

DÉFIS ÉTHIQUES EN PAGAILLE

L'intelligence artificielle va changer le monde. Peut-on encadrer ses révolutions ? Pour les experts le pire n'est pas certain.

L'Express (France) · 30 janv. 2019 · Par Victor Garcia

L'intelligence artificielle (IA) promet de révolutionner la société. Va-t-elle dépasser l'homme et l'écraser ou l'aider à relever les grands défis à venir ? Mettra-t-elle des millions de travailleurs au chômage ? Remplacera-t-elle les médecins ? Tirera-t-elle à la place des soldats ? Pendant longtemps, ces questions ont été l'apanage d'oeuvres de science-fiction imaginant souvent l'avènement d'une « Singularité », une IA anthropomorphe transcendant l'être humain, pour le meilleur ou pour le pire. Aujourd'hui, certains envisagent sérieusement ce scénario, comme Ray Kurzweil, directeur de l'ingénierie de Google. Mais la plupart des scientifiques ne partagent pas cette vision. « Les tenants de la Singularité disent que la vitesse des ordinateurs double tous les deux ans, et donc que l'IA va dépasser le cerveau humain d'ici à 2045 », résume Raja Chatila (1), directeur de l'Institut des systèmes intelligents et de robotique. Absurde, selon lui, puisqu'il y a longtemps que les ordinateurs calculent plus vite que l'homme, ce qui ne les empêche pas de rester enfermés dans un fonctionnement systématique, sans libre arbitre et sans conscience. Quel est l'ingrédient manquant ? « On n'en sait fichtre rien », balaie-t-il d'un revers de la main. Ce concept de singularité inquiète aussi, puisqu'il est intimement lié au transhumanisme, dont Ray Kurzweil est l'un des chantres. « Cette idéologie visant à améliorer l'humanité grâce aux technologies a des relents politiques qui rappellent les pires heures de l'Histoire. Elle ouvre la boîte de Pandore. Or on connaît les conséquences d'une légitimité basée sur la science comme si elle n'était pas discutable », s'agace Raja Chatila.



Jean-Gabriel Ganascia (2), président du comité d'éthique du CNRS et professeur d'informatique à Sorbonne Université, auteur du livre *Le Mythe de la Singularité* (Seuil), reconnaît que ces questions ont un intérêt philosophique. Il souligne néanmoins qu'elles ne doivent pas masquer les véritables enjeux de l'IA. « Raconter des histoires qui font peur empêche de se rendre compte de ce qu'il se passe, en l'occurrence de la prise de pouvoir des géants du numérique américains et chinois », relève-t-il. L'explosion de l'IA est en réalité limitée à quelques-uns de ses champs d'étude : principalement à l'apprentissage automatique (machine learning), en particulier à l'apprentissage par renforcement (reinforced learning) et surtout à l'apprentissage profond (deep learning), qui fait appel aux réseaux de neurones artificiels. Ces algorithmes apprennent par eux-mêmes à établir des corrélations statistiques en se nourrissant de bases de données colossales, qui, pour la plupart, sont la propriété des mastodontes d'Internet. Le scandale

Cambridge Analytica, une société britannique accusée d'avoir analysé illégalement 87 millions de profils Facebook pour influencer sur la dernière campagne présidentielle américaine et le vote sur le Brexit, en 2016, illustre la problématique. « Cette affaire montre non seulement qu'il est possible d'interférer avec des processus démocratiques, mais aussi que les responsables ne sont pas les machines, mais les personnes qui les emploient », analyse Jean-Gabriel Ganascia.

INQUIÉTANTE « BOÎTE NOIRE »

« L'autre problématique fondamentale est la manière dont les bases de données sont utilisées et les biais statistiques qu'ils peuvent induire, poursuit Raja Chatila. Comme exploiter des statistiques concernant un groupe pour classer des individus dans des cases. » Ainsi, par exemple, le déploiement d'une IA prédictive dans le système judiciaire américain censée aider les juges à évaluer les risques de récidive des condamnés en fonction de certains critères. Or, aux Etats-Unis, la majorité des détenus sont des Noirs ou des Latinos. « L'IA va envenimer les choses en perpétuant les statistiques et les situations existantes, au lieu d'y remédier », regrette le roboticien. Pire, même les chercheurs qui développent les algorithmes basés sur les réseaux de neurones artificiels ont du mal à expliquer comment ils prennent des décisions. Une « boîte noire » inquiétante, particulièrement quand cette technologie est appliquée aux assurances, à la médecine, à la justice ou à l'armée.

Conscients de ces problématiques, les leaders du domaine s'échinent à montrer patte blanche. Facebook a lancé le 20 janvier son Institut pour l'éthique de l'IA, à l'Université technique de Munich (Allemagne). Google s'est désengagé des recherches militaires refusant des contrats de plusieurs milliards avec le Pentagone et s'est doté d'un pôle « éthique des opérations machine learning ». Sa dirigeante, Jen Gennai (3), assure que son entreprise vise le bien commun. « Nous savons, par exemple, que les voitures autonomes feront des erreurs, mais aussi qu'elles seront plus sûres : le nombre de vies sauvées sera plus grand que le nombre d'erreurs commises, le compromis en vaut la peine », résume-t-elle. Quant à la gestion des données, Jen Gennai rappelle que Google demande toujours le consentement de ses utilisateurs et protège leurs informations. Une promesse qui ne résiste pas toujours à l'épreuve des faits, en témoigne la récente amende de 50 millions d'euros infligée par la Commission nationale de l'informatique et des libertés, ou la décision de DeepMind, filiale de Google, de transmettre à sa maison mère les données de santé de 1,6

million de Britanniques, après avoir promis de ne jamais le faire. Comment s'assurer que l'IA reste éthique ? « D'abord en travaillant sur la neutralité des bases de données, en appliquant des règles aussi strictes que pour les sondages d'opinion et leurs échantillons représentatifs », propose Raja Chatila. Ne pourrait-on pas coder l'éthique ? Si les machines ne peuvent comprendre ce qu'est l'éthique, il n'est pas pour autant impossible de les concevoir de manière qu'elles prennent des décisions allant dans ce sens, explique-t-il. L'enjeu majeur consistera à mener des arbitrages sur l'équilibre entre trois grandes exigences contradictoires : la protection de la vie privée, la sécurité et la transparence. « Jusqu'à maintenant, l'Europe a raté la marche industrielle en ne parvenant pas à créer de géant du Web, remarque de son côté Nicolas Mialhe (4), président de The Future Society, un incubateur d'idées et de projets lancé à la Harvard Kennedy School of Government. Mais elle reste en avance sur les questions éthiques, un domaine qu'elle pourrait mettre à profit pour reprendre la main. » Bonne nouvelle, l'idée d'un « Giec de l'IA » (G2IA ou IPAI, en anglais), proposée par Emmanuel Macron en mars 2018, devrait voir le jour cette année. L'occasion pour le Vieux Continent d'imposer, enfin, ses idées ?