciel permettant le chiffrage automatique des données du disque dur de l'ordinateur ;
- Coupe-Feu : Autorisation/Blocage de connexions entrantes pour les appareils et applications.

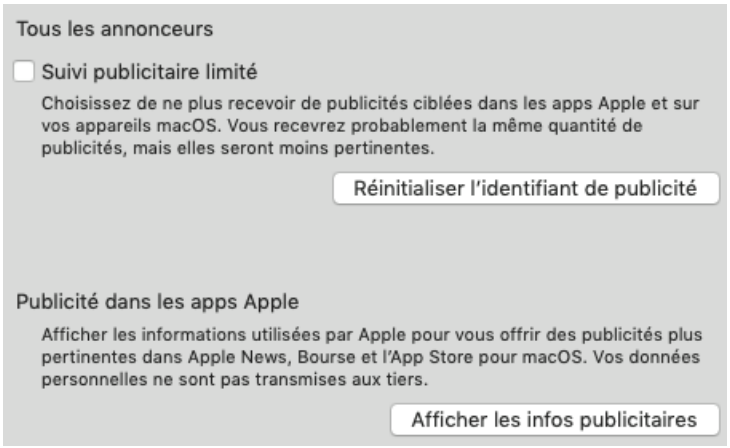
Confidentialité

Apple propose moins d'options de configuration que son concurrent car la gestion des données de ses clients passe par l'AppleID qui est nécessaire pour utiliser n'importe quel produit de la marque. De plus, de nombreuses applications proposées par la firme sont utilisables sur Mac et sur iPhone, ce qui demande une configuration des applications de manière générale. Pour toutes les options suivantes, l'utilisateur a le choix de donner ou non l'accès à ces données aux applications de son choix :

- Service de localisation
- Contacts
- Calendriers
- Rappels
- Photos
- Appareil Photo
- Microphone
- Accessibilité
- Accès complet au disque
- Automatisation

MacOS comporte encore 2 autres paramètres se démarquant du reste car ils concernent les données que l'utilisateur peut accepter d'envoyer à Apple. Premièrement, il est possible de choisir les données d'analyse envoyées à la firme. L'utilisateur peut configurer le partage de 3 types de données différents : les analyses de performance de sa machine, celles de ses applications, et de son cloud. Dans les 3 cas, il a la liberté d'envoyer ou non ces données à Apple (cf. Figure 5).

Figure 5: Options de publicité MacOS



Le deuxième paramètre donne la possibilité à l'utilisateur de configurer le type de publicités qu'il va recevoir lors de l'utilisation des différentes applications proposées par Apple (cf. Figure 6). La même quantité de publicités est proposée ; que l'utilisateur ait choisi un suivi publicitaire limité ou non. Tout comme Windows, un identifiant publicitaire est associé à chaque détenteur d'un compte Apple. Il est tout de même possible d'afficher quelles informations sont utilisées pour proposer les publicités obtenues lors de la navigation.

Figure 6: Options d'analyse MacOS

Linux

La suite Ubuntu de Linux distribuée par l'entreprise Canonical se démarque directement des deux autres systèmes d'exploitation analysés dans ce travail par son caractère libre et Open Source. Le choix d'Ubuntu plutôt qu'une autre suite de Linux a été relativement simple car c'est la suite la plus célèbre et la plus facile d'accès pour les utilisateurs.

Linux est, de manière générale, reconnu comme étant le système d'exploitation le plus « Privacy Friendly » puisque son code n'appartient pas à une entreprise.

La première option de confidentialité est proposée à l'utilisateur lors de l'installation. On donne le choix d'envoyer ou non des informations concernant le système à Canonical (cf. Figure 7) tout en montrant une copie d'un rapport typique.

Figure 7: Demande d'envoi d'informations système Ubuntu

Pareillement à Microsoft et Apple, Canonical explique que la récolte de ces données permet d'améliorer l'expérience des utilisateurs et précise même ne pas retenir l'adresse IP. Il faut savoir que la majorité des personnes utilisant Ubuntu tiennent énormément à leurs informations personnelles. Lorsque Canonical a implanté la fonction de rapport, celle-ci a provoqué de nombreuses discussions au sein de la communauté, raison pour laquelle cette option est désactivée par défaut lors de l'upgrade ou de l'installation de la version 18.04 d'Ubuntu.

En dehors de ces informations techniques, Ubuntu ne collecte pas d'autres données concernant ses utilisateurs, mais il existe tout de même d'autres options de confidentialité personnalisables :

- Verrouillage de l'écran : Un utilisateur peut paramétrer le verrouillage automatique de son écran lorsqu'il s'éloigne de son ordinateur pour éviter l'accès à ses fichiers ;
- Services de localisation : Il est possible de désactiver entièrement la localisation de la machine via les différents points d'accès ;
- Utilisation et historique : Cette option permet de configurer le temps de stockage de l'historique de consultation des fichiers ;

- Vider la corbeille et les fichiers temporaires : Ubuntu laisse la possibilité à l'utilisateur de configurer de manière automatique le nettoyage de la corbeille et la suppression des fichiers temporaires ;
- Signalement de problème : Chaque personne utilisant Ubuntu peut choisir de programmer l'envoi de rapport d'erreurs de manière automatique ou manuelle à Canonical ;
- Vérification de la connectivité : Ce dernier paramètre permet de configurer son réseau et d'utiliser ou non un VPN.

L'utilisateur peut choisir de synchroniser certains de ses comptes à Ubuntu, mais le système d'exploitation n'utilise pas de compte de la même façon qu'Apple ou Microsoft le ferait.

1.3 **Gestion des données**

A la suite du paramétrage des systèmes d'exploitation étudiés, il est important que l'utilisateur sache comment les informations qu'il communique à sa machine, sont gérées par les entreprises possédant les licences de ces OS. Cette partie est divisée selon 3 axes. Nous traitons en premier lieu de l'utilisation qui est faite des données récoltées. Pour qu'un utilisateur sache s'il veut partager ses informations ou non, il est nécessaire qu'il comprenne pourquoi celles-ci sont récoltées et l'utilisation qui en est faite. Le deuxième axe important concerne la manière dont les données sont protégées et stockées. L'utilisateur doit savoir ce qui est mis en place pour assurer la protection de ses informations et où elles sont stockées. Après avoir pris connaissance de tout cela, l'utilisateur doit savoir quelles actions sont mises à sa disposition concernant le contrôle de ses données. Peut-il les modifier ou en demander la suppression ? A-t-il un droit de regard ?

Windows

Après avoir effectué la configuration des différents paramètres concernant la récolte de données, l'utilisateur peut aussi vérifier le traitement et l'utilisation que Microsoft fait des différentes informations qui lui sont envoyés via Windows.

Utilisation des données

Dans sa déclaration de confidentialité[3], Microsoft donne des précisions sur la manière dont les données récoltées par Windows sont utilisées par l'entreprise. Il est important de noter que Microsoft stipule que pour bénéficier de certaines fonctionnalités, le système d'exploitation a besoin de certains types de données, comme par exemple la saisie vocale pour Cortana. Voici comment sont utilisées les données des utilisateurs :

- La fourniture de produit ;
- L'amélioration et le développement de produits ;
- La personnalisation des produits et la recommandation ;
- La publicité et la communication marketing.

Ces 4 utilisations représentent les différents services que Microsoft propose aux clients à partir des informations que ceux-ci leur donnent. La firme précise également qu'elle effectue des analyses de performances et des recherches grâce aux données puisées chez les utilisateurs.

Stockage et protection des données

Il existe plusieurs stratégies mises en place pour protéger les données de clients de Microsoft. Il faut tout d'abord savoir que, n'importe quel client peut savoir à quel endroit sont stockées ses informations à tout moment. En effet, Microsoft possède plusieurs serveurs à travers le monde pour certifier à ses utilisateurs un stockage local de leurs informations.

La firme s'assure aussi qu'une copie de toutes les données d'un client existe pour garantir la récupération de celles-ci lors d'un problème ou d'une panne. De plus, toutes les données récoltées, lorsqu'un client n'est pas connecté à son compte utilisateur, sont stockées séparément du reste des informations personnelles liées au compte. Afin de préserver l'intégrité des informations de chacun, Microsoft utilise l'isolation logique pour empêcher que les données d'un utilisateur se mélangent avec celles d'un autre. Effectivement, les données clients peuvent être stockées sur le même hardware, raison pour laquelle des logiciels permettant de séparer les fichiers sont utilisés afin d'éviter ce genre de problèmes.

[3] Microsoft, (2019) *Déclaration de confidentialité Microsoft*. En ligne https://privacy.microsoft.com/fr-FR/privacystatement#mainwindowsmodule, consulté le 5 avril 2019

Lorsque les données d'un utilisateur sont stockées sur le Cloud, Microsoft va utiliser un système de cryptage dénommé *Azure Information Protection* qui va chiffrer toutes les données avec une clé unique. Le problème de cette approche réside dans le fait qu'il est possible d'obtenir l'accès à ses données via certaines autorisations, ce qui provoque le décryptage des données par Microsoft. Même si la compagnie précise que très peu d'employés ont cet accès.

Actions possibles

Une personne utilisant Windows se voit proposer toute une série d'actions concernant la gestion de ses données personnelles. L'utilisateur a un droit de regard sur chacune des informations récoltées par Microsoft et associées à son compte. En plus de pouvoir obtenir une liste complète de toutes les entreprises tierces qui peuvent obtenir un accès à ces données, un utilisateur peut :

- Consulter chaque information récoltée sur lui par diverses applications ;
- Télécharger une copie de ses données ;
- Modifier ses informations personnelles ;
- Effacer les données associées à n'importe quelle application.

De plus, si un client décide d'arrêter d'utiliser n'importe quel produit Microsoft, dont Windows, la firme lui donne la possibilité de récupérer ses données dans une durée de 90 jours après la fermeture de son compte. S'il décide de désactiver celui-ci, Microsoft procède à la destruction complète de toutes les informations possédées sur le détenteur du compte, copies comprises.

MacOS

Tout comme Microsoft, Apple donne la possibilité à ses utilisateurs d'obtenir des réponses quant à la manière dont l'entreprise gère leurs données. Les différences majeures observées résident surtout dans la manière dont Apple gère la protection des informations personnelles et les actions mises en place pour le client.

Utilisation des données

L'utilisation des données des utilisateurs faites par Apple, possède certaines similarités à celles faites par Microsoft. Mais il existe tout de même certaines différences dans la manière dont Apple utilise les données de ses clients et dans la formulation de leur déclaration de confidentialité. Voici ce que la firme fait avec les informations de ses utilisateurs :

- Diffusion d'événements et mises à jour ;
- Amélioration des produits et prévention contre la fraude ;
- Vérification de l'identité via les données personnelles ;
- Envoi d'informations importantes comme la modification de conditions d'utilisation ;
- Gestion de concours ou d'évènements ;
- Evaluer une éventuelle candidature chez Apple.

Nous pouvons constater qu'Apple possède une approche présentant certaines différences comparées à celle de Microsoft même si la finalité que l'entreprise semble défendre reste la personnalisation de l'expérience de l'utilisateur de ses produits.

Stockage et protection des données

La manière de faire d'Apple concernant le stockage et la protection des données de ses utilisateurs est différente de celle de Microsoft dans plusieurs domaines. En premier lieu, Apple ne met pas à la disposition de ses clients le moyen de savoir dans quel serveur leur données sont stockées. Néanmoins, la firme nous explique de manière plus détaillée que Microsoft les moyens mis en place pour empêcher l'identification des utilisateurs de leurs appareils.

De plus, l'utilisation de MacOS se fait sur des machines Apple, ce qui permet le chiffrage systématique de toutes les données présentes sur l'appareil de l'utilisateur grâce au système de protection FileVault. Les utilisateurs bénéficient donc d'une protection supplémentaire des fichiers stockés sur leurs appareils. Tout comme pour Windows, les données de MacOS sont chiffrées lorsqu'elles sont envoyées sur le Cloud ou récoltées par Apple. Le géant à la pomme permet aussi aux développeurs d'avoir accès à leur technologie de protection des données pour certifier que les informations des clients soient aussi protégées lors de l'utilisation d'applications.

La plus grande différence réside dans la manière dont Apple récolte les données de ses clients pour permettre à ceux-ci de rester anonyme. En effet, l'entreprise utilise la confidentialité différentielle. Cette technologie fonctionne de la manière suivante : toute une série d'informations aléatoires sont ajoutées aux données avant de pouvoir être analysées par Apple. Cela donne la possibilité à la compagnie de dégager des tendances globales à grande échelle tout en garantissant le caractère privé des données de ses utilisateurs.

Actions possibles

Il est évident qu'Apple possède une manière de gérer les données de ses utilisateurs se démarquant de ses concurrents. Cela vient notamment de l'offre que l'entreprise veut proposer à ses clients. Effectivement, comme nous l'avons mentionné plus tôt, le but d'Apple est d'offrir une expérience multi-plateforme personnalisée qui permet la synchronisation immédiate des différents produits en possession de l'utilisateur. L'AppleID étant unique, n'importe qui pourra, par exemple, consulter ses iMessages sur son téléphone ou son ordinateur.

Dans cet esprit, Apple permet à ses clients de gérer les données associées à leur AppleID de différentes manières. Il existe 4 actions mises en avant lorsqu'un utilisateur consulte la partie « Gestion de vos données » du site d'Apple[4] :

- Obtention d'une copie des données du compte : N'importe quel utilisateur a le droit de demander et de télécharger une copie complète de toutes les informations qu'Apple a récoltées et associées à son AppleID ;
- Correction des données du compte : Après avoir consulté ses données, un utilisateur peut corriger n'importe lesquelles de celles-ci s'il juge qu'elles sont incorrectes ;
- Désactivation temporaire du compte : S'il le souhaite, un utilisateur peut demander à Apple de suspendre l'activité de son compte, ce qui limite l'accès aux données. L'utilisateur ne peut pas non plus accéder au compte et bénéficier des différents services proposés par Apple ;
- Suppression du compte : Un utilisateur peut, à tout moment, demander la suppression totale de son compte et de toutes les données qui y sont associées dans tous les services Apple.

[4] Apple, (2019). *Data and Privacy*. En ligne https://privacy.apple.com, consulté le 5 avril 2019

Il faut remarquer l'absence d'une option qui était pourtant présente chez Microsoft. L'utilisateur ne peut pas demander la suppression de ses données sans pour autant garder l'AppleID qui lui est associé. En cas d'une demande de désactivation temporaire de celui-ci, le client perd l'accès à ses données de compte et aux services qui lui étaient proposés auparavant.

Linux

Comme précisé lors de l'analyse des différents paramètres de confidentialité d'Ubuntu, le système ne récolte pas, ou très peu de données concernant les utilisateurs. Les seules informations pouvant éventuellement être récoltées par Canonical concerne les rapports de diagnostics qui peuvent être désactivés.

Néanmoins, par soucis de transparence, Canonical rend publique et disponible toutes les informations obtenues lors de l'obtention de leurs rapports. Les données collectées concernent surtout des sujets tels que le temps d'installation moyen d'Ubuntu sur une machine ou la quantité de RAM possédée par un ordinateur. Ces rapports rendent possible l'amélioration du système et permettent à Canonical de se concentrer sur les logiciels les plus utilisés par sa communauté. L'entreprise précise[5] également que les données géographiques proviennent du fuseau horaire sélectionné lors de l'installation d'un produit et non via l'adresse IP qui n'est pas enregistrée par la firme.

1.4 **Comparaison des systèmes d'exploitation**

Maintenant que nous avons analysé chacun des paramètres de confidentialités et la manière dont les systèmes d'exploitation gèrent les données des utilisateurs, nous pouvons créer plusieurs tableaux nous permettant de comprendre les majeures similitudes et/ou différences existant entre eux. Cela nous permettra d'établir une comparaison concernant leur manière d'aborder la confidentialité et de respecter les données de leurs utilisateurs. Dans un premier temps, établissons un tableau servant de carte d'identité de ces OS.

[5] Cooke, W. *A first look at desktop metrics*. En ligne https://blog.ubuntu.com/2018/06/22/a-first-look-at-desktop-metrics, consulté le 5 avril 2019

Tableau 1: Carte d'identité des systèmes d'exploitation

	Windows	***Mac OS***	***Linux***
Famille	Windows NT	Unix	Unix
Version	Windows 10 (10.0.17763.404)	Mojave (10.14.4)	Ubuntu (18.04 LTS)
Type de licence	Propriétaire	Propriétaire	Libre et Open Source
Développeurs	Microsoft	Apple	Canonical

La première différence notoire pouvant avoir un impact sur l'approche de la confidentialité de ces différents systèmes d'exploitation est le type de licence de chacun d'entre eux (cf. Tableau 1). En effet, Microsoft et Apple possèdent tous deux, tous les droits et toutes les versions de leur OS respectif et ceux-ci font partie intégrante de leur gamme de produits.

Maintenant que nous en savons un peu plus sur les systèmes d'exploitation que nous avons analysés, nous pouvons présenter les différents critères de notre tableau comparatif avant d'interpréter les résultats que nous avons obtenus. Les différentes catégories présentes dans notre grille traitent chacune d'un aspect différent de la gestion des informations de l'utilisateur.

Paramètres de confidentialité

Le tout premier critère d'importance de notre analyse concerne la manière dont l'utilisateur peut configurer les paramètres de confidentialité de son système d'exploitation et la possibilité qu'il a de choisir quelles données sont récoltées lors de l'utilisation. Est-il possible de reconfigurer ses choix de confidentialité à tout moment ? Quand est-il de la protection des données de l'appareil ?

Nécessité d'un compte client pour les fonctionnalités

Ce paramètre précise si le système d'exploitation utilisé nécessite la possession d'un compte client pour bénéficier de toutes les fonctionnalités proposées.

Récolte de données

Le critère suivant reprend le type d'information qui est récolté lors de l'utilisation du système d'exploitation. L'OS récolte-t-il seulement des données techniques et des rapports de performance ? Ou les données des utilisateurs sont-elles aussi collectées ?

Utilisation des données

Dans cet onglet, nous apportons une précision quant à la suite réservée aux données récoltées par l'entreprise possédant le système d'exploitation. Les données sont-elles simplement utilisées pour améliorer l'expérience des utilisateurs ? La firme responsable de l'OS utilise-t-elle les informations de ses clients pour vendre d'autres de ses produits ?

Stockage et protection des données

Dans cette partie de la grille, nous trouvons des informations concernant ce que l'utilisateur sait sur le stockage de ses données et sur la manière dont elles sont protégées. Nous obtenons des informations sur les procédés utilisés pour assurer l'intégrité et la sécurité des données. L'utilisateur a-t-il le droit de savoir où sont stockées ces données ? Comment la firme responsable du système d'exploitation protège-t-elle les données de ses utilisateurs ?

Actions proposées

Les 3 derniers critères repris dans notre grille comparative, concernent les actions possibles pour les utilisateurs concernant les données récoltées par le système d'exploitation qu'ils ont choisi. Ils se divisent de cette manière :

- Consultation des données : Les utilisateurs possèdent-ils bien un droit de regard sur leurs informations ? Ont-ils d'autres actions disponibles ?
- Modification des données : Quel est le champ d'action de l'utilisateur lorsqu'une information qui est collectée sur lui est erronée ?
- Suppression des données : Quelles actions peuvent être entreprises par un utilisateur qui souhaiterait que les données récoltées par son système d'exploitation soient détruites ?

Tableau 2: Paramètres de confidentialité et Gestion des données des systèmes d'exploitation

	Windows	***MacOS***	***Linux***
Paramètres de confidentialité	• Configurables à l'installation et pendant l'utilisation • Personnalisation du contenu	• Configurables à l'installation et pendant l'utilisation • Personnalisation du contenu • Possibilité de chiffrage des fichiers	• Configurables à l'installation et pendant l'utilisation • Possibilité de ne transmettre aucune données
Nécessité d'un compte client pour les fonctionnalités	Oui	Oui	Non
Récolte de données	• Performances • Utilisateurs	• Performances • Utilisateurs	• Performances
Utilisation des données	• Personnalisation et améliorations des produits • Publicités ciblées	• Personnalisation et améliorations des produits • Publicités ciblées • Vérification d'identité	• Obtenir des données sur les applications les plus utilisées
Stockage et protection des données	• Possibilité de localisation des données • Isolation logique • Chiffrage des données clients sur les serveurs (Azure Information Protection) • Copie des données en cas de problème	• Chiffrage des données clients sur les serveurs • Chiffrage possible sur les machines (FileVault) • Récolte de données permettant la préservation de la vie privée (Confidentialité différentielle)	• Données récoltées anonymes
Consultation des données	• Possibilité d'obtenir une copie des données	• Possibilité d'obtenir une copie des données	-
Modification des données	• Possibilité de correction	• Possibilité de correction	-
Suppression des données	• Possible à n'importe quel moment • 90 jours après suppression du compte	• Possibilité de bloquer les données à la désactivation du compte • Seulement à la suppression du compte Apple	-

1.5 **Quel système d'exploitation choisir ?**

Après avoir présenté tous les paramètres offerts par les OS que nous avons analysés et comparés via notre grille, nous pouvons nous poser la question du choix. Un système d'exploitation est-il meilleur qu'un autre ? L'utilisateur peut se demander pourquoi il devrait choisir une option plutôt qu'une autre. Windows étant le système d'exploitation le plus utilisé sur le marché, il est important de savoir s'il existe une réelle différence et un réel danger pour l'individu qui ne souhaite pas partager ses données.

Il existe plusieurs réponses à cette question (cf. Tableau 2), la première étant qu'il est important de savoir ce que valorise l'utilisateur en premier. Windows et MacOS sont des options viables, car les entreprises se trouvant derrière ces deux systèmes sont tenues de respecter les données de leurs clients et de leur offrir certaines garanties depuis l'instigation de la RGPD en Europe. Un utilisateur préférant avoir accès à un contenu plus personnalisé et n'ayant pas forcément le temps de configurer lui-même son système va pouvoir opter pour Windows 10 ou MacOS.

D'autre part, un individu ne voulant absolument pas que ses informations personnelles soient récoltées et utilisées par l'entreprise éditant son OS, va plutôt devoir se diriger vers Ubuntu qui reste la meilleure option au niveau de la confidentialité, bien que Microsoft et Apple proposent des services de chiffrage de données pouvant rassurer l'utilisateur. Microsoft possède tout de même un avantage par rapport à Apple concernant les différentes actions proposées aux clients. En effet, l'utilisateur peut supprimer ses données personnelles récoltées par Windows à tout moment tandis qu'un client Apple va devoir supprimer entièrement son compte.

En conclusion, Linux est une excellente option pour les utilisateurs un peu plus aguerris qui se soucient de ce que les grandes entités peuvent faire de leurs informations personnelles. Quant à Windows et MacOS, ils possèdent tout 2 leurs avantages et leurs inconvénients. MacOS prend les devants concernant la sécurité des données de ses utilisateurs en proposant un logiciel de chiffrage intégré à l'OS et en utilisant la confidentialité différentielle tandis que Windows permet l'accès et la suppression de données à tout moment.

2. Navigateurs

2.1 Choix des navigateurs :

Pour pouvoir accéder à n'importe quel site internet, un individu va devoir utiliser un navigateur. Cet outil s'occupe d'envoyer une requête au serveur pour récupérer les différents composants de la page Web à laquelle l'utilisateur veut se connecter. De nos jours, il existe toute une série de navigateurs différents qu'il est possible d'installer sur son ordinateur. Chacun possède ses propres particularités et des options qui lui sont propres. Certains d'entre eux sont plus populaires que les autres et possèdent une grande partie du marché. Mais pour quelles raisons doit-on en privilégier un plutôt qu'un autre ?

La deuxième partie d'analyse de ce travail se concentre sur ces outils utilisés chaque jour par des milliards d'internautes à travers le monde. Après s'être concentrés sur les données récoltées par les systèmes d'exploitation, nous allons analyser comment un utilisateur peut configurer son navigateur de manière à diminuer le nombre de données qui est récoltées par toute une panoplie de traqueurs présents sur la toile. Est-il possible de limiter les informations envoyées par notre navigateur ? Quelles options de configuration sont mises à la disposition de l'internaute et sont-elles suffisantes ? Nous tentons de donner des éléments de réponse à ces questions.

Cette partie sera divisée de manière à proposer une comparaison des navigateurs sur les paramètres de confidentialités qu'ils proposent aux utilisateurs et sur la protection qu'il est possible d'avoir en configurant correctement ces options. Nous établissons plusieurs tableaux comparatifs permettant d'obtenir une idée sur la performance de chacun des outils utilisés lors de la navigation.

Malgré des dizaines d'options différentes, nous avons choisi de nous limiter aux navigateurs les plus populaires lorsque nous naviguons sur le Web. En effet, plusieurs d'entre eux se démarquent du reste, que cela soit via le nombre d'internautes qui les utilisent ou parce qu'ils sont proposés par défaut sur certains systèmes d'exploitation. Nous allons brièvement les présenter ainsi que les raisons pour lesquelles ceux-ci ont été choisis.

Google Chrome

Le premier navigateur choisi dans le cadre de ce travail est celui le plus utilisé sur le marché. Avec 65,64 % des utilisateurs en Avril 2019 selon NetMarketShare (2019), le navigateur de la firme Google se place en tête dans le classement. Il est disponible sur tous les systèmes d'exploitation que nous avons analysé plus tôt depuis septembre 2008 et appartient à la catégorie des « freeware ».

Selon la Fondation pour le logiciel libre (2019) un freeware est : « un logiciel propriétaire distribué gratuitement sans toutefois conférer à l'utilisateur certaines libertés d'usage associées au logiciel libre ». Le navigateur et son code appartiennent donc entièrement à Google, même si celui-ci les met à la disposition des internautes de manière gratuite. Son avantage est qu'il est directement relié au moteur de recherche de Google ce qui peut faciliter la vie des utilisateurs. De nombreuses extensions sont aussi disponibles pour que les internautes puissent personnaliser leur expérience de navigation.

Mozilla Firefox

Le deuxième navigateur le plus utilisé sur le marché avec 10,23 % selon NetMarketShare (2019) se trouve être Mozilla Firefox. Il se démarque des autres navigateurs car tout comme Linux pour les systèmes d'exploitation, Firefox est entièrement libre et Open Source. Son code source est mis à disposition par Mozilla et enregistré sous la licence MPL-2.0 (Mozilla Public License). N'importe quelle personne peut donc y avoir accès.

Tout comme Google Chrome, le navigateur peut être installé de manière entièrement gratuite sur les 3 systèmes d'exploitation que nous avons analysés. C'est d'ailleurs celui qui est proposé par défaut sur la plupart des variantes de Linux, vu son côté libre. A l'effigie de ses concurrents, Firefox propose aussi toute une gamme d'applications permettant de personnaliser la navigation de l'utilisateur. Son principal atout est que n'importe quel internaute peut choisir d'installer une version du navigateur incorporant toute une série d'outils de développement permettant de coder pour encore davantage personnaliser l'expérience proposée. La politique de confidentialité[6] de Firefox est aussi présentée lors de l'installation du navigateur.

[6] Cf. Figure 22 Annexe 2

Microsoft Edge

Microsoft Edge est le troisième navigateur choisi dans le cadre des analyses effectuées dans ce travail. Bien qu'utilisé par seulement 5,53% des internautes, il constitue un choix important car c'est à la fois le successeur d'Internet Explorer et le navigateur par défaut proposé sur Windows 10. Avec une première version datant de juillet 2015, Edge est le navigateur le plus récent parmi ceux que nous allons étudier. Il n'est malheureusement pas disponible sur Linux et MacOS, étant un logiciel propriétaire appartenant à Microsoft et fourni avec Windows.

De même que les autres navigateurs étudiés, Microsoft Edge propose aussi d'installer toute une série d'extensions permettant de créer son expérience de navigation personnalisée. Il donne aussi accès à l'assistant intelligent de Windows 10, Cortana, qui est intégré au navigateur. Edge est relié directement au moteur de recherche de Microsoft, Bing.

Safari

Le dernier navigateur sur lequel nous allons travailler n'est autre que Safari, le navigateur propriétaire distribué par défaut sur tous les produits Apple dont le système d'exploitation MacOS. Il est apparu en Janvier 2003, à la suite de la fin du support de Microsoft concernant Internet Explorer sur Mac. Depuis, environ 3,58% des internautes ont fait le choix de l'utiliser.

Tout comme les autres navigateurs choisis pour ce mémoire, Safari possède une galerie d'extensions qui permet à ses utilisateurs de personnaliser leur expérience. Le navigateur n'est disponible que sur les produits Apple. Effectivement, il existait une version pour Windows mais Apple a décidé d'abandonner son support après 2012.

2.2 Critères de comparaison

Avant d'effectuer des tests pour comparer à quel point les différents navigateurs que nous avons choisis protègent nos données, il est important de définir sur quels critères nous allons comparer les paramètres de confidentialité proposés par ceux-ci. L'analyse se déroule en 2 parties. En premier lieu, nous comparons les options de confidentialité sélectionnées par défaut lors de l'installation d'un navigateur. Lorsque nous aurons obtenu un tableau récapitulatif, nous

pourrons effectuer les tests pour vérifier si chacune des configurations par défaut permet de protéger les informations des utilisateurs.

En effet, il n'est pas rare d'entendre parler du « Privacy Paradox » ces dernières années. Il semble que de plus en plus d'utilisateurs disent se soucier de leurs données, mais ne mettent rien en place pour protéger celles-ci. Il est donc légitime d'analyser les paramètres proposés par un navigateur par défaut avant de configurer celui-ci de la manière la plus optimale. Il est évident que tous ne vont pas proposer les mêmes options de sécurité concernant les informations récoltées en ligne. C'est pour cela que nous établissons une grille d'options que nous allons détailler.

Blocage de contenu

La toute première option que nous allons comparer dans le tableau concerne le paramétrage du contenu proposé lors de la navigation. En effet, il n'est pas rare que des navigateurs proposent un mode de « navigation privée » qui permet de ne pas enregistrer l'historique et de ne pas activer les cookies lors de la navigation. Notre but est de vérifier si les navigateurs que nous avons choisis se limitent à un blocage dit « Standard » qui ne fonctionne que lors de la navigation privée, ou permettent un blocage « Personnalisé » avec différentes options.

Blocage d'applications tierces

Le deuxième paramètre que nous avons choisi est relativement proche du tout premier. Il concerne le blocage des « Cookies » et applications dites tierces. Lorsqu'un utilisateur navigue sur le Web, il est amené à rencontrer toute une série de publicités ciblées. Les « Cookies » et traqueurs tiers sont ceux qui sont utilisés par les entreprises proposant ces publicités aux internautes. Ils sont appelés « tiers » car ils n'appartiennent pas au site visité par l'internaute. L'apparition de cette option dans notre grille est primordiale car c'est à travers ces traqueurs que les entreprises vont récolter des informations sur ce qu'un individu consulte sur le net.

Do Not Track

Depuis un certain temps, une nouvelle technologie de protection appelée Do Not Track est apparue dans certains navigateurs (Ruiz-Martinez, 2012). Si un internaute choisit d'activer cette

option lors de la navigation, une requête va être envoyée par le navigateur au serveur pour déclarer que l'utilisateur ne veut pas que ses informations soient récoltées par des annonceurs tiers. Bien que cette solution paraisse attrayante, beaucoup de sites internet ne respectent pas encore son application.

Navigation sécurisée

Il est nécessaire pour un utilisateur de pouvoir naviguer de manière entièrement sécurisée, sans avoir à se soucier de logiciels malveillants ou de téléchargements douteux. Le navigateur représente la première barrière de défense lorsqu'un internaute est sur le Web. Si le contenu dangereux est automatiquement bloqué, un individu ne doit pas faire autant attention aux sites qu'il consulte.

Utilisation de Cookies

Savoir si notre navigateur utilise ou non des Cookies pour stocker des informations lors de notre navigation est extrêmement important. Ils vont permettre à un site Web de garder certaines informations concernant un internaute pour pouvoir lui proposer du contenu personnalisé lors de sa prochaine visite. Certains utilisateurs n'apprécient pas l'utilisation des Cookies car ceux-ci peuvent stocker toute une série d'informations considérées comme privées.

Enregistrement des identifiants

Pour personnaliser et faciliter encore plus l'expérience des internautes, de nombreux navigateurs permettent l'enregistrement systématique des identifiants permettant de se connecter à certains sites Web. L'ajout de cette option dans notre grille va nous permettre de déterminer si l'enregistrement des identifiants se fait de manière systématique sur les navigateurs ou si cela requiert l'autorisation de l'utilisateur.

Règle de conservation de l'historique

L'historique des pages qui ont été visitées par un utilisateur est une information qui contient énormément de valeurs pour de nombreuses entreprises. En effet, si une firme possède une liste complète des sites consultés par un internaute, elle peut facilement établir les différents intérêts

d'une personne et le contenu qu'elle aime voir apparaître lors de sa navigation. Ce paramètre nous permet de savoir si l'historique de navigation est conservé par défaut par les différents navigateurs.

Envoi de données techniques

Les navigateurs envoient des données techniques et des rapports de performance à leurs développeurs de la même manière que les systèmes d'exploitation. Celles-ci peuvent contenir des informations sur les composants de l'ordinateur ou encore certaines données de navigation. Tout comme pour les OS, un utilisateur peut ne pas vouloir partager d'informations concernant sa machine ou ses habitudes de navigation.

JavaScript

JavaScript est un langage de programmation qui est très utilisé sur le Web car il donne la possibilité de rendre les pages internet interactives grâce à l'utilisation de script selon Mozilla[7]. Les sites Web peuvent envoyer toute une série de scripts au navigateur, comme des publicités. Cela pose certains problèmes de sécurité car, selon ADsafe[8], n'importe quel script envoyé peut avoir accès à toutes les informations disponibles sur une page. C'est pour cela qu'il existe certains mécanismes permettant de limiter l'accès de JavaScript ou même de ne pas l'utiliser.

Flash

Certaines pages Web utilisent ce qu'on appelle la technologie Flash, qui peut être lue par Adobe Flash Player[9], pour envoyer du contenu à un navigateur. Mais cela peut également servir à stocker des données concernant un utilisateur sur un site internet utilisant le lecteur d'Adobe (Hassan et Hijazi, 2018). Cependant, tous les navigateurs ne possèdent pas un lecteur de fichier Flash dès l'installation ou ne permettent pas d'en désactiver l'utilisation.

[7] Mozilla, (2019) *JavaScript*. En ligne https://developer.mozilla.org/fr/docs/Web/JavaScript, consulté le 7 avril 2019
[8] ADsafe, (2019) *Making JavaScript Safe for Advertising*. En ligne http://www.adsafe.org/, consulté le 9 avril 2019
[9] Adobe, (2019). *Adobe Flash Player*. En ligne https://get.adobe.com/fr/flashplayer/about/, consulté le 14 mai 2019

Aide à la saisie

La possibilité d'obtenir des suggestions lorsqu'un utilisateur cherche un site internet ou écrit un message est devenue une fonctionnalité de base de nombreux produits. Mais cette aide se base généralement sur les anciennes actions de l'individu pour pouvoir l'assister au mieux lors de sa recherche. Nous essayons de déterminer si elle est présente par défaut dans les navigateurs et s'il est possible de la désactiver.

Blocage de fenêtre pop-up

Lors de la navigation, il n'est pas rare de consulter un site et de voir apparaître une fenêtre de publicité dans un des coins de son écran. Cet ajout à la grille nous permet de savoir si cette fonctionnalité existe sur tous les navigateurs sélectionnés et si elle est activée par défaut pour rendre l'expérience de l'utilisateur plus agréable.

Accès à la localisation, caméra, microphone, notifications

Ce dernier élément regroupe 4 critères de comparaison différents mais nous les avons regroupés car ils ont, généralement, une fonctionnalité commune. L'utilisateur peut choisir d'autoriser l'accès de chacune de ces informations à une liste de site Web ou non. Nous tentons de déterminer si l'accès est autorisé par défaut ou si le navigateur installé effectue une demande à l'utilisateur lorsqu'une page veut avoir accès à ces données.

2.3 Comparaison des navigateurs

Afin de s'assurer que les données correspondent pour chacun des navigateurs et que les paramètres n'aient pas été altérés, nous avons procédé à une installation de base pour Google Chrome et Mozilla Firefox. Pour rester dans la même optique, tous les paramètres de Safari et de Microsoft Edge ont été réinitialisés à ceux fournis de base avec leur système d'exploitation respectif. C'est seulement après s'être assuré de n'avoir que les navigateurs avec leurs options par défaut et sans aucune extension supplémentaire que nous avons commencé à compléter notre grille de comparaison.

Configuration par défaut des navigateurs

La grille que nous allons présenter maintenant contient le réglage par défaut de chacune des options qui animent notre sujet. Cette configuration est bien celle donnée par le navigateur sans aucune modification. Ce qui signifie que ce sera donc celle que l'internaute utilisera comme paramètres s'il se contente d'une installation normale de ces navigateurs sans pour autant consulter les options disponibles (cf. Tableau 3).

Tableau 3: Configuration par défaut des navigateurs

	Google Chrome	***Mozilla Firefox***	***Microsoft Edge***	***Safari***
Blocage de contenu	Standard	Standard	Standard	Standard
Blocage d'applications tierces	Non	Non	Non	Non
Do Not Track	Non	Non	Non	Non
Navigation sécurisée	Oui	Oui	Oui	Oui
Utilisation de Cookies	Oui	Oui	Oui	Oui
Enregistrement des identifiants	Demande d'autorisation	Demande d'autorisation	Oui	Oui
Règle de conservation de l'historique	Conservé	Conservé	Conservé	Conservé
Envoie de rapports de performance et de données techniques	Oui	Oui	Option liée à Windows	Oui
JavaScript	Oui	Oui	Oui	Oui
Flash	Demande d'autorisation	Pas installé de base	Oui	Non
Aide à la saisie	Oui	Oui	Oui	Oui
Blocage Pop-up	Oui	Oui	Oui	Oui
Localisation	Demande d'autorisation	Demande d'autorisation	Option liée à Windows	Demande d'autorisation

Caméra	Demande d'autorisation	Demande d'autorisation	Option liée à Windows	Demande d'autorisation
Microphone	Demande d'autorisation	Demande d'autorisation	Option liée à Windows	Demande d'autorisation
Notifications	Demande d'autorisation	Demande d'autorisation	Option liée à Windows	Demande d'autorisation

Pour chacun des navigateurs que nous avons sélectionnés, la configuration par défaut est relativement similaire. Lorsqu'un utilisateur choisit d'installer un navigateur, il est paramétré pour permettre la récolte et l'utilisation des données de navigations. La seule réelle différence semble résider dans l'enregistrement automatique des identifiants pour les navigateurs étant présents par défaut sur leur système d'exploitation respectif.

Tests des navigateurs

Afin de pouvoir effectuer une comparaison objective des différents navigateurs que nous avons choisis d'analyser dans ce travail, nous allons effectuer le test Panopticlick (cf. Figure 8[10]) proposé par l'Electronic Frontier Foundation. Celui-ci permet de vérifier si notre navigateur nous protège ou non des différents traqueurs présents sur internet. Pour ce faire, il nous propose de tenter de nous connecter à des domaines étant connus pour traquer les utilisateurs.

Figure 8: Test Panopticlick

[10] Electronic Frontier Foundation, (2019). *Is your browser safe against tracking ?*. En ligne https://panopticlick.eff.org, consulté le 20 mars 2019

Ce test va nous permettre de comparer nos navigateurs via 5 questions qui se présentent de la manière suivante :

- Votre navigateur bloque-t-il les traqueurs de publicités ?
- Votre navigateur bloque-t-il les traqueurs invisibles ?
- Votre bloqueur de publicité arrête-t-il les traqueurs placés dans la liste des publicités acceptables ? *Cette option n'apparaît pas toujours si le test considère que notre navigateur n'utilise pas de bloqueur de publicités*
- Est-ce que votre navigateur débloque les parties tierces qui ont promis d'honorer l'accord « Do Not Track » ?
- Est-ce que votre navigateur vous protège du fingerprinting ?

Il existe 3 réponses possibles à chacune de ces questions : Oui, Non, et Protection partielle.

Dans un premier temps, nous allons effectuer ce test avec chacun de nos navigateurs et leurs paramètres par défaut. Lorsque cela sera fait, nous configurerons les navigateurs pour qu'ils bloquent le plus de traqueurs possibles et qu'ils respectent au mieux les données des utilisateurs. Nous proposerons à nouveau le test fourni par l'EFF pour voir si nous observons des changements au niveau des résultats.

Test avec configuration par défaut

Tableau 4: Résultats du test Panopticlick sur configuration par défaut

	Google Chrome	***Firefox***	***Microsoft Edge***	***Safari***
Est-ce que votre navigateur bloque les traqueurs de publicités ?	Non	Non	Non	Non
Est-ce que votre navigateur bloque les traqueurs invisibles ?	Non	Non	Non	Non
Est-ce que votre bloqueur de publicité arrête les traqueurs placés dans la liste des publicités acceptables ?	-	-	-	-
Est-ce que votre navigateur débloque les parties tierces qui ont promis d'honorer l'accord « Do Not Track » ?	Non	Non	Non	Non
Est-ce que votre navigateur vous protège du fingerprinting ?	Non	Non	Non	Non

Source : Electronic Frontier Foundation, (2019). *Is your browser safe against tracking?*. En ligne https://panopticlick.eff.org, consulté le 20 mars 2019

Nous pouvons tout de suite remarquer que les résultats du test sont sans appel. Aucun des navigateurs que nous avons choisis ne protège l'utilisateur avec sa configuration par défaut (cf. Tableau 4). Jusque maintenant, tous proposent des services relativement similaires aux internautes. Aucun d'entre eux ne semble être plus performant concernant la protection des données de navigation.

Test sur navigateurs configurés

Maintenant que nous avons comparé chacun des navigateurs et leur configuration par défaut. Nous allons les configurer pour qu'ils respectent au mieux les données de navigation de l'utilisateur et qu'ils bloquent le plus de logiciels de traque possibles (cf. Tableau 5).

Tableau 5: Configuration la plus stricte possible des navigateurs

	Google Chrome	***Mozilla Firefox***	***Microsoft Edge***	***Safari***
Blocage de contenu	Personnalisé	Personnalisé	Personnalisé	Personnalisé
Blocage d'applications tierces	Oui	Oui	Oui	Oui
Do Not Track	Oui	Oui	Oui	Oui
Navigation sécurisée	Oui	Oui	Oui	Oui
Utilisation de Cookies	Non	Non	Non	Non
Enregistrement des identifiants	Non	Non	Non	Non
Règle de conservation de l'historique	Suppression automatique	Suppression automatique	Suppression automatique	Suppression automatique
Envoie de rapports de performance et de données techniques	Non	Non	Non	Non
JavaScript	Non	Oui	Oui	Non
Flash	Non	Pas installé de base	Non	Non
Aide à la saisie	Non	Non	Non	Non
Blocage Pop-up	Oui	Oui	Oui	Oui
Localisation	Non	Non	Non	Non
Caméra	Non	Non	Non	Non
Microphone	Non	Non	Non	Non
Notifications	Non	Non	Non	Non

Nous constatons que lorsqu'ils sont configurés de manière à augmenter la confidentialité de l'utilisateur, nos navigateurs proposent un paramétrage relativement similaire. Les seules différences résidant dans l'utilisation de Flash ou de JavaScript. En effet, il est impossible de changer les paramétrages concernant JavaScript pour Microsoft Edge et Firefox sans utiliser la console. Ce qui nous intéresse ici étant les options de confidentialité accessibles à n'importe quel utilisateur, nous avons décidé de laisser JavaScript activé. Sa désactivation peut, en effet, poser un problème pour le chargement et le bon fonctionnement de certaines pages car cela permet généralement de charger les scripts pour les menus déroulants, etc (cf. Figure 9[11]).

Figure 9: Page YouTube sans JavaScript

Maintenant que nous avons pris connaissance de la façon dont nos navigateurs sont configurés, nous pouvons recommencer le test proposé par Panopticlick pour voir à quel niveau l'utilisateur est protégé.

[11] Youtube, (2019). Youtube. En ligne https://www.youtube.com, consulté le 15 avril 2019

Tableau 6: Résultats du test Panopticlick sur configuration stricte

	Google Chrome	***Firefox***	***Microsoft Edge***	***Safari***
Est-ce que votre navigateur bloque les traqueurs de publicités ?	Oui	Oui	Protection Partielle	Protection Partielle
Est-ce que votre navigateur bloque les traqueurs invisibles ?	Oui	Oui	Protection Partielle	Protection Partielle
Est-ce que votre bloqueur de publicité arrête les traqueurs placés dans la liste des publicités acceptables ?	-	Oui	Non	Non
Est-ce que votre navigateur débloque les parties tierces qui ont promis d'honorer l'accord « Do Not Track » ?	Non	Non	Non	Non
Est-ce que votre navigateur vous protège du fingerprinting ?	Non	Non	Non	Non

Source : Electronic Frontier Foundation, (2019). *Is your browser safe against tracking ?*. En ligne https://panopticlick.eff.org, consulté le 20 mars 2019

Lorsque les navigateurs sont configurés pour respecter au mieux la vie privée de l'utilisateur et partager le moins de données de navigation possible, les résultats du test s'en voient fortement impactés (cf. Tableau 6). Nous pouvons déjà constater une démarcation entre les options par défaut proposées par Windows 10 et MacOS et les navigateurs disponibles en téléchargement sur le Web, Firefox et Google Chrome. Nous constatons aussi que Firefox bloque même les traqueurs présents sur les pages proposant des publicités acceptables, information qui n'a pas été trouvée par le test pour Google Chrome.

Il y a encore une autre observation qui peut être faite grâce à notre tableau comparatif. Effectivement, aucun des navigateurs configurés ne respecte l'accord « Do Not Track » alors que ceux-ci sont paramétrés en conséquence. Cela n'impacte pas la navigation de l'utilisateur, mais il est important de noter que cette option n'est pas respectée alors qu'elle leur est proposée. En plus de cela, aucun des navigateurs n'a réussi à protéger les utilisateurs du fingerprinting. Ils peuvent donc être identifiés malgré le blocage des différents traqueurs.

2.4 Quel navigateur choisir ?

De la même façon que pour les systèmes d'exploitation, nous pouvons nous poser la question du choix du navigateur. Si un utilisateur veut être certain de pouvoir protéger ses données de navigation de la manière la plus optimale possible, il va devoir se diriger vers Firefox ou Google Chrome plutôt que Microsoft Edge ou Safari. Nous verrons plus tard dans ce travail s'il est possible d'améliorer leurs performances de confidentialité via certaines extensions.

Quelle est l'option la plus viable entre Google Chrome et Firefox ? Un navigateur se démarque-

Figure 10: Onglet "Sécurité et vie privée" de Firefox

t-il plus que l'autre ? Lorsque nous observons le tableau de résultat du test, nous constatons que les 2 navigateurs bloquent les traqueurs publicitaires et les traqueurs invisibles. Nous voyons aussi que Firefox arrête les logiciels de surveillance pouvant être placés dans les publicités acceptables alors que nous n'avons pas cette information pour Google Chrome. Mais ce n'est pas tout, le navigateur de Mozilla constitue aussi une meilleure option si un utilisateur souhaite protéger ses informations car il offre un guide à la confidentialité lors de son installation et est entièrement Open Source alors que Chrome dépend de Google.

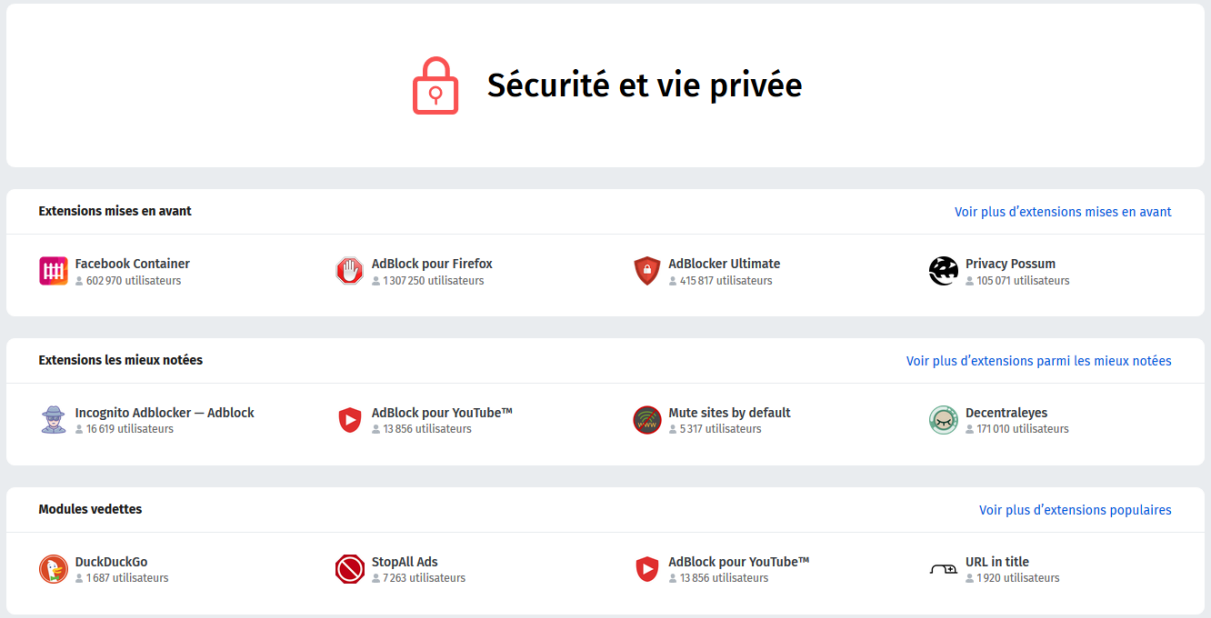
3. Applications

Lorsque nous naviguons sur le Web, il n'est plus possible de se rendre sur une page ou de visionner une vidéo sans avoir de publicités. Ces annonces sont d'autant plus dérangeantes quand elles essayent de nous vendre des produits dont nous venons de discuter avec quelqu'un lors d'une discussion instantanée ou via mail. En effet, il n'est pas rare pour un internaute de se faire traquer lors de sa navigation.

C'est pour cela que la troisième partie de l'analyse concernant les données des utilisateurs se concentre sur les applications que ceux-ci peuvent utiliser pour protéger leurs informations en pleine navigation. Nous nous concentrons principalement sur les différents bloqueurs disponibles en ligne qui permettent d'éviter qu'une entreprise enregistre ce qu'un utilisateur a consulté sur une page ou ses préférences.

Afin de réaliser au mieux cette analyse, certains choix doivent d'abord être posés. Lorsqu'un utilisateur va vouloir se protéger de la récolte de données en ligne et ne plus recevoir de

publicités, il va sans doute se diriger vers un bloqueur de publicités. Mais comment va-t-il effectuer un choix parmi la myriade d'applications disponibles sur les catalogues des navigateurs ? La page « Sécurité et vie privée » des extensions du navigateur Firefox contient un grand nombre d'extensions qu'un internaute peut installer (cf. Figure 10[12]). Mais comment savoir si une extension est meilleure qu'une autre ? Et comment savoir ce qu'une extension peut bloquer ou non ? C'est ce que notre analyse va déterminer.



3.1 Choix des applications

Afin de travailler de la manière la plus optimale possible, nous avons décidé de comparer les applications ayant le plus d'utilisateurs dans la catégorie « Sécurité et vie privée » de Firefox. Le choix de ce navigateur pour l'étude des applications se justifie par ses caractéristiques libres et sa propension à la personnalisation. Nous allons donc comparer 6 bloqueurs de publicités entre eux grâce à une grille de critères pour juger s'il existe une extension plus efficace que les autres.

En premier lieu, nous allons présenter les différents bloqueurs que nous avons choisis et certaines de leurs particularités. Avant de passer à la partie d'analyse à proprement parler, nous présentons et expliquons l'importance des divers critères que nous avons implémenter dans

[12] Mozilla, (2019). *Sécurité et vie privée*. En ligne https://addons.mozilla.org/fr/firefox/extensions/privacy-security/, consulté le 13 avril 2019

notre tableau comparatif. Une fois la grille présentée, nous pourrons comparer les différentes extensions par rapport au contrôle et à la protection qu'elles apportent à l'utilisateur.

Adblock Plus

Avec un peu plus de 10 millions d'utilisateurs rien que sur Firefox, Adblock Plus est certainement le bloqueur de publicités le plus populaire présent sur internet. Il a un fonctionnement plutôt particulier car il est en majorité financé par la publicité. En effet, selon ses créateurs[13], Eyeo, les revenus d'Adblock Plus proviennent principalement de l'initiative de « Publicité Acceptable »[14] qui propose des publicités ne perturbant pas la navigation de l'utilisateur et permettant de faire vivre des sites Web indépendants.

Ublock Origin

Le deuxième bloqueur de publicités le plus utilisé sur internet avec environ 5 millions d'utilisateurs sur Firefox est Ublock Origin. Il ne se définit pas lui-même comme un adblocker mais bien comme un bloqueur beaucoup plus large (Hill, 2019) qui peut aussi faire le travail d'un bloqueur de publicités normal. Contrairement à Adblock Plus, l'extension Ublock ne dépend pas d'une entreprise et ne nécessite donc pas de revenus pour persister. Le code source de celle-ci est d'ailleurs disponible de manière entière sur GitHub.

NoScript Security Suite

La troisième extension la plus populaire sur laquelle nous allons travailler est un peu particulière car elle s'adresse aux utilisateurs un peu plus aguerris. Nous avons choisi de l'analyser car c'est une extension qui est généralement fournie avec le navigateur Tor, reconnu pour ses propriétés de confidentialité. Cette application n'a pas, pour but premier, le blocage de publicités mais bien la protection des internautes contre l'exécution de scripts malveillants pouvant les traquer ou récolter des informations et contre certains types d'attaques.

---

[13] Adblock Plus, (2019) *Publicité acceptable.* https://adblockplus.org/fr/about#acceptableads, consulté le 10 avril 2019

[14] AcceptableAds, (2019) *What are acceptable ads ?*. En ligne https://acceptableads.com/en/about/, consulté le 10 avril 2019

Adblock

Etant le bloqueur de publicités le plus populaire de Google Chrome et de Safari, et se plaçant en quatrième position sur la liste des extensions de Firefox, nous ne pouvions pas passer à côté d'Adblock dans notre analyse.

Ghostery

Développée et éditée par Cliqz, Ghostery est une application qui permet elle aussi de bloquer les publicités et les traqueurs durant la navigation. La particularité de cette extension est qu'elle propose une version gratuite, et une version payante permettant d'obtenir plus de fonctionnalités comme des statistiques sur les publicités rencontrées et bloquées.

Privacy Badger

Pour compléter le tableau comparatif que nous allons établir, nous avons choisi de travailler avec l'extension proposée par l'EFF lors du test effectué sur Panopticlick que nous avons déjà utilisé pour analyser si nos navigateurs étaient fiables ou non lors de la navigation. Cette application possède un peu moins d'utilisateurs que les autres (environ 550 000) mais suscite l'intérêt car elle a été développée par les codeurs de l'association tentant de défendre les droits des consommateurs à la confidentialité la plus grande. Son fonctionnement est un peu particulier car le bloqueur va « apprendre » tout au long de la navigation. En effet, Privacy Badger utilise un algorithme qui va détecter si un domaine nous traque ou non et prendra une décision en conséquence.

3.2 Critères de comparaison

Comme lors de l'analyse des navigateurs, il est important d'établir une grille de critères objectifs qui permet d'obtenir une comparaison des différentes extensions testées. Nous allons donc présenter certaines des caractéristiques de notre tableau en précisant pourquoi nous avons choisi ce critère en particulier.

Contrairement au chapitre précédent, notre analyse est contenue dans un seul tableau récapitulatif permettant de comparer les bloqueurs de publicités et leurs fonctionnalités. Nous allons tenter de savoir si l'utilisateur doit privilégier une application plutôt qu'une autre, et si ses données sont belles et bien protégées.

Aide à la configuration

Le premier critère sur lequel nous avons choisi de comparer nos différentes extensions est l'aide à la configuration. En effet, nous ne pouvons pas négliger les utilisateurs qui ne possèdent pas forcément de connaissance concernant les divers traqueurs existant sur le Web. L'existence d'un tutoriel ou d'une aide lors de l'installation de l'application est un atout non-négligeable.

Fonctionnement

Dans cette caractéristique, nous allons retrouver la manière donc chaque extension fonctionne pour trouver et bloquer les logiciels de surveillance et de traque. Il existe plusieurs types de fonctionnement relativement communs pour les bloqueurs de publicités. En général, les moyens de blocage utilisés sont les suivants :

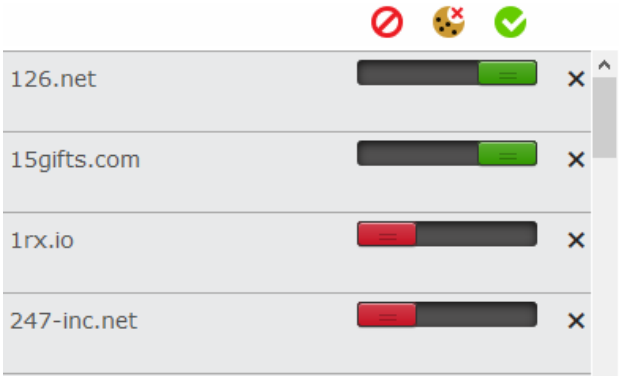
- Utilisation de listes filtrantes : Certains bloqueurs possèdent une série de listes leur dictant les entités qui doivent être bloquées ou non. Les utilisateurs peuvent utiliser les listes précréées par les autres internautes ou bien créer la leur ;
- Filtrage dynamique : Cette option n'est généralement pas disponible par défaut sur les bloqueurs les plus célèbres car elle concerne les utilisateurs les plus aguerris. Ce filtrage permet à l'internaute de choisir lui-même quelles entités il veut bloquer ou non.

Ces deux options sont généralement les plus utilisées parmi les bloqueurs, bien qu'il existe certains cas particuliers pour les utilisateurs s'y connaissant un peu plus.

Possibilité de consulter ce qui a été bloqué

Ce critère possède son importance car il peut aider l'utilisateur à connaître les entités qui suivent son activité en ligne et à se méfier de celles-ci. Certains bloqueurs donnent la possibilité aux internautes d'obtenir la liste des domaines ayant été bloqués (cf. Figure 11).

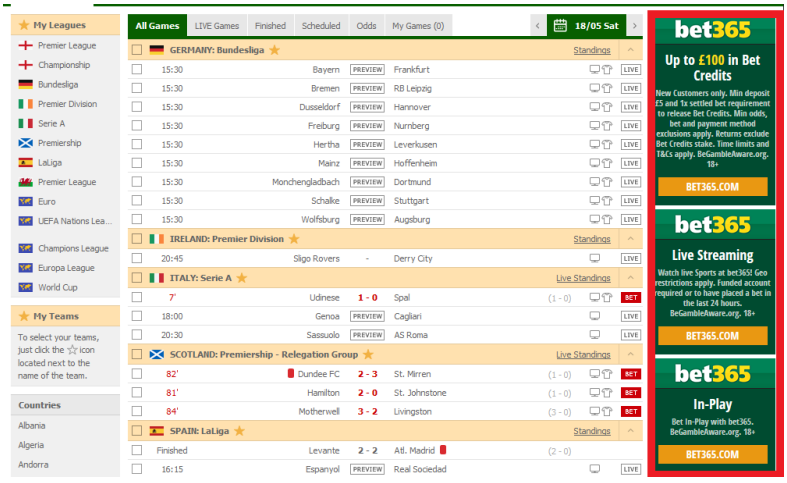
Figure 11: Liste des domaines traceurs Privacy Badger



Laisser passer les publicités acceptables

Comme expliqué plus tôt, il existe certains bloqueurs de publicités qui laissent passer des messages de réclames n'étant pas considérés comme intrusifs à la navigation de l'utilisateur (cf. Figure 12[15]).

Figure 12: Publicités Acceptables



[15] Flash Scores, (2019). Football. En ligne https://www.flashscores.co.uk, consulté le 29 mars 2019

Do Not Track

Tout comme pour les navigateurs que nous avons analysés, il existe certains bloqueurs qui proposent d'envoyer un signal « Do Not Track » qui précise à la page consultée que le visiteur ne souhaite pas être traqué pendant sa navigation. Cela fonctionne seulement pour les domaines qui ont accepté de participer à l'initiative « Do Not Track ».

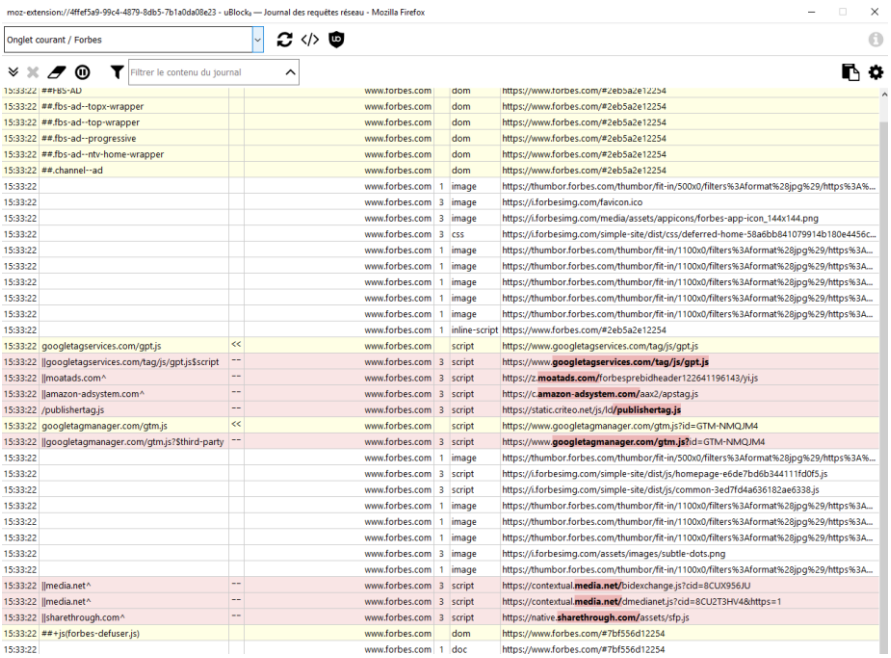
Possibilité de bloquer des scripts

En plus de leurs fonctions de blocage de publicités et de traqueurs en tout genre, certaines extensions ajoutent la possibilité de bloquer l'exécution des scripts effectués par une page. L'activation de cette option peut parfois provoquer le dysfonctionnement de certains sites qui utilisent JavaScript par exemple.

Journal de requêtes réseaux disponible

Lorsque notre navigateur charge un site Web, il va envoyer une requête HTTP au serveur pour obtenir ce qu'il doit afficher. Le serveur va ensuite envoyer la réponse au client contenant la page à afficher. Certains bloqueurs vont donner la possibilité à l'utilisateur d'afficher le journal de l'échange des requêtes entre le serveur et le client (cf. Figure 13).

Figure 13: Journal de requêtes réseaux



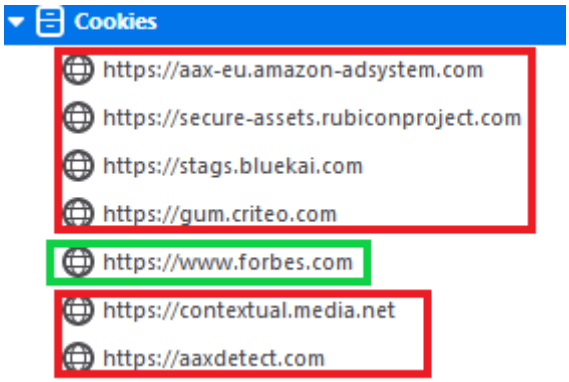
Blocage Pop-up

De la même manière que les différents navigateurs dont nous avons discutés plus tôt dans ce travail, certains bloqueurs de publicités permettent de bloquer les pop-ups qui peuvent être obtenus lors de la navigation. Ce critère signifie que l'extension bloque le pop-up avant son apparition, et non seulement la publicité contenue dans celui-ci.

Blocage Cookies

Ce critère en regroupe deux différents, le blocage des Cookies appelés First-Party et ceux appelés Third-Party. Comme nous l'avons expliqué dans la revue de littérature, les Cookies First-Party appartiennent au domaine du site sur lequel nous nous trouvons tandis que les Third-Party viennent d'entité extérieure et qui traquent généralement notre activité (cf. Figure 14). Nous pouvons voir un exemple de ces deux types de Cookies lors de la visite du site du Forbes[16] . Les Cookies First-party sont en vert tandis que les Third-party sont en rouge.

Figure 14: Cookies rencontrés sur le site du Forbes



Blocage Web Bug

Il existe un type de traqueur particulier appelé « Web Bug » qui se cache généralement sous la forme d'un pixel ou d'une image invisible sur certaines pages Web. C'est pour cela que certains bloqueurs ont mis en place la possibilité de bloquer ce dispositif de traque en plus des publicités et des traqueurs plus « standards ».

[16] Forbes, (2019) *Forbes Now*. En ligne https://www.forbes.com/, consulté le 13 avril 2019

3.3 **Comparaison des applications**

Afin que les résultats des extensions choisies soient garantis représentatifs, les tests ont été réalisés sur le navigateur Mozilla Firefox configuré avec les paramètres par défaut. Chaque fois qu'une application a été analysée, nous avons réinitialisé Firefox pour être certains que les bloqueurs n'interfèrent pas entre eux. De la même manière que pour les navigateurs, nous avons utilisé le test Panopticlick pour certaines parties de notre grille.

Tableau 7: Comparaison Applications

	Adblock Plus	**Ublock Origin**	**NoScript Security Suite**	**Adblock**	**Ghostery**	**Privacy Badger**
Développeurs	EYEO	Raymond Hill	Giorgo Maone	BetaFish Incorporated	Cliqz	EFF
License	GPLv3	GPLv3	GPLv3	GPLv3	MPL 2.0	GPLv3
Aide à la configuration	Non	Non	Non	Non	Oui	Oui
Fonctionnement	Listes	Listes ou filtrage dynamique	Listes	Listes	Listes	Apprends pendant la navigation + Filtrage dynamique
Consultation de ce qui a été bloqué	Non	Oui	Oui	Non	Oui	Oui
Publicité acceptables	Oui	Non	Non	Oui	Non	Non
Do Not Track	Non	Non	Non	Non	Non	Oui
Blocage de scripts	Non	Oui	Oui	Non	Non	Non
Journal de requêtes réseaux disponibles	Non	Oui	Non	Non	Non	Non
Blocage Pop-up	Oui	Oui	Oui	Oui	Non	Non
Blocage Cookies First-Party	Non	Non	Non	Non	Non	Non
Blocage Cookies Third-Party	Partiel	Oui	Oui	Partiel	Oui	Oui
Blocage Web Bug	Non	Oui	Non	Partiel	Non	Oui

Contrairement aux différents navigateurs que nous avons comparés entre eux, les bloqueurs de publicités obtiennent des résultats relativement différents concernant les tests effectués et les fonctionnalités qu'ils proposent (cf. Tableau 7). Mais il existe tout de même 2 critères que toutes ces extensions ont en commun :

- Elles sont toutes en licence libre et leur code est accessible
- Aucune d'entre elle ne bloque les Cookies First-Party

La raison simple pour laquelle nos bloqueurs laissent passer les Cookies First-Party, c'est-à-dire, ceux appartenant au domaine visité par l'utilisateur : ils ne sont pas utilisés pour traquer la navigation d'un individu mais bien pour proposer du contenu personnalisé relatif uniquement au domaine où se trouve le Cookie. Malgré ces 2 critères communs, nous pouvons tout de même séparer les bloqueurs étudiés en deux groupes bien distincts. Le premier se constitue d'Adblock Plus et Adblock, tandis que l'autre contient le restant des bloqueurs.

Cette séparation s'opère car nous pouvons tout de suite observer qu'Ablock Plus et Adblock sont moins performants que leurs concurrents. En effet, ils laissent tous deux passer les publicités acceptables et ne bloquent que partiellement les Cookies venant de domaines tiers. Nous pouvons observer la différence dans les illustrations ci-dessous pour le site du New York Times[17] avec Ublock Origin, Adblock, et Adblock Plus (cf. Figure 15 à 17). Nous voyons très clairement qu'Adblock et Adblock Plus laissent passer toute une série de Cookies de domaines tiers qui sont bloqués par Ublock Origin.

Figure 15: Cookies avec Ublock Origin

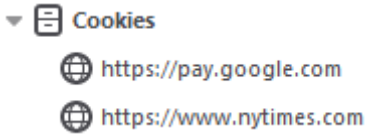


Figure 16 : Cookies avec Adblock

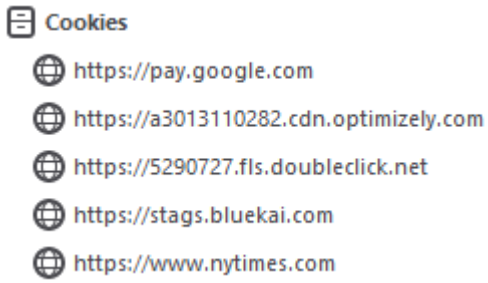
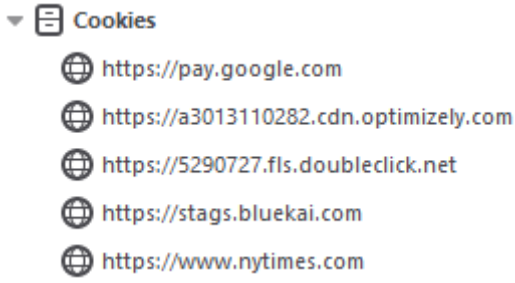


Figure 17: Cookies avec Adblock Plus



[17] New York Times, (2019) *The New York Times*. En ligne https://www.nytimes.com/, consulté le 13 avril 2019

3.4 **Quelle extension choisir ?**

Nous pouvons, de prime abord, éliminer le premier groupe d'applications, contenant Adblock et Adblock Plus, des extensions à choisir pour les utilisateurs voulant conserver leurs données de navigations privées. En effet, ces extensions proposent une protection limitée si nous les comparons à leurs contreparties. Ce qui est très surprenant sachant qu'Adblock Plus est le bloqueur le plus utilisé sur Firefox tandis qu'Adblock est celui le plus utilisé sur Google Chrome et sur Safari.

Il nous reste donc les 4 autres options proposant une protection nettement plus complète que ses concurrents. L'application que nous pouvons éliminer en premier lieu de ce quatuor de tête n'est autre que Ghostery. Effectivement, comparés aux 3 autres candidats, Ghostery est l'extension proposant le moins de fonctionnalités, permettant seulement de bloquer les Cookies provenant de domaine tiers sans bloquer d'éventuels traqueurs.

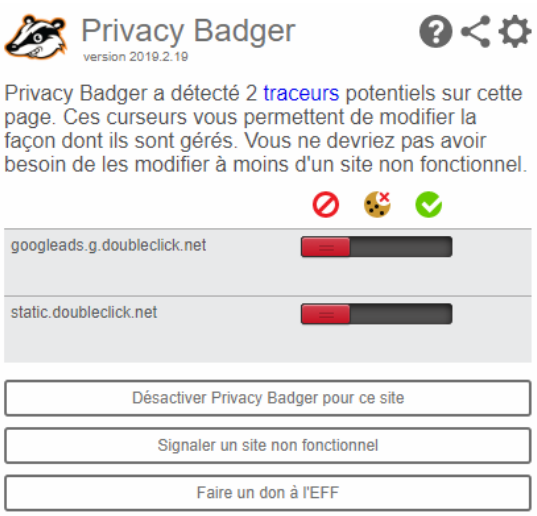
A la suite de Ghostery, nous retrouvons NoScript Security Suite. En effet, cette application se concentre principalement sur le blocage de différents scripts malveillants et ne vise pas les divers traqueurs pouvant suivre les internautes lors de la navigation. De plus, cette extension s'adresse à un public d'utilisateurs beaucoup plus avisé sachant comment la configurer pour une utilisation efficace.

Cela nous laisse 2 options finales, Ublock Origin, et Privacy Badger. Ces deux extensions offrent, toutes deux, une protection complète contre les différentes publicités ciblées qu'un utilisateur rencontre en ligne, et contre les traqueurs. Leur principale différence réside dans leur manière respective de fonctionner et l'aide qu'ils proposent à l'utilisateur.

Effectivement, Privacy Badger donne la possibilité d'apprendre de quelle manière l'application procède[18] pour permettre à l'internaute de naviguer de manière sécurisée. L'utilisateur n'a absolument rien à faire. L'extension va apprendre lors de la navigation et décider de bloquer un traqueur à partir du moment où celui-ci va être repéré sur 3 pages Web différentes. En plus de cette fonctionnalité, l'application donne aussi la possibilité de filtrer dynamiquement les différents traqueurs en choisissant entre trois actions : bloquer le domaine, bloquer les Cookies, ou autoriser le domaine (cf. Figure 18).

[18] cf. Figure 24 Annexe 3

Figure 18: Filtrage dynamique de Privacy Badger

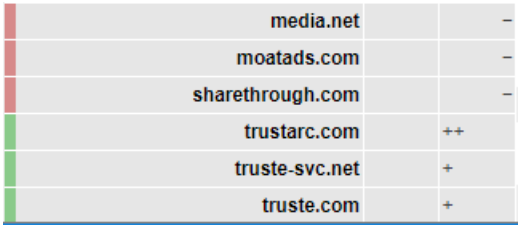


En plus de ces fonctions, Privacy Badger est la seule extension qui respecte entièrement l'accord Do Not Track parmi toutes celles que nous avons analysées lors de ce travail.

De son côté, Ublock Origin propose à ses utilisateurs 2 manières d'améliorer leur sécurité lors de la navigation. La première constitue la configuration de base de l'application et utilise un système de listes de domaines à bloquer. L'application possède plusieurs listes relativement complètes et donne aussi la possibilité à l'internaute de créer les siennes. La deuxième peut être activée via les options avancées et propose de filtrer dynamiquement les domaines visités de la même manière que Privacy Badger. En plus de cette fonctionnalité, l'utilisateur peut savoir combien de requêtes réseaux ont été faites par le domaine (cf. Figure 19) :

- - ou + de 1 à 9 requêtes refusées ou acceptées
- -- ou ++ de 10 à 99 requêtes refusées ou acceptées
- --- ou +++ plus de 100 requêtes refusées ou acceptées

Figure 19: Filtrage Dynamique Ublock Origin



En plus de ces options, Ublock permet aussi de bloquer JavaScript et l'apparition de pop-ups intempestifs, ce qui n'était pas possible avec Privacy Badger. Il constitue donc l'extension la plus complète pour un utilisateur qui souhaiterait ne pas laisser ses données de navigation dans les mains de n'importe quel domaine.

Nous apprécierons surtout son accessibilité, que cela soit pour un internaute lambda qui pourra bénéficier des listes préinstallées, ou pour un utilisateur avancé qui pourra lui-même configurer sa propre protection. Privacy Badger reste, tout de même, une option viable grâce à sa capacité d'apprentissage lors de la navigation en permettant également d'apporter une protection décente à un utilisateur moins expérimenté.

Conclusion

Lors de ce chapitre, nous avons pu effectuer des analyses sur les différents outils constituant nos 3 catégories. Premièrement, nous avons comparé les diverses options de confidentialité proposées par nos systèmes d'exploitation. Il a été constaté que chaque OS possédait une approche différente de la confidentialité ; certains préférant se concentrer sur la sécurité des informations de l'utilisateur. De la même manière, les grandes instances qui gèrent ces systèmes d'exploitation ont une attitude différente concernant les données collectées. Même si elles respectent toutes la loi et proposent des possibilités de gestion des informations personnelles. C'est pour cela que nous en avons déduit que Linux était la meilleure option concernant la confidentialité, bien que Windows et MacOS restent des choix viables grâce à leurs paramètres de sécurité.

Ensuite, nous avons étudié les navigateurs les plus utilisés sur le Web. Dans un premier temps, une comparaison des configurations de confidentialité par défaut a été effectuée pour ensuite donner lieu à un test comparant leur protection. Il s'est avéré qu'aucun des navigateurs n'offre une protection suffisante, lorsqu'utilisé avec les paramètres de base. A la suite de ces résultats, nous avons paramétré nos différents outils pour qu'ils procurent la plus haute défense possible. Cette fois-ci, il est ressorti du test que Safari et Edge ne possédaient qu'une protection partielle comparés à Chrome et Firefox. Le navigateur proposé par Mozilla s'est tout de même démarqué de son concurrent car il reste Open Source et propose un guide de la confidentialité qui n'apparaît nulle part ailleurs.

Enfin, la troisième partie de notre étude était dédiée aux applications permettant de bloquer les traqueurs et publicités rencontrées lors de la navigation. Dans la même démarche que les navigateurs, les bloqueurs les plus populaires ont été analysés. Nous avons observé qu'il existait une hiérarchie relativement claire parmi ceux-ci. Certains d'entre eux proposent une protection faible comparée aux autres et ne permettent pas à l'utilisateur de naviguer sans se soucier des traqueurs. Plusieurs extensions se sont démarquées car elles proposent un grand nombre de

fonctionnalités, et permettent aux utilisateurs expérimentés de personnaliser leur protection. Grâce à ces résultats, nous avons pu déterminer qu'Ublock Origin et Privacy Badger constituaient de bonnes protections pour les utilisateurs.

Chapitre 4 : Discussion

Maintenant que nous avons terminé les analyses de tous les outils que nous avions choisis d'étudier dans ce travail, nous pouvons enfin émettre une réponse aux différentes hypothèses que nous avons posées.

1. Les utilisateurs ne sont pas obligés de transmettre des données

Tout au long de ce mémoire, nous avons tenté de déterminer si les utilisateurs pouvaient limiter les traces numériques qu'ils laissaient lors de leur navigation et s'ils pouvaient s'abstenir de transmettre des données les concernant en ligne. Nous avons distingué 2 situations distinctes de transmission de données à laquelle chaque utilisateur est confronté. Nous allons donc répondre à cette hypothèse pour chacune de ces situations.

Données récoltées lors de l'utilisation

Dans cette situation, nous parlons de toutes les données qui sont récoltées lorsque l'utilisateur effectue n'importe quelle tâche sur sa machine, qu'elle soit en ligne ou non. Cela concerne donc principalement les informations qui sont collectées par les systèmes d'exploitation que nous avons étudiés. Nous avons découvert qu'il était possible pour un utilisateur de ne transmettre aucune donnée s'il choisissait d'utiliser Linux.

Cependant, il est ressorti que les utilisateurs de MacOS et Windows étaient toujours obligés d'au moins envoyer certaines données de diagnostics comme des rapports de performance et d'autres données techniques à Apple ou Microsoft. Même si des options de modification et de suppression des données récoltées leur étaient proposées. Nous avons aussi remarqué que plus un utilisateur envoyait d'informations, plus il avait accès à un service personnalisé et à différentes fonctionnalités (comme Cortana).

Données récoltées lors de la navigation

Le second cas concerne principalement les informations collectées lorsqu'un utilisateur se trouve sur le Web. C'est ce que nous avons analysé en comparant les divers navigateurs et extensions de ceux-ci. Nous avons découvert qu'il était possible de configurer son navigateur pour qu'il ne transmette plus les données de l'internaute. A cela s'est ajouté la possibilité d'installer toute une panoplie d'applications permettant de renforcer la sécurité et la confidentialité des navigateurs ne proposant pas une sécurité satisfaisante contre les traqueurs généralement présents. Mais malheureusement, aucune des applications utilisées ne protégeaient du fingerprinting. Ce qui signifie que notre navigateur est toujours identifié comme unique.

2. Les utilisateurs peuvent avoir le contrôle sur les données qu'ils laissent filtrer en ligne

Cette hypothèse est directement liée à la première. Nous pouvons donc y répondre de la même manière que la précédente, en différenciant les données transmises lors de l'utilisation de la machine de l'utilisateur et celles transmises lors de la navigation sur le Web.

Données récoltées lors de l'utilisation

Lors de nos analyses, nous avons observé que les utilisateurs se voyaient proposer différentes options quant à la gestion de leurs informations. Pour nos 3 systèmes d'exploitation, chaque utilisateur est libre de choisir les données personnelles qu'il souhaite envoyer à l'entreprise responsable de l'OS. De plus, chaque information transmise peut être modifiée ou supprimée par le client à n'importe quel moment. Les utilisateurs possèdent donc un certain niveau de contrôle sur les informations qu'ils décident de transmettre, même si cela se fait parfois au dépend de certaines fonctionnalités.

Données récoltées lors de la navigation

Nous avons remarqué lors de notre étude que la situation est différente lorsque l'utilisateur navigue en ligne. En effet, un internaute est contraint d'envoyer certaines informations lorsqu'il

consulte une page Web pour pouvoir obtenir un contenu lui correspondant. La désactivation de tous les scripts provoquant le dysfonctionnement de certains sites, l'utilisateur ne peut pas, par exemple, se passer de l'utilisation de JavaScript.

En revanche, nous avons remarqué qu'il est possible pour l'utilisateur de se protéger de manière complète contre tous les traqueurs présents sur internet via toute une panoplie d'applications améliorant sa sécurité, et de bloquer les Cookies pouvant contenir des informations concernant ses habitudes de navigation. Certaines extensions proposent même à l'utilisateur de configurer lui-même les domaines qu'il souhaite bloquer ou non.

3. N'importe quel outil n'offre pas une protection des données équivalente à l'utilisateur

Notre dernière hypothèse concerne directement tous les outils dont nous avons discutés. Nous avons observé durant nos comparaisons qu'il existe une hiérarchie relativement claire parmi les applications disponibles.

Concernant les systèmes d'exploitation, cela dépend principalement des préférences des utilisateurs. Linux constitue un choix de qualité pour toute personne ne voulant partager aucune de ses informations, mais les deux autres options présentées offrent toute deux, une protection décente.

En ce qui concerne les navigateurs, il en est ressorti que Firefox et Google Chrome se présentaient tout deux comme de meilleures options si utilisées sans aucune extension. Si un utilisateur possède une préférence pour Safari ou Microsoft Edge, nous lui conseillons de coupler à son utilisation une des extensions que nous avons proposées lors de ce travail.

La dernière comparaison que nous avons effectuée nous a permis d'obtenir un classement relativement précis sur les applications à privilégier. Parmi les extensions les plus utilisées sur le Web, Ublock Origin et Privacy Badger sont celles qui offrent la meilleure protection lors de la navigation. Ghostery et NoScript Suite offrent, tous deux, une protection correcte mais devraient être complétés avec d'autres extensions. Quant à AdBlock Plus et AdBlock, ils proposent chacun une protection relativement faible, en laissant passer la majorité des traqueurs de domaines tiers et certaines publicités.

Conclusion

Résumé de notre étude

A travers l'étude réalisée, nous avons tenté d'établir une analyse et une comparaison des outils permettant de limiter ses traces numériques. Dans un premier temps, il a fallu se renseigner sur la littérature traitant du sujet pour recadrer précisément notre problématique. C'est à travers le premier chapitre que nous détaillons l'approche des différents experts concernant notre problématique. Cette analyse approfondie nous a permis de dégager certaines hypothèses. Nous nous sommes, avant tout, interrogés sur l'obligation qu'ont les utilisateurs à transmettre leurs informations. Ensuite, nous supposons qu'ils possèdent pleinement la capacité de contrôler les données qu'ils souhaitent filtrer en ligne, quel que soit l'outil utilisé. En dernier lieu, nous nous questionnons sur la capacité des outils à offrir la même protection aux utilisateurs. Il existerait une hiérarchie prédéfinie dans les applications à utiliser et à conseiller aux internautes.

A la suite de cette toute première partie, toutes une série d'analyses ont été effectuées pour pouvoir mettre à l'épreuve les différentes hypothèses posées. Pour pouvoir obtenir des résultats d'analyse concluants, différents tests ont été effectués sur les paramètres de confidentialité des systèmes d'exploitation utilisés par la majorité des utilisateurs, sur les configurations des navigateurs et leur sécurité, et pour finir, sur diverses extensions proposées pour améliorer la conservation de la vie privée des internautes et compléter la protection offerte par leur navigateur. Des grilles comparatives ont été établies sur base de ces tests pour déterminer les meilleures options parmi les outils étudiés.

Les résultats de tous ces différents tests et de ces différentes grilles de comparaison nous ont permis d'observer que les individus ne sont pas obligés de transmettre leurs informations personnelles en ligne, même si certains outils leur proposent plus de fonctionnalités s'ils choisissent de partager leurs données. De plus, il a été constaté que les consommateurs peuvent aussi choisir de filtrer les informations qu'ils souhaitent communiquer lors de leur navigation. Pour finir, il a été montré qu'il existe bel et bien une hiérarchisation dans les outils qui sont mis à la disposition des utilisateurs. Certains outils sont plus performants et offrent une meilleure protection aux internautes.

L'analyse de nos différentes grilles de comparaison et des paramètres de nos outils nous ont permis de répondre en partie à notre problématique. Cependant, celle-ci possède tellement

d’aspects différents qu’il n’est pas possible d’y répondre de manière complète en une seule analyse.

Limites

Malgré la rigueur et l’esprit critique utilisés, il est évident que l’étude réalisée présente tout de même certaines limites pouvant apporter des modifications aux résultats et conclusions ayant été établis. Bien entendu, il est impossible de comprendre toutes les particularités pouvant apporter de nouveaux aspects à notre questionnement.

En amont, le choix fut porté sur la récolte des données découlant directement des actions de l’utilisateur et non transmises directement par ceux-ci. C’est pour cela que les moteurs de recherche ont, par exemple, été écartés des outils à étudier dans le cadre de ce travail. Si cette partie avait été intégrée à l’étude effectuée, il est probable que les résultats obtenus aient été très différents de ceux obtenus actuellement. C’est dans cette optique qu’il faut émettre certaines réserves quant à nos résultats car ils pourraient varier avec d’autres types de données récoltées.

Si nous désirons rester critiques, de nombreux outils n’ont pas été étudiés. Il s’est effectué un tri dans les différents types testés pour permettre de donner des résultats et des analyses plus précises. Nous avons choisi de travailler sur les outils utilisés par le plus grand nombre d’utilisateurs. Il est possible qu’il en existe certains méconnus qui présenteraient des résultats et des fonctionnalités différentes de ceux que nous avons étudiés. Il faudra garder cela à l’esprit lors de la lecture des conclusions.

Pour terminer, il existe d’autres types d’applications permettant de garantir l’anonymité des utilisateurs que nous avons choisis de ne pas étudier. En effet, il est important de bien comprendre la scission existante entre la notion de confidentialité et d’anonymité. Notre analyse avait pour but de comparer les outils qui permettent à un utilisateur de limiter ses traces numériques et non de les faire disparaître complètement. Il n’était en aucun cas question de déterminer si un individu avait la possibilité de se cacher ou non en ligne, mais bien d’analyser comment certains outils pouvaient modifier les données récoltées lors d’une utilisation normale. Les résultats obtenus auraient très bien pu être très différents si nous avions choisi de traiter un tout autre type d’outil, uniquement dans le but de masquer notre présence en ligne.

Suggestions de recherche future

Le sujet du respect de la vie privée en ligne et de la confidentialité est tellement large que nous pourrions en débattre pendant des heures. Le développement de l'internet des objets ouvre d'autres possibilités de recherche par rapport aux multiples objets connectés pouvant dorénavant récolter des données sur leurs utilisateurs. Analyser les applications disponibles sur Smartphone ou même les options de confidentialité et de sécurité proposées par ceux-ci pourrait être très intéressant. La technologie reste en perpétuelle évolution, ce qui permet d'imaginer différentes perspectives d'avenir.

Bibliographie

Articles Scientifiques

Arnaud, M. (2009). Authentification, Identification, et Tiers de Confiance. *Hermès, La Revue 53 (1)*, 127-136. En ligne https://www.cairn.info/revue-hermes-la-revue-2009-1-page-127.htm

Barth, S., & de Jong, M.D.T. (2017) The privacy-paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review. *Telematics and Informatics 34*, 1038-1058. DOI: 10.1016/j.tele.2017.04.013

Belanger, F., & Crossler, R.E. (2018). Dealing with digital traces: Understanding protective behaviors on mobile devices. *Journal of Strategic Information Systems 28 (1)*, 34-49. DOI: 10.1016/j.jsis.2018.11.002

Bradbury, D. (2014), Anonymity and privacy: a guide for the perplexed. *Network Security 10*, 10-14. DOI: 10.1016/S1353-4858(14)70102-3

Bujlow, T., Carela-Espanol, V., Solé-Pareta, J., & Barlet-Ros P. (2015) Web Tracking: Mechanisms, Implications, and Defenses, En ligne https://arxiv.org/abs/1507.07872

Chen, D., & Zhao H. (2012) Data Security and Privacy Protection Issues in Cloud Computing. Communication présentée au *International Conference on Computer Science and Electronics Engineering*, 647-651. DOI: 10.1109/ICCSEE.2012.193

Englehardt, S., & Narayanan, A. (2016) Online tracking: A 1-million-site Measurement and Analysis. Communication présentée au *ACM SIGSAC Conference on Computer and Communications Security* 1388-1401. DOI: 10.1145/2976749.2978313

Estrada-Jiménez, J., Parra-Arnau, J., Rodriguez-Hoyos, A., & Forné, J. (2017) Online advertising: Analysis of privacy threats and protection approaches. *Computer Communications 100*, 32-51. DOI: 10.1016/j.comcom.2016.12.016

Galinon-Mélénec, B., & Zlitni, S. (2013) L'Homme-trace, producteur de traces numériques. *Traces numériques, de la production à l'interprétation*, 7-19. En ligne https://halshs.archives-ouvertes.fr/halshs-01080027

Galinon-Mélénec, B. (2015) A la recherche de la trace. *Communication et organisation 47*, 31-50. En ligne http://communicationorganisation.revues.org/4876

Gandomi, A., & Haider, M. (2015) Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management 35*, 137-144. DOI: 10.1016/j.ijinfomgt.2014.10.007

Heng, X., Crossler, R.E., & Bélanger, F. (2012) A Value Sensitive Design Investigation of Privacy Enhancing Tools in Web Browsers. *Decision Support Systems 54*, 424-433. DOI: 10.1016/j.dss.2012.06.003

Heurix, J., Zimmerman, P., Neubauer, T., & Fenz S. (2015) A taxonomy for privacy enhancing technologies. *Computers & Security 53*, 1-17. DOI: 10.1016/j.cose.2015.05.002

Hou, J., Qu, L., & Shi, W. (2018) A survey on internet of things security from data perspectives. *Computers Networks 148*, 295-306. DOI: 10.1016/j.comnet.2018.11.026

Janecek, V. (2018) Ownership of personal data in the Internet of Things. *Computer Law & Security Review 34*, 1039-1052. DOI: 10.1016/j.clsr.2018.04.007

Karuna Jyothi, K., & Indira Reddy, B. (2018) Study on Virtual Private Network (VPN), VPN's Protocols And Security. *International Journal of Scientific Research in Computer Science 3 (5)*, 919-932. En ligne http://ijsrcseit.com/paper/CSEIT1835225.pdf

Kessous, E., & Rey, B. (2009) Economie numérique et vie privée. *Hermès, La Revue 53(1)*, 49-54. En ligne https://www.cairn.info/revue-hermes-la-revue-2009-1-page-49.htm

Kobusinska, A., Pawluczuk, K., & Brzezinski, J. (2018) Big Data fingerprinting information analytics for sustainability. *Future Generation Computer Systems 86*, 1321-1337. DOI: 10.1016/j.future.2017.12.061

Kouicem, D.E., Bouabdallah, A., & Lakhlef, H. (2018) Internet of things security: A top-down survey. *Computer Networks 141*, 199-221. DOI: 10.1016/j.comnet.2018.03.012

Li, B., Erdin, E., Gunes, M.H., & Shipley, T. (2013) An overview of anonymity technology usage. *Computer Communications 36*, 1269-1283. DOI: 10.1016/j.comcom.2013.04.009

Longworth, J. (2018) VPN: From an obscure network to a widespread solution. *Computer Fraud & Security 4*, 14-15. DOI: 10.1016/S1361-3723(18)30034-4

Lopez, J., Rios, R., Bao, F., & Wang, G. (2017) Evolving privacy: From sensors to the Internet of Things. *Future Generation Computer Systems 75*, 46-57. DOI: 10.1016/j.future.2017.04.045

Malandrino, D., Scarano, V. (2013) Privacy leakage on the Web: Diffusion and countermeasures. *Computer Networks 57*, 2833-2855. DOI: 10.1016/j.comnet.2013.06.013

Maline, L., Hajny, J., Fujdiak, R., & Hosek, J. (2016) On perspective of security and privacy-preserving solutions in the internet of things. *Computer Networks 102*, 83-95. DOI: 10.1016/j.comnet.2016.03.011

Manach, J-M. (2009) Contourner les systèmes de traçabilité. *Hermès, La Revue 53 (1)*, 167-173. En ligne https://www.cairn.info/revue-hermes-la-revue-2009-1-page-167.html

Mayer, J. R., Mitchell, J. C. (2012) Third-Party Web Tracking: Policy and Technology. *IEEE Symposium on Security and Privacy*, 413-427. DOI: 10.1109/SP.2012.47

Merzeau, L. (2009) Du signe à la trace : L'information sur mesure. *Hermès, La Revue 53 (1)*, 21-29. En ligne https://www.cairn.info/revue-hermes-la-revue-2009-1-page-21.html

Mille, A. (2013) Des traces à l'ère du Web. *Intellectica 59 (1)*, 7-28. En ligne https://www.persee.fr/doc/intel_0769-4113_2013_num_59_1_1083

Oussous, A., Benjelloun, F-Z., Lahcen, A.A., Belfkih, S. (2018) Big Data technologies: A survey. *Journal of King Saud University – Computer and Information Sciences 30*, 431-448. DOI: 10.1016/j.jksuci.2017.06.001

Perriault, J. (2009) Traces numériques personnelles, incertitude et lien social. *Hermès, La Revue 53 (1)*, 13-20. En ligne https://www.cairn.info/revue-hermes-la-revue-2009-1-page-13.html

Poullet, Y. (2018) Is the general data protection regulation the solution? *Computer Law & Security Review 34*, 773-778. DOI: 10.1016/j.clsr.2018.05.021

Pras, B. (2012) Entreprise et Vie Privée : Le « privacy paradox » et comment le dépasser ?. *Revue française de gestion 224 (5)*, 87-94. En ligne https://www.cairn.info/revue-francaise-de-gestion-2012-5-page-87.html

Rey, B. (2014) Les intelligences numériques des informations personnelles : Vers un changement de perspective pour garantir le droit à la vie privée ?. *Les Cahiers du numérique 10 (1)*, 9-18. En ligne https://www.cairn.info/revue-les-cahiers-du-numerique-2014-1-page-9.html

Roesner, F., Kohno, T., & Wetherall, D. (2012) Detecting and Defending Against Third-Party Tracking on the Web. *9th USENIX Symposium on Networked Systems Design and*

Implementation, En ligne https://www.usenix.org/conference/nsdi12/technical-sessions/presentation/roesner

Rossi, J., & Bigot, J.-E. (2018) Traces numériques et recherche scientifique au prisme du droit des données personnelles. *Les Enjeux de l'information et de la communication 19 (2)*, 167-177. En ligne https://www.cairn.info/revue-les-enjeux-de-l-information-et-de-lacommunication-2018-2-page-161.html

Ruiz-Martinez, A. (2012) A survey on solutions and main free tools for privacy enhancing Web communications. *Journal of Network and Computer Applications 35*, 1473-1492. DOI: 10.1016/j.jnca.2012.02.011

Sanchez, D., & Viejo, A. (2018) Privacy-perserving and advertising-friendly Web surfing. *Computer Communications 130*, 113-123. DOI: 10.1016/j.comcom.2018.09.002

Serres, A. (2012) Problématiques de la trace à l'heure du numérique. *Sens-Dessous 10 (1)*, 84-94. DOI: 10.3917/sdes.010.0084

Sfar, A.R., Natalizio, E., Challal, Y., & Chtourou, Z. (2018) A roadmap for security challenges in the Internet of Things. *Digital Communications and Networks 4*, 118-137. DOI: 10.1016/j.dcan.2017.04.003

Smith, B. (2007) A Quick Guide to GPLv3. *GLPv3 Free Software Free as in Freedom*, 1-6. En ligne https://www.gnu.org/licenses/quick-guide-gplv3.fr.html

Su, J., Li, Z., Grumbach, S., Ikram, M., Salamatian, K., & Xie, G (2019) A Cartography of Web Tracking using DNS Records. *Computer Communications 134*, 83-95. DOI: 10.1016/j.comcom.2018.11.008

Sylvestre, G. (2017) Être furtif et anonyme sur internet. *I2D – Information, données & documents 54 (3)*, 34-35. En ligne https://www.cairn.info/revue-i2d-information-donnees-etdocuments-2017-3-page-34.html

Zhou, D., Yan, Z., Fu, Y., & Yao, Z. (2018) A survey on network data collection. *Journal of Network and Computer Applications 116*, 9-23. DOI: 10.1016/j.jnca.2018.05.004

Livres

Fischer-Hbner, S., & Berthold S. (2017) Privacy-Enhancing Technologies. *Computer and Information Security Handbook,* 759-778. Britain: John R. Vacca DOI: 10.1016/B978-0-12-803843-7.00053-3

Hassan, N. A., & Hijazi, R. (2017) *Digital Privacy and Security Using Windows.* Berkeley: CA Apress. DOI: 10.1007/978-1-4842-2799-2

Jajodia, S., & Zhou, J. (2010) Security and Privacy in Communication Networks. *International ICST Conference*, Berlin: Springer.

Lazar, J., Feng, J.H., & Hochheiser, H. (2017) Chapter 12 – Automated data collection methods. *Research Methods in Human Computer Interaction (2nd edition)*, 329-368. Britain: Elsevier DOI: 10.1016/B978-0-12-805390-4.00012-1

Sites Internet

AcceptableAds, (2019). *What are acceptable ads ?*. En ligne https://acceptableads.com/en/about/, consulté le 10 avril 2019

Adblock Plus, (2019). *Publicité acceptable.* https://adblockplus.org/fr/about#acceptableads, consulté le 10 avril 2019

Adobe, (2019). *Adobe Flash Player*. En ligne https://get.adobe.com/fr/flashplayer/about/, consulté le 14 mai 2019

ADsafe, (2019). *Making JavaScript Safe for Advertising*. En ligne http://www.adsafe.org/, consulté le 9 avril 2019

Apple, (2019). *Data and Privacy*. En ligne https://privacy.apple.com, consulté le 5 avril 2019

Apple, (2019). *Gérer votre confidentialité*. En ligne https://www.apple.com/befr/privacy/manage-your-privacy/, consulté le 5 avril 2019

Apple, (2019). *Notre approche de la confidentialité.* En ligne https://www.apple.com/befr/privacy/approach-to-privacy/, consulté le 5 avril 2019

BrowserLeaks, (2019). *Web Browser Security*. En ligne https://browserleaks.com/, consulté le 18 février 2019

Canonical, (2019). *Data privacy*. En ligne https://www.ubuntu.com/legal/data-privacy, consulté le 5 avril 2019

Cooke, W. (2018). *A first look at desktop metrics*. En ligne https://blog.ubuntu.com/2018/06/22/a-first-look-at-desktop-metrics, consulté le 5 avril 2019

Duffez, O. (2018) *Tous les résultats financiers Google*. En ligne https://www.Webrankinfo.com/dossiers/google/resultats-financiers, consulté le 15 avril 2019

Electronic Frontier Foundation, (2019). *Is your browser safe against tracking ?*. En ligne https://panopticlick.eff.org, consulté le 20 mars 2019

Electronic Frontier Foundation, (2019) *Privacy Badger*. En ligne moz-extension://7e372406-c8f9-465d-b467-ede4c31a6f62/skin/firstRun.html, consulté le 15 avril 2019

Electronic Frontier Foundation, (2019). *The leading nonprofit defending digital privacy, free speech, and innovation.* En ligne https://www.eff.org/, consulté le 12 mars 2019

Flash Scores, (2019). *Football*. En ligne https://www.flashscores.co.uk, consulté le 29 mars 2019

Forbes, (2019) *Forbes Now*. En ligne https://www.forbes.com/, consulté le 13 avril 2019

FSF, (2019) *Graticiel (ou freeware)*. En ligne http://www.gnu.org/philosophy/categories.html#freeware, consulté le 3 mai 2019

Garay, J. (2011). *Anti-Tracking : Mozilla dévoile sa solution pour Firefox*. En ligne https://www.generation-nt.com/firefox-mozilla-tracking-http-publicite-ciblee-actualite-1147851.html, consulté le 7 mars 2019

Hill, R. (2019). *uBlock Origin - An efficient blocker for Chromium and Firefox. Fast and lean.* En ligne https://github.com/gorhill/uBlock, consulté le 10 avril 2019

Kemp, S. (2019) *Digital 2019*. En ligne https://wearesocial.com/global-digital-report-2019, consulté le 15 avril 2019

Microsoft, (2019). *Déclaration de confidentialité Microsoft*. En ligne https://privacy.microsoft.com/fr-FR/privacystatement#mainwindowsmodule, consulté le 5 avril 2019

Microsoft, (2019). *Garder le contrôle de votre confidentialité*. En ligne https://account.microsoft.com/account/privacy?refd=privacy.microsoft.com&ru=https%3A%2

F%2Faccount.microsoft.com%2Fprivacy%3Frefd%3Dprivacy.microsoft.com&destrt=privacy-dashboard, consulté le 5 avril 2019

Microsoft, (2019). *La confidentialité chez Microsoft*. En ligne https://privacy.microsoft.com/fr-fr, consulté le 5 avril 2019

Mozilla, (2019). *JavaScript*. En ligne https://developer.mozilla.org/fr/docs/Web/JavaScript, consulté le 7 avril 2019

Mozilla, (2019). *Firefox – Politique de confidentialité*. En ligne https://www.mozilla.org/fr/privacy/firefox/, consulté le 7 avril 2019

Mozilla, (2019). *Sécurité et vie privée*. En ligne https://addons.mozilla.org/fr/firefox/extensions/privacy-security/, consulté le 13 avril 2019

NetMarketShare, (2019). *Browser Market Share*. En ligne https://netmarketshare.com/?options=%7B%22filter%22%3A%7B%22%24and%22%3A%5B%7B%22deviceType%22%3A%7B%22%24in%22%3A%5B%22Desktop%2Flaptop%22%5D%7D%7D%5D%7D%2C%22dateLabel%22%3A%22Trend%22%2C%22attributes%22%3A%22share%22%2C%22group%22%3A%22browser%22%2C%22sort%22%3A%7B%22share%22%3A-1%7D%2C%22id%22%3A%22browsersDesktop%22%2C%22dateInterval%22%3A%22Monthly%22%2C%22dateStart%22%3A%222018-05%22%2C%22dateEnd%22%3A%222019-04%22%2C%22segments%22%3A%22-1000%22%7D, consulté le 15 mai 2019

NetMarketShare, (2019). *Operating System Market Share*. En ligne https://netmarketshare.com/operating-system-market-share.aspx?options=%7B%22filter%22%3A%7B%22%24and%22%3A%5B%7B%22deviceType%22%3A%7B%22%24in%22%3A%5B%22Desktop%2Flaptop%22%5D%7D%7D%5D%7D%2C%22dateLabel%22%3A%22Trend%22%2C%22attributes%22%3A%22share%22%2C%22group%22%3A%22platform%22%2C%22sort%22%3A%7B%22share%22%3A-1%7D%2C%22id%22%3A%22platformsDesktop%22%2C%22dateInterval%22%3A%22Monthly%22%2C%22dateStart%22%3A%222018-05%22%2C%22dateEnd%22%3A%222019-04%22%2C%22segments%22%3A%22-1000%22%7D, consulté le 12 mai 2019

New York Times, (2019). *The New York Times*. En ligne https://www.nytimes.com/, consulté le 13 avril 2019

Privacy Tools, (2019). *Privacy Tools*. En ligne https://www.privacytools.io/, consulté le 3 février 2019

Randriavaniaina, A. F. (2011). *Tutoriel : Activer la fonction « Do Not Track » des navigateurs Web*. En ligne https://www.tomsguide.fr/tutoriel-activer-la-fonction-do-not-track-des-navigateurs-Web/, consulté le 8 mars 2019

Wikipédia, (2019). *Freeware*. En ligne https://fr.wikipedia.org/wiki/Freeware, consulté le 30 avril 2019

Wikipédia, (2019). *Google Chrome*. En ligne https://fr.wikipedia.org/wiki/Google_Chrome, consulté le 27 mars 2019

Wikipédia, (2019). *Microsoft Edge*. En ligne https://fr.wikipedia.org/wiki/Microsoft_Edge, consulté le 27 mars 2019

Wikipédia, (2019). *Mozilla Firefox*. En ligne https://fr.wikipedia.org/wiki/Mozilla_Firefox, consulté le 27 mars 2019

Wikipédia, (2019). *Safari (navigateur Web)*. En ligne https://fr.wikipedia.org/wiki/Safari_(navigateur_Web), consulté le 27 mars 2019

Youtube, (2019). *Youtube*. En ligne https://www.youtube.com, consulté le 15 avril 2019

Annexes

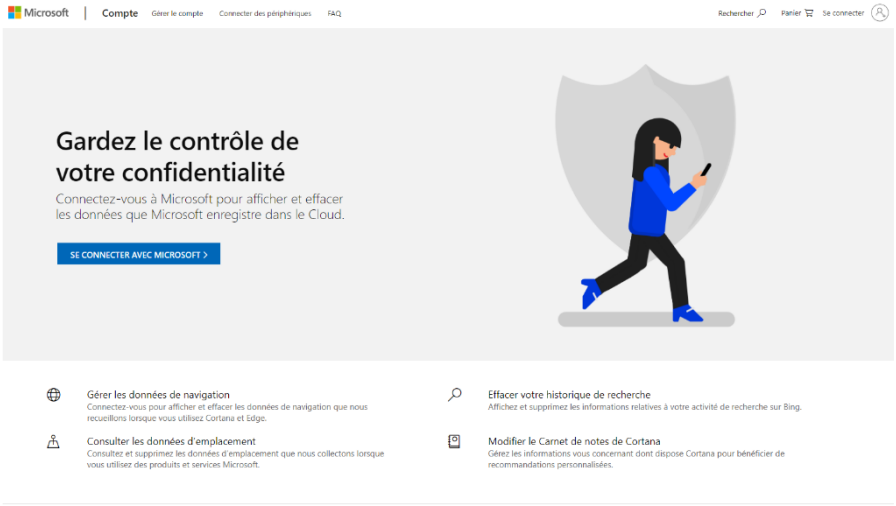
Annexe 1 : Système d’exploitation

Annexe 2 : Navigateurs

Annexe 3 : Applications

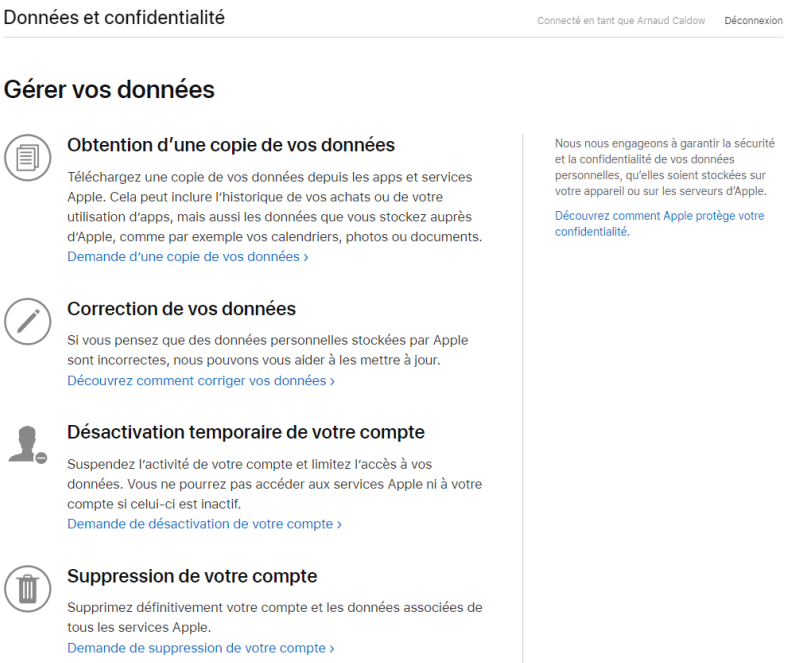
Annexe 1 : Système d'exploitation

Figure 20: Gestion des données Microsoft



Source : Microsoft, (2019). Garder le contrôle de votre confidentialité. En ligne https://account.microsoft.com/account/privacy?refd=privacy.microsoft.com&ru=https%3A%2F%2Faccount.microsoft.com%2Fprivacy%3Frefd%3Dprivacy.microsoft.com&destrt=privacy-dashboard, consulté le 5 avril 2019

Figure 21: Gestion des données Apple



Source : Apple, (2019). *Gérer votre confidentialité*. En ligne https://www.apple.com/befr/privacy/manage-your-privacy/, consulté le 5 avril 2019

Annexe 2 : navigateurs

Figure 22: Politique de confidentialité Firefox

Source : Mozilla, (2019). *Firefox – Politique de confidentialité*. En ligne https://www.mozilla.org/fr/privacy/firefox/, consulté le 7 avril 2019

Figure 23: Exemple de test Panopticlick

Test	Result
Is your browser blocking tracking ads?	✓ yes
Is your browser blocking invisible trackers?	✓ yes
Does your blocker stop trackers that are included in the so-called "acceptable ads" whitelist?	✓ yes
Does your browser unblock 3rd parties that promise to honor Do Not Track?	✗ no
Does your browser protect from fingerprinting?	✗ your browser has a unique fingerprint

Note: because tracking techniques are complex, subtle, and constantly evolving, Panopticlick does not measure all forms of tracking and protection.

Your browser fingerprint **appears to be unique** among the 200,505 tested in the past 45 days.

Currently, we estimate that your browser has a fingerprint that conveys **at least 17.61 bits of identifying information.**

Source : Electronic Frontier Foundation, (2019). *Is your browser safe against tracking ?*. En ligne https://panopticlick.eff.org, consulté le 20 mars 2019

Annexe 3 : Applications

Figure 24: Guide à l'utilisation Privacy Badger

Source : Electronic Frontier Foundation, (2019) *Privacy Badger*. En ligne moz-extension://7e372406-c8f9-465d-b467-ede4c31a6f62/skin/firstRun.html, consulté le 15 avril 2019