

L'INTELLIGENCE ARTIFICIELLE ET LES DONNÉES À CARACTÈRE PERSONNEL

LE BIG DATA

Le big data ou « données massives » désigne l'ensemble des données numériques. Une donnée numérique est une information qui a du sens, une signification (elle peut désigner un lieu, une date, un site internet visité, un nom...). Elles peuvent être de nature personnelle, professionnelle, institutionnelle, etc. Ces données sont produites par l'utilisation de tous types d'outils numériques sans que l'utilisateur n'en soit toujours conscient. L'intelligence artificielle a besoin d'une importante quantité de données à analyser pour être fonctionnelle et produire des résultats, le big data est donc une ressource essentielle pour l'IA.

L'analyse de toutes ces données est difficile à produire par l'humain seul, ainsi que par des logiciels de gestion des données classiques. À l'inverse, l'intelligence artificielle est en mesure de produire une analyse, de catégoriser, de réaliser des prédictions, d'apprendre et d'adapter ses réponses. Cela apporte de nombreuses perspectives pour l'activité humaine assistée par l'IA (recherche, santé, commerce, économie, écologie...).

L'exploitation de l'ensemble de ces données se structure autour de différentes chartes, déclarations, serments, lois afin de protéger l'individu.

LES DONNÉES À CARACTÈRE PERSONNEL

La donnée à caractère personnel se définit différemment d'un pays à un autre. Aux États-Unis, elle est caractérisée par les données d'identification qui sont utilisées pour accéder à un compte numérique alors qu'en Europe les données personnelles se définissent comme toutes les données « en relation avec des personnes physiques identifiables de façon directe ou indirecte » [déclaration de Montréal, 2018].

En mai 2018, l'Union européenne a instauré un règlement européen sur la protection des données (RGPD). Depuis le 25 mai 2018, ce règlement pose le cadre légal sur l'utilisation et le traitement des données personnelles recueillies par les institutions qui résident ou qui ont des utilisateurs au sein de l'Union européenne. Il constitue un pas vers la transparence de l'utilisation des datas. Les organisations sont dans

l'obligation d'indiquer quelles sont les données collectées sur un individu et de lui demander son accord (droit à l'accès) ; de supprimer toutes les données qui ne sont pas tenues de respecter d'autres obligations lorsque cela est demandé (droit à l'oubli) ; de fournir une explication du traitement effectué sur les données du client (droit à l'explication). Ce règlement a été conçu, entre autres, pour protéger la vie privée des individus. Les préoccupations sur les impacts sociétaux et individuels ne sont pas uniquement européennes mais bien des préoccupations mondiales.

LA PROTECTION DE LA VIE PRIVÉE

Pour structurer le développement de l'intelligence artificielle au niveau international, le serment Holberton-Turing, qui peut être prêté par tous les professionnels de l'IA, a pour volonté de fédérer ces derniers au niveau mondial autour de valeurs éthiques et morales communes. Il existe également la déclaration de Montréal, pensée dans le but de poursuivre l'avancée technologique tout en présentant certaines vigilances. Le but est d'inciter l'individu travaillant ou utilisant l'intelligence artificielle à être vigilant, dans sa pratique, aux valeurs éthiques, morales et aux problématiques que l'IA serait susceptible de créer. Le respect de la vie privée est un point mis en lumière, tout comme l'importance d'une collaboration entre les humains et les machines, le respect de l'équité, de l'égalité, le maintien de l'autonomie de tous, la liberté de penser, de choisir pour ainsi anticiper de potentielles conséquences négatives.

Les points sur les données personnelles dans le **serment Holberton-Turing** donnent les précisions suivantes.

- « Je me souviendrai que je ne manipule pas que des données, des 0 et des 1, mais des informations en relation avec des êtres humains dont les interactions avec mon logiciel d'intelligence artificielle peuvent affecter la liberté, la famille ainsi que la stabilité économique. »
- « Je respecterai les secrets qui me sont confiés. »

Les données personnelles et le respect de la vie privée sont également évoqués dans la **déclaration de Montréal**.

- « Le développement des systèmes d'IA doit prévenir les risques d'une utilisation néfaste des données d'utilisateurs et protéger l'intégrité et la confidentialité des données personnelles. »

- « Des espaces d'intimité dans lesquels les personnes ne sont pas soumises à une surveillance ou à une évaluation numérique doivent être protégés de l'intrusion de système d'IA ou de systèmes d'acquisition et d'archivage des données personnelles (SAAD). »

D'autres documents participent au cadrage du développement de l'intelligence artificielle. Laurence Devilliers, par exemple, propose d'étoffer les **trois lois de la robotique d'Asimov** [destinées à donner un cadre aux comportements et capacités de tout robot] pour les adapter aux robots sociaux et d'y intégrer des commandements en faveur du respect de la vie privée.

Toutes ces préoccupations se jouent également dans le système éducatif.

DANS LE SYSTÈME ÉDUCATIF

Les établissements scolaires sont aussi concernés par la protection des données puisque l'utilisation d'outils numériques dans le contexte scolaire est une pratique accessible aujourd'hui.

Les données personnelles, c'est-à-dire les informations qui peuvent identifier un élève, concernent le nom et prénom, les adresses mail et postale, les résultats scolaires ou encore les évolutions d'apprentissage. Les données personnelles, avec l'arrivée du numérique, peuvent être facilement distribuées et diffusées via internet. Dans le cadre scolaire, les établissements peuvent être amenés à travailler avec des organismes ou des professionnels qui sont en mesure de collecter les données telles que les entreprises proposant des plateformes d'apprentissage en ligne ou des logiciels. Ces supports numériques sont utilisés par les élèves dans le cadre de leur vie scolaire, mais également par les professionnels de l'éducation pour la gestion administrative par exemple.

Le rapport « Données numériques à caractère personnel au sein de l'Éducation nationale », rédigé par l'inspection générale de l'Éducation nationale et l'inspection de l'administration de l'Éducation nationale et de la recherche (février 2018), indique que les élèves ont peu de connaissances sur le traitement ou le devenir de leurs données personnelles. L'utilisation de ces données, leur recueil, leur analyse, leur stockage sont des processus peu visibles, rendant difficile la prise de conscience et la compréhension de ces processus immatériels. Les élèves sont exposés aux risques d'utilisations abusives de leurs données. Pour faire face à ces risques de diffusion de ces données, il convient de favoriser l'accès à l'information aux enseignants et aux chefs d'établissements, pour ensuite les transmettre aux élèves et les sensibiliser aux dimensions éthiques, sociales et économiques de l'utilisation et du recueil de leurs données numériques. Cela signifie donner à tous une culture générale et un esprit critique sur l'intelligence artificielle et l'algorithmique [Villani, 2018].

En France, la CNIL¹, pour répondre aux enjeux de transparence et de respect de la vie privée dans le système éducatif, a signé, en décembre 2018, une convention sur la protection des données personnelles dans les usages numériques de l'Éducation nationale avec le ministère de l'Éducation nationale de la Jeunesse. Les axes s'articulent vers la formation ou la création de nouvelles ressources pour accompagner la communauté éducative dans la maîtrise et la compréhension de l'écosystème qui gravite autour d'internet et des données personnelles. Cette convention a aussi pour but de valoriser les données personnelles pour, par exemple, accompagner des projets innovants dans un cadre sécurisé et respectueux. Ces données peuvent fournir de nombreuses informations pour faire évoluer les méthodes d'apprentissage. Leur utilisation, leur récolte ou leur analyse ne sont pas prohibées mais contraintes à la transparence.

RÉFÉRENCES

Unesco, *Conférence internationale sur l'intelligence artificielle et l'éducation*. Planifier l'éducation à l'ère de l'IA : Un bond en avant, 2019 [en ligne].

Villani Cédric, « *Transformer l'éducation* », in Donner un sens à l'intelligence artificielle. Pour une stratégie nationale et européenne, 2018.

IGEN, IGAENR, *Données numériques à caractère personnel au sein de l'Éducation nationale*, 2018 [en ligne].

1. La Commission nationale de l'informatique et des libertés (CNIL) est un organisme public qui veille en France à la protection des données personnelles contenues dans les documents informatiques ou papier qui évoluent au rythme de l'apparition de nouveaux usages.