



Intelligence artificielle

• Éléments scientifiques et juridiques

Le dernier rapport¹ sur « l'Intelligence artificielle » (IA) de la Commission Nationale Informatique et Libertés (CNIL) de décembre 2017 souligne que l'IA est le « grand mythe de notre temps » et que sa définition reste très imprécise. Quand on parle d'IA, on pense aux publicités ciblées sur internet, aux objets connectés et à l'internet des objets, aux traitements de données massives et hétérogènes sorties d'une enquête, aux machines numériques et aux robots humanoïdes capables d'apprentissage et d'évolution, aux véhicules autonomes, aux interfaces cerveau-machine et aux algorithmes.

Les algorithmes sont en quelque sorte les « squelettes de nos outils informatiques ». Ils sont des systèmes d'instructions permettant à la machine numérique de fournir des résultats à partir des données fournies. Ils sont à l'œuvre lorsqu'on utilise sur Internet un moteur de recherche, ou bien lorsqu'on propose un diagnostic médical à partir de données statistiques, mais aussi pour choisir un itinéraire en voiture et sélectionner des informations sur les réseaux sociaux en fonction des goûts des réseaux d'amis. Les algorithmes appartiennent au concepteur et restent très souvent méconnus des utilisateurs. Ils deviennent capables de tâches de plus en plus complexes grâce à la puissance de calcul qui s'accroît de manière exponentielle et grâce aux techniques d'apprentissage automatique (réglage automatique des paramètres d'un algorithme pour qu'il produise les résultats attendus à partir des données fournies). « L'apprentissage profond » rencontre ainsi de nombreux succès. Il permet déjà de reconnaître des images et des objets, d'identifier un visage, de piloter un robot dit intelligent ...

Les liens entre neurosciences et IA sont à la base du projet européen *Human Brain Project* dont l'un des objectifs est de simuler le comportement du cerveau humain. On peut aussi utiliser l'IA pour mieux comprendre des maladies neuronales comme des troubles compulsifs ou des dépressions. Il s'agit donc de construire des machines dites intelligentes à la fois pour piloter des systèmes évolutifs et pour participer à la compréhension du cerveau humain.

La loi Informatique et libertés du 6 janvier 1978, modifiée régulièrement depuis, dit que « l'informatique doit être au service de chaque citoyen...ne doit porter atteinte ni à l'identité humaine, ni aux droits de l'homme, ni à la vie privée, ni aux libertés individuelles ou publiques ». Elle définit notamment les principes à respecter pour la collecte, le traitement et la conservation des données personnelles. Elle précise les pouvoirs et capacités de sanction de la CNIL. Le nouveau règlement européen sur la protection des données personnelles (RGDP), adopté le 27 avril 2016, entre en fonction le 25 mai 2018 dans les États membres de l'Union Européenne pour renforcer la régulation juridique. C'est pourquoi un projet de loi (n. 490) a été déposé à l'Assemblée nationale le 13 décembre 2017.

• Les problèmes posés

Pour beaucoup, l'IA est une formidable occasion à saisir en termes d'économie de la connaissance. Ses apports dans les domaines de la médecine, de la robotique, de l'apprentissage et des sciences notamment sont déjà considérables. Mais comment apprivoiser l'IA pour qu'elle soit vraiment au service de tous ?

Parmi les craintes et les risques les plus exprimés, on trouve les problèmes de suppression d'emplois avec les robots. On constate aussi une méfiance, voire une « perte d'humanité » devant la « boîte noire » que représentent les algorithmes « qui nous gouvernent », sur Internet, sur les réseaux sociaux, dans le e-commerce et jusque dans notre vie privée. Mais ces algorithmes pourraient aussi gouverner le médecin et l'employeur qui se reposeraient sur eux pour prendre leurs décisions. « Qui contrôle quoi ? », telle est la question souvent soulevée. Ainsi on se demande quels sont les « biais » par lesquels les jugements sont portés pour recruter un employé grâce à l'IA, avec suspicion de discrimination.

Outre la protection des données personnelles, les grandes questions du rapport CNIL sont :

- Devant la puissance des machines, comment appréhender les formes de dilution de responsabilités dans les décisions ? Jusqu'où peut-on accepter « l'autonomie des machines » qui peuvent décider pour nous ?
- Comment faire face au manque de transparence des algorithmes quant aux biais qu'ils utilisent pour traiter les données et « décider des résultats » ?
- Comment appréhender cette nouvelle classe d'objets que sont les robots humanoïdes susceptibles de conduire à des formes d'attachement (par exemple des personnes âgées aidées par ces robots dans leur vie quotidienne) ?

¹ Comment permettre à l'homme de garder la main ?, Rapport de la CNIL, publié le 15 décembre 2017.

Quel statut donner aux robots dits intelligents, et quelles conséquences au niveau juridique en terme de responsabilité en cas de problème ?

Devant les risques d'une forme possible de « dictature de la technologie numérique » plus ou moins invisible, le rapport de la CNIL plaide pour **deux principes fondateurs pour l'éthique de l'IA** :

- **loyauté collective** (pour la transparence et l'utilisation démocratique des algorithmes par exemple) ;
- **vigilance/réflexivité** par rapport à l'autonomie des machines et aux biais qu'elles propagent ou génèrent, afin que l'homme ne « perde pas la main » sur l'IA.

• Les visées anthropologiques et éthiques

Poser la question du statut des robots est signifiant d'un « trouble » introduit par l'IA concernant le rapport que l'homme entretient avec ses « machines apprenantes ». Une résolution du Parlement européen encourage les recherches sur l'octroi du statut de « personne électronique » à certains robots². Cette expression juridique relativiserait la notion de personne, qui s'enracine dans la dignité de l'être humain³. L'expression « robot cognitif », par exemple, serait préférable. Cependant, jusqu'où les capacités des machines pourront-elles se rapprocher de celles des humains, puis les dépasser ? Certains transhumanistes extrêmes attendent ce moment où l'IA dépassera l'intelligence humaine, sorte de « singularité » à partir de laquelle une fusion homme-machine constituera un « cyborg » qui prendra la suite de l'*homo sapiens* !

Sans entrer dans de tels fantasmes, des célébrités comme Stephen Hawking, Bill Gates et Elon Musk ont plusieurs fois fait part de leurs inquiétudes sur l'IA⁴. Ils expriment leur peur que les machines apprenantes nous contrôlent, car elles auront des compétences statistiques et combinatoires bien supérieures aux nôtres ainsi qu'un accès à des bases de données gigantesques que l'homme ne peut traiter directement. **C'est essentiellement sur l'aspect calculatoire que la puissance des machines numériques est aujourd'hui appréhendée. Seule cette forme d'intelligence est en jeu, alors que l'homme a de nombreuses formes d'intelligence (rationnelle, émotionnelle, artistique, relationnelle, ...)**. Certes, on comprend que les puissants calculateurs permettent à la machine de trouver les combinaisons pour battre les champions du jeu de Go. Cependant, l'IA se situe aujourd'hui sur le terrain de la simulation. Or, il y a un seuil entre « simuler » une émotion et l'éprouver. L'émotion, avec sa dimension communicationnelle, conduit l'homme qui l'éprouve à attribuer une valeur aux choses à partir de laquelle il pose des choix de vie quotidienne. L'émotion exprime la richesse de l'homme vulnérable. La machine apprenante n'en est pas là ! L'IA humanise-t-elle ? Elle est de fait un « pouvoir » qui doit être soumis au discernement face à la fragilité et la vulnérabilité comme sources d'humanisation. De même, il est impossible de comparer la conscience humaine (existentielle, psychologique et morale) avec une éventuelle conscience des machines⁵.

À propos de l'IA, l'idée est souvent répandue que « penser, c'est calculer ». Cela entraîne bien des confusions. L'homme, doté d'une intelligence faite pour la vérité et la sagesse, a un autre registre de pensée beaucoup plus varié, vaste et subtil⁶. Notre conscience se situe dans un corps façonné par des millions d'années d'évolution, avec de belles capacités de raison, de création, de vie psychique et de profondeur spirituelle, qui vont bien au-delà des combinatoires les plus sophistiqués. **Certains soulignent qu'au lieu de pointer les risques d'une intelligence calculatoire et combinatoire des machines, il est plus urgent de rendre publiques les valeurs que les concepteurs d'algorithmes introduisent dans leurs logiciels. La transparence des algorithmes est une véritable question de fond. Leur conception a-t-elle toujours pour visée l'amélioration du soin et le service de la dignité humaine ?**

• Des références bibliographiques pour continuer le travail

IA, *Promesses et péril*, Cahier du Monde n. 22696, 31 décembre 2017-2 janvier 2018.

Frédéric Alexandre et Serge Tisseron, « Où sont les vrais dangers de l'IA ? », in *Les robots en quête d'humanité*, Pour la Science, n. 87, avril-juin 2015, p.102-107.

Milad Doueihy et Frédéric Louzeau, *Du matérialisme numérique*, Hermann, 2017.

Serge Abiteboul et Gilles Dowek, *Le temps des algorithmes*, Le Pommier, 2017.

2 février 2018

² <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+TA+P8-TA-2017-0051+0+DOC+PDF+V0//FR>

³ La notion de « personne morale » ne permet pas d'ambiguïté. Par ailleurs, elle n'est pas reconnue par toutes les traditions juridiques.

⁴ Voir Alexandre Picard, « L'intelligence artificielle, star inquiétante du web summit à Lisbonne », *Le monde économie*, 10 novembre 2017.

⁵ Voir par exemple Mehdi Khamassi et Raja Chatila, « La conscience d'une machine », in *Les robots en quête d'humanité*, Pour la Science, n° 87, avril-juin 2015. Cf. Vatican II, constitution *Gaudium et spes*, n. 16 ; Déclaration *Dignitatis humanae*, n. 1-3 ; Jean-Paul II, encyclique *La splendeur de la vérité*, 6 août 1993, n. 31-34.

⁶ Cf. Vatican II, constitution *Gaudium et spes*, n. 15 ; Jean-Paul II, encyclique *Foi et raison*, 14 septembre 1998, n. 16-33.

« Soyez toujours prêts à rendre raison de votre espérance mais avec douceur et respect » (1 Pierre 3, 15-16)