

Difficile aujourd'hui d'échapper aux algorithmes et à l'**intelligence artificielle** (IA). De plus en plus, ceux-ci modèlent la vie des individus, reconstruisent les sociétés, affectant jusqu'aux formes du droit et le cours de la justice.

**Leurs « prophéties autoréalisatrices » reconstruisent le monde ; et cela de manière non désintéressée.** C'est pourquoi ils comptent, depuis peu mais de manière croissante, au nombre des principales préoccupations des autorités publiques. En France, la Cnil est très naturellement en première ligne face à ces enjeux. « *Comment permettre à l'Homme de garder la main ?* » : telle est la question à laquelle l'autorité indépendante tente de répondre dans son récent rapport sur les enjeux éthiques des algorithmes et de l'**intelligence artificielle**. Il s'agit de la synthèse du débat public que la Cnil a animé durant l'année 2017 et que sa présidente, Isabelle Falque-Pierrotin, a présentée le 15 décembre 2017, en présence de Mounir Mahjoubi, secrétaire d'État chargé du numérique, et de Cédric Villani, député chargé par le gouvernement d'une mission sur l'**intelligence artificielle**. Ce texte vise principalement à fournir un panorama des questions d'ordre éthique que les algorithmes soulèvent.

La loi pour une République numérique du 7 octobre 2016 avait confié à la Cnil la mission de mener une réflexion sur les nouvelles questions de société engendrées par l'évolution des technologies numériques. Aussi, entre janvier et octobre 2017, elle a organisé de vastes discussions autour de ces problématiques. Celles-ci, ouvertes et décentralisées, ont réuni plus de 3 000 personnes à l'occasion de 45 manifestations partout en France, dont une journée de concertation citoyenne. Ont été abordés, entre autres, l'éducation, la santé, les médias, l'emploi, le secteur bancaire et financier, la justice, la sécurité et la défense ou encore la vie publique et politique. Uniques en Europe, ces débats ont abouti à la rédaction de ce texte comprenant quelque 80 pages, préconisant plus d'un moyen d'éviter que les hommes ne se retrouvent dépassés par les machines. Surtout, **il s'agit de démystifier l'informatique et la robotique, ainsi que leurs applications**, afin de porter un regard éclairé sur leurs potentialités et les éventuels périls qu'ils feraient surgir (destructions massives d'emplois, américanisation de la société et de l'économie défavorable à l'Europe, disparition de l'humanité au profit de l'automatisme etc.). Les algorithmes suscitent à la fois l'intérêt du public et un certain malaise en raison des craintes qui les accompagnent. À la faveur de son rapport, la Cnil souhaite nourrir et préciser les problèmes et les solutions afin de « *garantir que l'**intelligence artificielle** augmente l'homme plutôt qu'elle ne le supplante* ». Elle entend aussi contribuer à « **l'élaboration d'un modèle français de gouvernance éthique de l'**intelligence artificielle**** ». Les algorithmes ne doivent pas être autre chose qu'un outil dans la main de l'homme, utilisé de manière transparente et responsable.

Les algorithmes ne sont guère une nouveauté. Déjà, dans les années 1970, avant que la

Silicon Valley ne devienne leur eldorado, les systèmes informatiques experts fonctionnaient à base d'algorithmes. Aujourd'hui, l'[intelligence artificielle](#) va de pair avec leur perfectionnement exponentiel. Il s'agirait de la nouvelle prose de Monsieur Jourdain, pratiquée par chacun assez inconsciemment. C'est pour cela que les algorithmes d'aujourd'hui, d'une part, orientent insidieusement les conduites de ceux qui ne réalisent pas combien ils sont soumis au diktat algorithmique et, d'autre part, inquiètent profondément ceux qui prennent conscience de leur forte influence.

À l'ère du numérique et de l'internet, les algorithmes et l'[intelligence artificielle](#) sont partout. Ce sont eux qui suggèrent des « amis » et des activités sur les réseaux sociaux, qui fixent les prix sur les sites d'e-commerce et indiquent aux clients quoi acheter, qui affectent les étudiants dans les établissements d'enseignement supérieur, qui conduisent les automobiles autonomes et pilotent les avions, qui battent les plus grands champions du jeu de go ou des échecs, qui simulent l'évolution de la propagation de la grippe en hiver, qui établissent les bulletins météo, qui indiquent aux avocats quelle décision les juges devraient rendre en fonction des circonstances de l'espèce, qui suggèrent enfin aux forces de police les zones où patrouiller en priorité etc. Grâce au perfectionnement des techniques d'apprentissage automatique, les robots effectuent des tâches de plus en plus complexes, parfois critiques, et souvent mieux que des humains. Or, le fonctionnement des algorithmes est indissociable du [big data](#), de la production et de la diffusion de masses de données, dont une part non négligeable est constituée d'informations personnelles - ce qui préoccupe nécessairement la Cnil.

Dans son rapport, celle-ci prend soin tout d'abord de définir ce qu'est un algorithme et ce qu'est l'[intelligence artificielle](#). Cela n'est pas sans intérêt, tant il est souvent difficile de décrire avec précision ces objets technologiques. Un algorithme est ainsi défini comme « *une suite finie et non ambiguë d'instructions permettant d'aboutir à un résultat à partir de données fournies en entrée* ». Quant à l'[intelligence artificielle](#), il s'agit d' « *une nouvelle classe d'algorithmes, paramétrés à partir de techniques dites d'apprentissage* ». Les instructions à exécuter ne sont plus programmées par un développeur humain, mais générées automatiquement par la machine, laquelle se perfectionne ainsi au fur et à mesure en tirant les leçons de ses expériences passées. Par conséquent, leur logique et leur mode de fonctionnement deviennent de plus en plus opaques et incompréhensibles.

La Cnil, soulignant néanmoins que les peurs entourant les algorithmes et l'[intelligence artificielle](#) appartiennent largement à l'ordre des fantasmes, dresse la liste des problématiques éthiques qui se posent. Celles-ci seraient au nombre de six :

1° L'autonomie humaine au défi de l'autonomie des machines

Grâce aux algorithmes, les hommes peuvent s'en remettre toujours davantage à l'[intelligence artificielle](#), donc « robotiser » leurs activités et leurs prises de décision, y compris lorsque celles-ci sont de première importance. Or, souligne la CNIL, il faut prendre garde à ne pas se laisser aveugler par l'impression d'infaillibilité et de neutralité que les algorithmes peuvent dégager ; il convient de ne pas leur faire une confiance aveugle. Les hommes ne doivent pas profiter du recours aux algorithmes pour diluer leurs responsabilités. En somme, ils ne doivent pas s'abandonner aux algorithmes.

## 2° Biais, discrimination et exclusion

Ce sont, pour la CNIL, des effets importants des algorithmes et de l'[intelligence artificielle](#). Ces phénomènes, parfois recherchés volontairement par les programmeurs, sont généralement imperceptibles du point de vue de ceux qui en subissent les effets. Il semble nécessaire de les identifier, afin de mettre en garde contre les dérives algorithmiques. Mais, dans le monde fragmenté et obscur de l'algorithmie, pareille identification apparaît délicate.

## 3° Fragmentation algorithmique : la personnalisation contre les logiques collectives

Avec le web, les algorithmes servent surtout à personnaliser les contenus proposés aux internautes en fonction de leurs historiques de navigation. Bien que permettant de faciliter la vie en ligne des individus, ce profilage tend aussi à les enfermer dans des formes d'autopropagande invisible nuisant au pluralisme indispensable au bon fonctionnement de la mécanique démocratique.

## 4° Entre limitation des mégafichiers et développement de l'[intelligence artificielle](#) : un équilibre à réinventer

La Cnil s'inquiète des quantités incommensurables de données personnelles que les algorithmes brassent à des fins d'apprentissage. Toutefois, loin de poser un regard dogmatique sur cette problématique, elle note que le traitement des mégafichiers est aussi la source de nombreuses promesses. Elle pose dès lors la question de savoir « *comment exploiter ces données tout en maintenant l'impératif de protection des libertés individuelles de chacun ?* »

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

L'homme intervient en sélectionnant les données « nourrissant » les algorithmes et leur permettant de s'entraîner et de se perfectionner. Se posent dès lors la question des modalités et de la transparence de cette sélection, afin de retenir les données les plus utiles

sans risquer d'induire en erreur les machines.

#### 6° L'identité humaine au défi de l'[intelligence artificielle](#)

Enfin, l'[intelligence artificielle](#) interroge les spécificités de l'intelligence humaine, dès lors notamment qu'émergent des formes d'hybridation entre humains et machines. « *Comment appréhender cette nouvelle classe d'objets que sont les robots humanoïdes, susceptibles d'engendrer chez l'homme des formes d'affect ?* », s'interroge la Cnil.

Face à ces défis, la Haute Autorité propose de consacrer deux « principes fondateurs » supposés contenir l'ensemble des réponses éthiques à apporter afin de façonner une [intelligence artificielle](#) au service de l'homme : la loyauté et la vigilance. Ces principes s'inscriraient dans une nouvelle génération de droits fondamentaux capables d'organiser la gouvernance internationale du monde numérique.

Le principe de loyauté implique que tout algorithme, qu'il traite ou non des données personnelles, ne cherche pas à manipuler ses utilisateurs. Ceux-ci doivent être considérés non comme des consommateurs mais comme des citoyens. Institué en matière de traitement de données personnelles par la loi du 6 janvier 1978, le principe de loyauté a été repris par la loi du 7 octobre 2016 pour une République numérique, spécialement à l'égard des plateformes du web participatif ou collaboratif qui sont tenues à une obligation d'information des utilisateurs quant aux critères de classement et de référencement mis en œuvre. Et ce principe avait déjà été évoqué par le Conseil d'État en 2014 dans son étude annuelle sur le numérique et les droits fondamentaux (voir [La rem n°32, p.61](#)). Selon l'institution, « *la loyauté consiste à assurer de bonne foi le service de classement ou de référencement sans chercher à altérer ou à détourner à des fins étrangères à l'intérêt des utilisateurs* ».

La Cnil juge que le principe de loyauté devrait concerner tous les algorithmes et que la notion d' « *intérêt des utilisateurs* » ne devrait pas viser les utilisateurs des plateformes individuellement, mais en tant que « collectif ». Un algorithme dit « loyal » ne devrait donc pas susciter, reproduire ou renforcer des discriminations. Ainsi, le principe de loyauté, tel qu'il est repris par la Cnil, vise à défendre l'intérêt général et les principes fondamentaux de la démocratie, mais aussi certains groupes ou certaines communautés dont l'existence pourrait être remise en question par les prophéties autoréalisatrices des algorithmes. Pour la Cnil, « *l'intérêt des utilisateurs doit primer* ». Par exemple, un tel principe pourrait obliger à repenser le régime juridique des réseaux sociaux en période électorale ou, plus généralement, en cas de débat public, cela afin de limiter les effets de la segmentation du corps politique par le ciblage de l'information.

Cependant, l'autorité indépendante note qu'il peut être difficile de concilier le principe de loyauté avec la montée en puissance des algorithmes de *machine learning* dès lors que leur mode de raisonnement et de fonctionnement peut échapper y compris à leurs concepteurs. Le problème d'un algorithme qui, par exemple, oriente des demandeurs d'emploi vers certains postes ou affecte des étudiants dans certaines filières, est que la machine n'est pas capable d'expliquer pourquoi elle opère ces choix plutôt que d'autres.

Quant au principe de vigilance, prolongeant directement le principe de loyauté, il consiste à prévoir un questionnement réflexif, régulier et méthodique afin d'identifier les dérives possibles de ces objets technologiques mouvants et incertains que sont les algorithmes. Ce principe vise donc à se prémunir contre le risque de confiance excessive en la fiabilité de l'*intelligence artificielle* en instituant une forme d' « obligation de douter ». Il s'agit de responsabiliser les différents acteurs de la création et de l'utilisation des algorithmes.

Surveiller les résultats produits par les algorithmes doit permettre d'identifier les discriminations qu'elles peuvent produire et, par la suite, de corriger ces résultats et de corriger ces algorithmes. Et la Cnil de préciser que « l'ensemble des maillons de la chaîne algorithmique (concepteurs, entreprises, citoyens) doivent être mobilisés pour donner corps à ce principe, au moyen de procédures concrètes ».

À la lumière de ces enjeux et de ces principes, la CNIL en vient à formuler six « recommandations opérationnelles » à l'attention des autorités publiques ainsi que des entreprises :

- 1° Former à l'éthique tous les acteurs-maillons de la « chaîne algorithmique » (concepteurs, professionnels, citoyens).
- 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'*intelligence artificielle* 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).

À travers les problématiques soulevées et les préconisations prodiguées, la Cnil cherche à imposer sa vision du numérique et des algorithmes, soit une approche éthique dans laquelle les performances économiques sont soumises à la responsabilité. Ainsi ouvre-t-elle la voie vers une déontologie des algorithmes, laquelle a vocation à s'appliquer non aux machines mais aux hommes qui les créent et les exploitent.

**Comment permettre à l'Homme de garder la main ? Rapport sur les enjeux éthiques des algorithmes et de l'intelligence artificielle**, Commission nationale de l'informatique et des libertés (Cnil), décembre 2017.