

Peut-on vraiment lier l'éthique et le big data ?

"Un grand pouvoir implique de grandes responsabilités"

Les amateurs de comics reconnaîtront rapidement les mots qui ont encouragé Peter Parker à devenir Spider-Man, les puristes, eux, diront que c'est Voltaire qui les a prononcé en premier. Néanmoins, aussi importante que soit cette citation pour l'Histoire ou pour l'avenir de l'homme araignée, elle est encore plus pertinente aux yeux des spécialistes des technologies dans un monde centré sur les données.

La quantité de données personnelles recueillies par les entreprises aujourd'hui est en effet considérable. En termes d'abus de données personnelles, le récent scandale de Facebook et Cambridge Analytica en est le parfait exemple. C'est pour ces raisons que le concept d'éthique des données prend autant d'importance.

Qu'est-ce que l'éthique des données ?

La frontière entre utilisation appropriée et abusive des données est mince. Au fur et à mesure que la science des données et leurs technologies évoluent, « l'art du possible » change également avec elles. Même si l'analyse des données n'est pas nouvelle en soi, il est maintenant possible de traiter rapidement de grandes quantités de données et d'établir des corrélations et des prédictions à partir de sources variées. La facilité de ces démarches soulève en revanche de nombreuses questions liées à la protection de la vie privée, la confidentialité, la transparence et l'identité.

Le concept d'éthique des données se concentre sur les problèmes moraux qui y sont associés :

La manière dont les données sont générées, enregistrées et partagées

La façon dont les algorithmes de machine learning et d'intelligence artificielle utilisent les données

Les pratiques en matière de données qui sont mis en place dans les secteurs public et privé

L'éthique des données souligne la complexité des défis posés par la science des données et l'analyse du big data. Gartner avait prédit que la moitié des abus en matière d'éthique commerciale résulteraient de l'exploitation inappropriée du big data. En bref, nos cadres éthiques actuels ne sont plus adaptés à l'usage des données et nous devons maintenant penser différemment.

La protection de la vie privée en pratique

Bien que presque toutes les entreprises appliquent une politique de protection de la vie privée, cela ne signifie pas pour autant qu'elles font preuve d'éthique en matière de données. Avez-vous déjà réellement lu les règles de confidentialité lorsque vous visitez un site web ? Seulement 7% des interrogés français le font vraiment et le nombre réel est probablement plus faible. De plus, une politique de la vie privée ne garantit pas la confidentialité absolue de vos données : elle s'apparente simplement à un document juridique identifiant les utilisations potentielles de vos données et l'on peut facilement être d'accord avec toutes les clauses

Bien qu'ils soient généralement lents à réagir, plusieurs gouvernements ont récemment adopté des lois pour protéger les consommateurs. Le Règlement général sur la protection des données (GDPR) de l'Union Européenne a été qualifié de « vie privée par défaut », donnant aux citoyens un contrôle strict de leurs données. Plus tôt en 2018, la Californie a adopté la California Consumer Privacy Act (CCPA) pour protéger la vie privée en ligne et les données personnelles Identifiables (DPI). Aujourd'hui, la Federal Data Strategy relative à l'accès aux informations vise à faire de la gouvernance éthique l'un de ses principes fondamentaux. Les grandes entreprises se rejoignent ainsi pour défendre la protection de la vie privée dans l'espoir de façonner la législation future aux États-Unis.

La protection de la vie privée est en train de faire son retour, mais nous avons encore besoin de l'éthique des données pour nous guider vers le privacy by design.

Le défi français des données

Le 4 mars dernier, Audrey Azoulay, directrice générale de l'Unesco, a ouvert la première conférence mondiale pour la promotion d'une approche humaniste sur l'intelligence artificielle (IA). Une feuille de route qui inclut aussi le thème de la gestion éthique des données. Contrairement aux États-Unis et à la Chine, l'Union Européenne et la France rencontrent encore des difficultés à mettre en place une réelle doctrine. La France a pourtant des valeurs fortes inspirées de la philosophie des Lumières. L'éthique des données doit donc entrer en corrélation profonde avec les aspirations des Français si l'on souhaite améliorer leur quotidien mais aussi le maîtriser et préparer l'avenir.

De nombreuses initiatives sont en cours dans l'Hexagone, principalement dans le cadre du Grand Débat National, initié par Emmanuel Macron suite aux revendications des Gilets Jaunes. Actuellement, les doléances des Français sont collectées dans les communes et les villes sur tout le territoire et beaucoup portent sur l'amélioration des services publiques. Des projets voient le jour comme le Health Data Hub (pilote par Jean-Marc Aubert, directeur de la recherche, des études, de l'évaluation et des statistiques (Drees), qui permet de soutenir les start-up dans la collecte et la consolidation du patrimoine de données de santé et assurer un accès simplifié, effectif et accéléré aux données, tout en tissant un lien avec les citoyens et la société civile. Mais la France doit surtout adopter un positionnement et un discours forts si elle souhaite s'imposer à l'échelle européenne et internationale.

Malgré tout, l'éthique des données ne freinera pas l'innovation. Cette dernière est un savant mélange entre collaboration et empathie. Dans le cas des données, l'empathie est née, elle, de la réflexion sur la façon dont nous utilisons les données et leur impact sur la société. Associée à une discussion ouverte et collaborative avec les citoyens et les entreprises sur l'éthique des données, elle conduira à l'innovation. De plus, dans notre société actuelle, l'éthique peut constituer un avantage concurrentiel, au même titre que les technologies responsables pour les entreprises soucieuses de l'environnement.

Il y a une méfiance historique à l'égard des institutions, surtout celles qui abusent du pouvoir de l'information. Quoi qu'il en soit, nous avons tous une grande responsabilité en matière d'éthique des données car c'est l'avenir de notre économie numérique qui est en jeu.