

Comment permettre à l'Homme de garder la main ? Rapport sur les enjeux éthiques des algorithmes et de l'intelligence artificielle

15 décembre 2017

La CNIL publie le rapport de synthèse du débat public qu'elle a animé sur les enjeux éthiques des algorithmes et de l'intelligence artificielle. Isabelle FALQUE-PIERROTIN, présidente de la CNIL, le présente aujourd'hui en présence de Mounir MAHJOUBI, Secrétaire d'Etat chargé du Numérique et de Cédric VILLANI, Député, chargé par le gouvernement d'une mission sur l'intelligence artificielle.

L'irruption dans nos vies quotidiennes des algorithmes et de l'intelligence artificielle fait l'objet d'une attention publique soutenue, depuis que quelques cas emblématiques (APB, rôle des réseaux sociaux dans les présidentielles américaines etc.) ont marqué les citoyens, ici en France ou à l'étranger. A l'heure où se construisent les positions française et européenne sur l'intelligence artificielle, ce rapport a pour but de construire un panorama des enjeux éthiques qu'ils soulèvent.

Un débat public innovant sur les algorithmes et l'intelligence artificielle

Dans notre monde de plus en plus numérique, les algorithmes et l'intelligence artificielle sont partout : pour simuler l'évolution de la propagation de la grippe en hiver, pour recommander des livres à des clients, pour suggérer aux forces de police des zones où patrouiller en priorité, pour piloter de façon autonome des automobiles, pour élaborer automatiquement un diagnostic médical personnalisé, pour personnaliser un fil d'activité sur les réseaux sociaux etc. Des tâches complexes, parfois critiques, sont ainsi déléguées à des systèmes de plus en plus autonomes à mesure que les techniques d'apprentissage propres à l'intelligence artificielle se développent. Ces usages multiples et croissants, indissociables des masses de données inédites à disposition dans tous les secteurs, soulèvent de nombreuses questions.

Chargée par loi pour une République numérique de mener une réflexion sur les questions éthiques et de société posées par les nouvelles technologies, la CNIL s'est naturellement tournée vers le thème des **algorithmes à l'heure de l'intelligence artificielle**.

Le rapport s'appuie sur les résultats d'un débat public ouvert et décentralisé que la CNIL a animé de janvier à octobre 2017 : 3 000 personnes ont participé à [45 manifestations](#) qui se sont tenues à l'initiative de [60 partenaires](#) en France (Ax les Termes, Bordeaux, Caen, Lille, Lyon, Marseille, Paris, Toulouse) et à l'étranger (Etats-Unis). [Une concertation citoyenne](#) a également été organisée à Montpellier le 14 octobre pour compléter la démarche.

Médias, vie publique et politique, éducation, culture, santé, justice, sécurité et défense, banque et finance, emploi et recrutement ont été abordés comme champs d'exploration des algorithmes et de l'intelligence artificielle par les partenaires : des instituts de recherches (CREOGN, CNAM, ENSC, INSA), des institutions publiques (ministères, universités et grandes écoles), des fédérations professionnelles, des syndicats, des ordres d'avocat, des hôpitaux, des think tanks, des entreprises, etc.

Pour Isabelle Falque-Pierrotin, Présidente de la CNIL :

L'objectif de ce débat est de garantir que l'intelligence artificielle augmente l'homme plutôt qu'elle ne le supplante et participe à l'élaboration d'un modèle français de gouvernance éthique de l'intelligence artificielle. Nous devons collectivement faire en sorte que ces nouveaux outils soient à la main humaine, à son service, dans un rapport de transparence et de responsabilité

Quelles réponses éthiques au développement des algorithmes et de l'intelligence artificielle ?

Vers l'affirmation de deux principes fondateurs : loyauté et vigilance

Les débats ont permis de dégager deux principes fondateurs pour une intelligence artificielle au service de l'homme. Ces principes pourraient s'inscrire dans **une nouvelle génération de garanties et de droits fondamentaux à l'ère numérique**, des « droits-système » organisant la **gouvernance mondiale de notre univers numérique** :

- [Un principe de loyauté](#) appliqué à tous les algorithmes et intégrant les impacts collectifs, et pas seulement personnels, de ces derniers. Tout algorithme, qu'il traite ou non des données personnelles, doit être loyal envers ses utilisateurs, non pas seulement en tant que consommateurs, mais également en tant que citoyens, voire envers des communautés ou de grands intérêts collectifs dont l'existence pourrait être directement affectée. **L'intérêt des utilisateurs doit primer**. Par exemple, un tel principe pourrait avoir vocation à s'appliquer à l'impact potentiel des réseaux sociaux sur la structure du débat public dans nos démocraties (segmentation du corps politique par le ciblage de l'information) ou à celui d'algorithmes de police prédictive sur des communautés ou quartiers entiers.
- [Un principe de vigilance/réflexivité](#) : il s'agit d'organiser une forme de questionnement régulier, méthodique et délibératif à l'égard de ces objets mouvants. Ce principe constitue une réponse directe aux exigences qu'imposent ces objets technologiques du fait de leur nature imprévisible

(inhérente au *machine learning*), du caractère très compartimenté des chaînes algorithmiques au sein desquels ils s'insèrent et, enfin, de la confiance excessive à laquelle ils donnent souvent lieu. C'est l'ensemble des maillons de la chaîne algorithmique (concepteurs, entreprises, citoyens) qui doivent être mobilisés pour donner corps à ce principe, au moyen de procédures concrètes (par exemple, des comités d'éthique assurant un dialogue systématique et continu entre les différentes parties-prenantes).

Ces principes fondateurs sont complétés par des principes organisationnels ayant trait à l'intelligibilité et à la responsabilité des systèmes algorithmiques ainsi qu'à la nature de l'intervention humaine dans la prise de décision algorithmique.

Les 6 recommandations opérationnelles

Ces principes font l'objet d'une déclinaison sous la forme de **6 recommandations opérationnelles** à destination tant des pouvoirs publics que des diverses composantes de la société civile (entreprises, grand public, etc.) :

1. [Former à l'éthique tous les acteurs-maillons de la « chaîne algorithmique »](#) (concepteurs, professionnels, citoyens) : l'alphabétisation au numérique doit permettre à chaque humain de comprendre les ressorts de la machine ;
2. [Rendre les systèmes algorithmiques compréhensibles](#) en renforçant les droits existants et en organisant la médiation avec les utilisateurs ;
3. [Travailler le design des systèmes algorithmiques au service de la liberté humaine](#), pour contrer l'effet « boîtes noires » ;
4. [Constituer une plateforme nationale d'audit des algorithmes](#) ;
5. [Encourager la recherche sur l'IA éthique et lancer une grande cause nationale participative autour d'un projet de recherche d'intérêt général](#) ;
6. [Renforcer la fonction éthique au sein des entreprises](#) (par exemple, l'élaboration de comités d'éthique, la diffusion de bonnes pratiques sectorielles ou la révision de chartes de déontologie peuvent être envisagées).

Les problématiques soulevées par les algorithmes et l'IA

Pour formuler ces principes fondateurs et les recommandations, le rapport s'est basé sur 6 problématiques essentielles remontées lors des débats afin de relever le défi éthique posé par ces technologies :

1. L'autonomie humaine au défi de l'autonomie des machines :

Algorithmes et intelligence artificielle permettent la délégation croissante de tâches, de raisonnements ou de décisions de plus en plus critiques à des machines. Souvent jugées infaillibles et « neutres », celles-ci n'ouvrent-elles pas la voie à une confiance excessive et à la tentation pour chacun de ne pas exercer pleinement ses responsabilités ? Comment faire face aux formes nouvelles de dilution de la responsabilité qu'impliquent des systèmes algorithmiques complexes et très segmentés ?

2. Biais, discrimination et exclusion :

Biais, discrimination et exclusion constituent des effets clairement identifiés des algorithmes et de l'intelligence artificielle. Parfois volontaires, ces phénomènes sont, à l'heure du déploiement des algorithmes de *machine learning*, le plus souvent inconscients et difficilement repérables. Comment appréhender cette nouvelle problématique ?

3. Fragmentation algorithmique : la personnalisation contre les logiques collectives :

La personnalisation constitue l'une des grandes promesses de l'algorithme tel qu'il s'est construit avec le Web. Bien que des bénéfices certains pour l'individu émanent de la segmentation et du profilage, comment peuvent-ils tout autant affecter vigoureusement des logiques collectives essentielles à la vie de nos sociétés (pluralisme démocratique et culturel, mutualisation du risque) ?

4. Entre limitation des mégafichiers et développement de l'intelligence artificielle : un équilibre à réinventer :

L'intelligence artificielle nécessite d'importantes quantités de données à des fins d'apprentissage. Elle suscite également de nombreuses promesses. Comment exploiter celles-ci tout en maintenant l'impératif de protéger les libertés individuelles de chacun, intrinsèque à la législation européenne de protection des données personnelles ?

5. Qualité, quantité, pertinence : l'enjeu des données fournies à l'IA :

L'entraînement d'un algorithme d'apprentissage implique de l'alimenter avec des données sélectionnées avec soin, pertinentes au regard de l'objectif poursuivi et en quantité suffisante. Comment parvenir à une telle attitude critique en dépit de la tendance à une confiance excessive dans la machine ?

6. L'identité humaine au défi de l'intelligence artificielle :

L'idée d'une spécificité humaine irréductible se voit questionnée par l'autonomie grandissante des machines ainsi que par l'apparition de formes d'hybridation entre humains et machines. Faut-il et est-il possible de parler au sens propre d'« éthique des algorithmes » ? Comment appréhender cette nouvelle classe d'objets que sont les robots humanoïdes, susceptibles d'engendrer chez l'homme des formes d'affect ?

Document reference

Pour approfondir

[Comment permettre à l'Homme de garder la main ? Rapport sur les enjeux éthiques des algorithmes et de l'intelligence artificielle](#)

[PDF-1.5 Mo]

[COMPTE RENDU - Concertation citoyenne sur les enjeux éthiques liés à la place des algorithmes dans notre vie quotidienne](#)

[PDF-1.88 Mo]