

Étude de cas A : L'utilisation des données personnelles par l'IA dans un but commercial

Document n°1 : extraits de l'article « Des données personnelles très convoitées »

Le business des données personnelles sur Internet est en plein boom, mais les GAFA (les quatre géants du Web : Google, Apple, Facebook, Amazon) trustent 95 % du marché. Est-il encore possible de réguler le secteur pour protéger la vie privée des citoyens ?

Tous surveillés. Les traces que l'on laisse derrière soi sur le Net – préférences idéologiques, culinaires ou sportives, achats, peines de cœur, soucis de santé... – sont devenues une industrie extrêmement lucrative. Selon le cabinet IDC, le marché des données des citoyens de l'Europe des Vingt-Huit s'élevait à 60 milliards d'euros en 2016, et devrait atteindre 80 milliards en 2020.

Et ce n'est que la partie émergée de l'iceberg. Indirectement, c'est-à-dire en tenant compte du chiffre d'affaires supplémentaire et des emplois générés par les utilisateurs de ces informations, la valeur de « l'économie européenne des données » s'élèverait à quelque 300 milliards d'euros en 2016, et pourrait atteindre 430 milliards en 2020.

Valérie Segond, *Le Monde*, 28 mai 2017.

Document n°2 : extraits de l'article « Web tracking, sur la piste de vos données personnelles ? »

Le web tracking est un moyen par lequel les sites identifient des utilisateurs et collectent leurs données de navigation. Les trackers sont les outils (ex : cookies) permettant de collecter de la donnée sur l'utilisateur et de suivre son parcours à travers les différents sites visités.

Qui se cache derrière le web tracking ?

Principalement, les publicitaires. Après avoir collecté ces données, des services (réseaux sociaux, web analytics, Ad Tech etc.) peuvent donc définir un identifiant utilisateur qui sera stocké dans un cookie. Cet identifiant permet de suivre l'utilisateur à travers différents sites web visités via un même navigateur.

L'identification d'un utilisateur est plus précise lorsqu'il s'est connecté à un espace loggé comme certains services le nécessitent tels que Facebook ou Twitter. Ainsi, la présence d'un bouton "like" de Facebook ou d'un bouton "tweeter" de Twitter sur un site web permet aux réseaux sociaux de savoir que l'utilisateur a visité cette page internet car un appel est fait vers leurs serveurs avec ces informations.

Pourquoi le web tracking ?

Le tracking sert à l'affichage ciblé des publicités dans le but de générer des revenus. Une publicité ciblée génère en général plus de clics lorsqu'elle repose sur les centres d'intérêts de l'utilisateur. L'exploitation des données utilisateurs permet donc de rémunérer les publishers.

Les outils de tracking peuvent permettre d'améliorer le confort de navigation de l'utilisateur en retenant par exemple les préférences de langue ou de personnalisation de l'interface. Ce tracking est aussi nécessaire à la recommandation de produits sur les sites e-commerce.

Que dit la loi pour protéger les internautes ?

La notion de consentement est essentielle pour garantir la protection des données personnelles des utilisateurs, déjà inscrite dans la loi Informatique et Liberté en France puis consolidée par le RGPD (*Règlement Général sur la Protection des Données*).

Lola Met, *DPO consulting*, 30/08/2019.

Document n°3 : extraits de l'article « Consommateurs : attention à vos données personnelles ! »

En 2016, 7 703 plaintes ont été déposées à la Commission nationale de l'informatique et des libertés (Cnil) par des personnes qui, malgré leurs demandes, n'étaient pas arrivées à connaître, faire rectifier ou supprimer des données les concernant.

Les entreprises ont l'obligation d'effectuer le recueil de données personnelles de façon transparente et avec l'accord des consommateurs. « Lorsqu'une personne souscrit une carte de fidélité ou s'inscrit sur un compte Internet pour un achat, elle doit remplir un formulaire ». « Elle va ensuite avoir la possibilité de cocher une case pour, selon les cas, consentir ou s'opposer à la prospection commerciale et à la transmission de ses données à des partenaires. Il faut y être extrêmement attentif ».

Lors d'un achat ou d'une souscription à un abonnement, les données obligatoires sont notifiées d'un astérisque. En son absence, « les consommateurs doivent se demander si la demande d'information est pertinente dans leur relation client et si elle a un lien avec l'acte d'achat ». Les sociétés peuvent être tentées de collecter de très nombreuses données (catégorie socio-professionnelle, hobbies...) sans rapport direct avec leur activité, ce qui est contraire à la loi. « On ne demande pas un numéro de sécurité sociale si cela n'est pas pertinent », détaille l'experte de la Cnil. « Les données sensibles relatives à la sexualité, la religion, ou encore l'origine ethnique ou raciale, ne doivent pas être collectées, sauf consentement express de la personne et sous réserve que cela soit pertinent ».

Cécile Simon, *Le Parisien*, 20/10/2017.

Étude de cas B : L'utilisation des données personnelles par l'IA dans le monde professionnel

Document n°1 : extraits « L'industriel français Schneider Electric déploie depuis 2020 une plate-forme basée sur l'IA pour le suivi de l'évolution de carrières. Au programme, évolution par la formation, mais aussi mise en relation de compétences avec des projets ».

L'énergéticien français Schneider Electric a décidé de s'appuyer sur l'intelligence artificielle pour gérer l'évolution de carrière de ses employés. Pour ce faire, l'entreprise a lancé sa plateforme Open Talent Market (OTM) basée sur l'IA début 2020 pour faciliter l'identification de mentors, la sélection de programmes de formation continue et l'élargissement du réseau des employés au sein de l'entreprise dans le monde entier.

Des données absorbées sur LinkedIn

En pratique, soit les employés créent eux-mêmes leur profil, soit la plateforme se nourrit directement des informations LinkedIn, y compris les compétences actuelles et l'expérience. Le dispositif se distingue parce qu'il aide les employés à prendre leur carrière en main de façon autonome, sans être désavantagés en fonction du sexe, de l'âge ou de l'origine ethnique.

Découvrir de nouveaux domaines

Shannon Booth confirme en effet que les employés utilisent la plate-forme pour échanger et se connecter directement les uns avec les autres. Les réseaux de mise en relation locaux intègrent aussi des dispositifs de mentorat qui peuvent passer par la plate-forme. L'IA met en correspondance et en relation (matchmaking) des mentors et des mentorés potentiels en fonction de leurs intérêts et objectifs respectifs.

Sarah K. White, IDG NS (adapté par Dominique Filippone), *Le monde informatique*.

Document n°2 : extraits de l'article « Recruter avec l'IA : qu'en pensent les candidats ? »

Selon une étude parue en août 2023, environ deux tiers des entreprises utilisent d'ores et déjà l'intelligence artificielle dans le cadre de leur process de recrutement. Que ce soit pour sourcer des candidats, trier des CV, rédiger des offres d'emploi ou préparer des trames d'entretien d'embauche.

Des décisions plus objectives ?

L'expérience a montré que les algorithmes (ChatGPT ou Google Bard) pouvaient discriminer certains candidats. Amazon, qui avait été l'une des premières entreprises à recruter avec l'IA en 2018, avait dû faire marche arrière, après la découverte de biais écartant certains groupes de candidats du processus d'embauche.

Même si la plupart des technologies travaillent à réduire ces biais, la possibilité d'une discrimination à l'embauche existe toujours, car l'algorithme s'entraîne sur des données qui peuvent, elles-mêmes contenir des biais. Or ce sont des humains qui conçoivent les algorithmes et qui interprètent les résultats fournis par l'IA.

Une image d'entreprise innovante ?

Les travaux des deux chercheurs, ont mis en évidence que plus les candidats considèrent l'intelligence artificielle comme éthique et efficace, plus leur confiance envers les employeurs qui l'utilisent pour leurs recrutements sera élevée. Ces candidats considérant l'IA comme éthique seront plus enclins à postuler dans les entreprises ayant recours à cette technologie, car ils les jugent innovantes et attractives. À l'inverse, ceux qui mettent en doute l'éthique des décisions prises par l'algorithme seront moins tentés de candidater.

Une menace pour la vie privée ?

Les chercheurs montrent également que la perception des candidats dépend aussi de la finalité de l'utilisation de l'IA dans le cadre du recrutement. Si son usage pour établir l'adéquation d'un CV avec le poste à pourvoir est globalement bien accepté par les candidats, son utilisation pour aller traquer leurs profils sur les réseaux sociaux ou analyser leurs comportements lors d'un entretien vidéo suscite davantage de défiance. Les candidats redoutent, en effet, l'exploitation de leurs données personnelles par l'IA, et ont tendance à se détourner d'entreprises ayant ce type de pratiques.

L'humain doit rester au cœur du processus

En conclusion, les chercheurs constatent que « les candidats préfèrent encore être soumis à une décision humaine, ou à une collaboration entre l'homme et l'IA, plutôt que de voir leur sort influencé par une IA autonome ».

Maïté Helligo, *Hello workplace*, 15/11/23

Étude de cas C : L'utilisation des données par l'IA à des fins politiques

Document n°1 : extraits de l'article « Ce qu'il faut savoir sur Cambridge Analytica, la société au cœur du scandale Facebook »

« Cette entreprise britannique, utilisée par Donald Trump durant sa campagne pour la présidentielle de 2016 aux Etats-Unis, a analysé les données de dizaines de millions d'utilisateurs à leur insu.

Elle devait changer les règles du jeu politique ; voilà qu'elle est au centre d'une affaire de vol de données révélée par le Guardian, le New York Times et The Observer, qui éclabousse autant la Maison Blanche que Facebook, le plus grand réseau social du monde.

Pourquoi est-elle dans la tourmente ?

Cambridge Analytica est accusée d'avoir utilisé des données de 30 millions à 70 millions d'utilisateurs de Facebook, recueillies sans leur consentement, en passant par un quiz développé par un universitaire anglais, Aleksandr Kogan, et sa société Global Science Research (GSR).

Samedi 17 mars, le Guardian, The Observer et le New York Times ont révélé que les données récoltées par GSR pour le compte de Cambridge Analytica l'ont été à l'insu des internautes concernés, en présentant le quiz comme un simple exercice académique, alors que celui-ci absorbait les données non seulement des participants, mais aussi de leurs « amis » Facebook. »

William Andureau, *Le Monde*, 22/03/2018

Document n°2. Extraits de l'article : « Mais que s'est-il (véritablement) passé avec Cambridge Analytica ? »

« Fin août, Meta, la société mère de Facebook, a trouvé un accord avec les plaignants qui avaient engagé des poursuites judiciaires contre l'entreprise, suite à l'affaire Cambridge Analytica. Mais le vrai scandale n'est pas celui que l'on croit.

En 2016, l'entreprise britannique avait utilisé des données, siphonnées sur Facebook, pour aider la campagne présidentielle de Donald Trump à mieux cibler ses messages aux électeurs et influencer les élections américaines. L'affaire est révélée le 16 mars 2018. À l'origine, les révélations du lanceur d'alerte Christopher Wylie, 28 ans, un jeune analyste canadien qui a travaillé pour Cambridge Analytica de 2013 à 2014. "Nous nous sommes servis de Facebook pour collecter des millions de profils d'utilisateurs et bâtir des modèles ciblant leurs préférences et leurs démons intimes" révèle-t-il au journal The Guardian.

Dans les coulisses de l'affaire

Le vrai scandale de l'affaire Cambridge Analytica, c'est la manière dont les données de 87 millions d'utilisatrices et utilisateurs de Facebook ont été utilisées par la firme britannique. Elle a réussi à créer un redoutable système capable de cibler et d'orienter les votants, grâce à des messages cousus main.

L'arme fatale

Ce qui intéresse Cambridge Analytica ce n'est pas l'âge, le sexe ou la couleur politique des utilisatrices et utilisateurs de Facebook, mais plutôt leur caractère. Introvertis? Extravertis? Avoir cette information permettrait de micro-cibler le message électoral adéquat. C'est la personnalité qui influence le comportement des gens et les pousse à voter d'une certaine manière. Et pour la saisir, Cambridge Analytica utilise des tests de personnalité qui rencontrent beaucoup de succès sur le réseau social. Ces données, la firme les croise avec les "like" et les centres d'intérêt. L'entreprise travaille dur pour développer des douzaines de variations publicitaires sur différents thèmes politiques tels que l'immigration, l'économie et les droits des armes à feu, tous adaptés à différents profils de personnalité.

Miruna Coca-Cozma, *TRS*, 2/10/2022