

100
QUESTIONS/RÉPONSES



Prix "Guide
pratique"
du Forum
International de la
Cybersécurité
(FIC) 2022

Laurane Raimondo

LA PROTECTION DES DONNÉES PERSONNELLES

ellipses

2^e édition

I. L'UNIVERS DES DONNÉES

Les données personnelles sont utilisées quotidiennement et leurs sources multiples. Désormais au centre des priorités stratégiques d'un grand nombre d'entreprises et de certains États, elles font l'objet de nombreux fantasmes, aussi vrais les uns que les autres. Mais d'où vient cette masse de données ? Car même sans étaler notre quotidien sur le premier réseau social venu, chacun de nous est un Petit Poucet en puissance.

Les données personnelles sont une ressource extraordinaire, non seulement nous les produisons nous-mêmes pour l'essentiel, mais en plus, elles sont infinies et s'accroissent de manière exponentielle. Une seule recherche sur Internet donne une indication sur votre emploi du temps, vos intérêts, etc. Il y a donc les activités sur le Web mais aussi tout ce que nous ne maîtrisons pas.

Songez à ce que les organismes de santé connaissent de vous depuis votre naissance : poids ; caractéristiques physiques ; risques de développement de certaines affections... Toutes ces données ont été créées par votre seule existence. Une partie de ces informations sont enregistrées sous un identifiant de grande valeur dont le traitement est d'ailleurs soumis à restriction : votre numéro d'inscription au répertoire (NIR), appelé aussi numéro de sécurité sociale. Ce numéro donne accès à des services et un grand nombre d'informations, il peut aisément être utilisé pour une usurpation d'identité.

Les autres administrations ne sont pas en reste, de votre état civil à la liste de tous les services auxquels vous répondez en termes de droits et devoirs, elles détiennent et traitent des données que vous produisez. Administrations et surtout, entreprises : chacune de vos actions, traçables d'une manière ou d'une autre, crée des données. Elles sont une ressource précieuse dont vous êtes la matière première : un paiement en carte bancaire ; un trajet suivi sur Waze® ou Google Map® ; un stationnement dans un parking muni d'un système de reconnaissance automatique des plaques d'immatriculation, etc.

Allez dans un centre commercial, aux Quatre Temps de la Défense par exemple, et votre parcours sera tracé point par point via la géolocalisation de votre smartphone, savoir combien de temps vous avez regardé tel écran ou telle devanture de boutique, a une réelle valeur. Ces données, lorsqu'elles sont corrélées avec d'autres informations vous concernant,

viennent compléter votre profil conservé et constamment enrichi par des entreprises de courtage de données afin d'atteindre l'objectif premier de cette collecte de masse : la prédiction de vos comportements, ou pire encore, leur influence.

Toute interaction avec un outil numérique connecté à Internet, comme une grande partie des activités quotidiennes, crée des données : regarder la météo sur l'appli de Météo France ; payer le péage par carte bancaire ; faire une demande d'extrait d'acte de naissance ; rechercher des informations sur l'éradication des punaises de lit ; prendre un rendez-vous médical sur doctolib.fr (victime d'ailleurs le 21 juillet 2020 d'un piratage et d'un vol de données personnelles de plus de 6 128 rendez-vous médicaux)... Ces données existent désormais, sont manipulées de part et d'autre après avoir été collectées, elles sont stockées, corrélées, revendues. S'il est difficile de s'y opposer concrètement, il est néanmoins possible de limiter ces traitements : ne pas donner ses informations à tort et à travers ; connaître et exercer ses droits.

Depuis quand parle-t-on de données personnelles et quand leur protection a-t-elle émergé ?

Les données personnelles existent depuis aussi longtemps que la nécessité d'identifier les personnes, mais les utiliser à cette fin ne signifiait pas qu'il leur était accordé une réelle importance à l'échelle individuelle. Plus concrètement, il semble que cet intérêt ait augmenté en corrélation avec l'évolution des services de renseignement et des besoins de contrôle des individus en France à la fin du XVIII^e siècle. En témoigne le livret ouvrier instauré en 1781 : surveiller les travailleurs au nom d'un impératif de sécurité publique. Olivier Tesquet le détaille dans son ouvrage *À la Trace*, sorti début 2020, « *Sur ce petit cahier de treize feuillets figurent non seulement le nom de son porteur, sa date de naissance, sa profession, mais aussi sa taille, la forme de son visage, la densité de ses sourcils ou d'éventuelles cicatrices* ». Un travailleur se déplaçant et qui n'est pas en mesure de produire ce carnet, conservé par l'employeur sauf permission, est arrêté. Identifier revient alors à contrôler des populations considérées comme potentiellement dangereuses, ou plutôt « à risque » de provoquer des mouvements sociaux.

Si les services de police ont eu besoin d'identifier et de classer des informations sur des personnes, hier comme aujourd'hui, la première nécessité impliquant des données personnelles fut d'ordre économique. Parmi les écrits les plus anciens retrouvés à ce jour, des tablettes sumériennes datées d'environ 3 500 ans avant notre ère indiquant l'identité d'une personne et la liste de ses biens. Toujours dans un objectif économique, le « recensement des paroisses et feux des Bailliages et Sénéchaussées de France », le premier en 1328, l'ancêtre de notre « foyer fiscal », permettant de collecter l'impôt. Ces feux représentaient une moyenne d'habitants mais pas leur nombre exact ni l'identité de chacun des membres du foyer. Un préliminaire au recensement, rapidement obsolète, la Grande Peste ayant ravagé le pays, et toute l'Europe au passage, entre 1346 et 1353.

Un autre moyen né d'un besoin d'identifier et de « tracer » les individus remonte au Moyen Âge, le registre paroissial. Le plus ancien conservé est celui de la commune de Givry en Saône-et-Loire, il détaille baptêmes, mariages et sépultures entre 1334 et 1357. Ces registres ô combien précieux pour les historiens mentionnaient les noms, prénoms, date et lieu de

naissance ainsi que l'identité des parents de l'enfant baptisé. Aujourd'hui l'état civil s'en charge et l'acte de naissance est l'un des documents nécessaires pour justifier de son identité et réaliser des démarches officielles.

Ces données sont précieuses pour un État dans l'exercice des droits régaliens. C'est à travers les dérives constatées par l'histoire que le sujet des données personnelles a réellement émergé. La police, selon qui la dirige, voit ses pouvoirs considérablement augmentés au fil du temps et ses objectifs varier selon le gouvernement en place. Un fichier de données sensibles comme la religion ou l'origine ethnique des individus a déjà conduit à de telles dérives dont les conséquences se payent encore à l'heure actuelle. L'électrochoc date du 21 mars 1974 lorsque *Le Monde* révèle au grand public le projet SAFARI (Système automatisé pour les fichiers administratifs et le répertoire des individus). Imaginez un fichier regroupant toutes les informations détenues par les différents services de l'administration française à votre sujet. Maintenant imaginez ce fichier entre les mains d'un État de droit transformé en État policier à la suite d'événements malheureux... La Sainte Inquisition aurait rêvé d'un tel fichier, la Stasi aussi, quoique d'après de récents témoignages, le système de surveillance actuel dépasse déjà les rêves les plus fous de ses anciens agents (*Nothing to Hide*, documentaire franco-allemand de Marc Meillassoux et Mihaela Gladovic, 2017).

La volonté de protéger les données en France est-elle seulement due à la révélation du projet SAFARI ? C'est en tout cas à la même époque qu'apparaissent en Europe les premières lois sur le sujet, à l'instar du Landtag de Hesse (parlement du Land de Hesse – alors situé en République fédérale d'Allemagne jusqu'en 1990) votant en 1970 une loi sur la protection des données, suivi par la Suède en 1973, la RFA en 1977, puis enfin la France, accompagnée de l'Autriche et du Danemark en 1978. Le Luxembourg se dote également d'une loi en 1979 avec pour objet, « *la protection des personnes physiques et morales contre une utilisation abusive de données nominatives* » dont l'évolution technologique a rendu l'application « *illusoire* » (Patrick Santer & Toinon Hoss, avocats à la Cour, Luxembourg).

Le besoin de protéger les personnes est notamment apparu avec la « *montée en puissance des banques de données et des ordinateurs* » (Sophie Kwasny, Conseil de l'Europe). Ces banques de données sont nées aux États-Unis, dans les années 1960, en même temps que le courtage

de données et les *data brokers* (cf. Question 92), fruit des progrès technologiques et d'une volonté gouvernementale d'efficacité dans les différents domaines stratégiques (défense, espace, énergies). Soutenues par l'État américain, des entreprises privées se sont développées autour de la mise à disposition de serveurs, leur praticité a poussé les États et organismes étrangers à les utiliser, confiant donc leurs données à ces serveurs.

Les législations, comme celles du Luxembourg, se sont rapidement révélées inefficaces : pour Santer et Hoss, « *la loi de 1979 soumettait toute banque de données, c'est-à-dire tout lieu où sont stockées des données nominatives, à une autorisation ministérielle préalable [...]. L'omniprésence et la démocratisation de l'outil informatique ont rendu impossible le respect de cette loi et il fallait être idéaliste ne serait-ce que pour tenter de s'y conformer.* » C'est lors de la publication de la directive 95/46/CE que certains États ont profité de l'occasion pour adapter leur législation, avec une problématique permanente : concilier les besoins d'une société de l'information avec les droits fondamentaux des personnes.

Fondamentalement, qu'est-ce qu'une donnée personnelle ? Un nom, une adresse, un numéro de téléphone... Oui mais pas seulement. Une première définition du Conseil de l'Europe la considère comme « toute information concernant une personne physique identifiée ou identifiable ». Le RGPD, comme le Conseil de l'Europe dans son rapport explicatif, précisent qu'une personne physique identifiable est « une personne physique qui peut être identifiée, directement ou indirectement, notamment par référence à un identifiant, tel qu'un nom, un numéro d'identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale ».

Une donnée à caractère personnel est donc toute information permettant de vous identifier personnellement à travers une ou plusieurs informations, des plus anodines aux plus précises : de votre rythme cardiaque à votre adresse IP en passant par vos déplacements les plus courants, dont quatre suffisent d'ailleurs à vous identifier. La majorité des personnes ne se doutent pas de la quantité de données personnelles qu'elles produisent au quotidien et laissent derrière elles à la manière du Petit Poucet. Elles se doutent encore moins que lesdites données sont captées par des moyens pratiquement invisibles pour les non-initiés, faisant la fortune de quelques milliers d'entreprises.

En effet, leur valeur est grande, il est couramment dit que « les données personnelles sont considérées comme l'or noir du ^{XXI}^e siècle », mais à tort : elles sont amplement plus précieuses que le pétrole. Alimentant des algorithmes sans cesse plus performants avec des objectifs plus ou moins légitimes, allant de la surveillance des personnes à l'évaluation de leur solvabilité jusqu'à l'influence de leur comportement sans qu'elles en prennent conscience, les données représentent un marché mondial estimé entre une et plusieurs centaines de milliards de dollars en 2023. Les chiffres en la matière restent à prendre avec du recul, il est extrêmement difficile de donner une fourchette précise au regard de l'opacité liée au secteur de la donnée. Le marché est cependant en constante augmentation avec un taux de croissance annuel composé avoisinant les 20 % entre 2020 et 2026 (ResearchAndMarkets : 17,3 %

entre 2020 et 2025 aux États-Unis; MarketsandMarkets: 18,6 % entre 2020 et 2026 en Europe)

Cet intérêt pour les données des personnes se disant ordinaires paraît démesuré et conduit ces dernières à penser que ce ne sont pas leurs propres informations qui comptent, qu'elles n'ont « rien à craindre » car « rien à cacher ». C'est justement parce que ce sont des personnes « ordinaires » que leurs données sont infiniment précieuses tant qu'elles sont collectées en quantité et mise en rapport avec d'autres pour obtenir des résultats permis par l'intelligence artificielle. Savoir que chacune de ses recherches sur Internet, chacun de ses déplacements, paiements, types d'achat, contacts, contenus de conversations, soucis de santé, salaires, types de véhicule, intérêts et chacune de ses photos, caractéristiques physiques, pensées, origine, etc., constituent le fonds de commerce d'entreprises de courtage de données ou de gouvernements est particulièrement inquiétant.

La technologie permet pour la première fois de l'histoire de récolter et traiter en masse des données; en ce sens et depuis les années 1970, diverses législations nationales et européennes viennent encadrer leur traitement hors du cadre privé et impliquant toutes les données à caractère personnel contenues dans des fichiers. À travers leur exploitation, les données sont une source d'information sur chacun d'entre nous en mesure de mettre en péril nos libertés fondamentales dont le droit à la vie privée, inscrit à l'article 8 de la Convention européenne des droits de l'Homme signée en 1950.

Qu'il existe des lois protégeant les données personnelles est insuffisant, une prise de conscience des personnes sur la valeur de leurs données et des dangers auxquels les traitements abusifs ou insuffisamment sécurisés les exposent est primordiale. Tout commence avec le réflexe le plus anodin qui soit: transmettre ses informations. Ne pas en dire plus que le nécessaire à qui vous le demande et savoir sur quoi ladite demande est fondée et si elle est légitime.

Plus insidieusement, faire une recherche sur Internet en dit beaucoup. Sans protection, de cette recherche sont extraites des informations comme la géolocalisation; l'heure de la recherche et son objet; le type de terminal utilisé, etc. Peuvent en être déduits directement ou par corrélation des données: l'état de santé; la situation familiale; le niveau de vie; la profession, etc. La liste n'est pas exhaustive et révèle dans