

## La course à l'éthique est lancée

AU CŒUR DE L'IA. Pour accompagner l'avancée des technologies, entreprises et chercheurs s'interrogent en amont sur les limites à fixer.

LE MONDE | 02.01.2018 à 08h00 | Par David Larousserie ([journaliste/david-larousserie/](#))



MARIO WAGNER

Un programme de voiture autonome choisira-t-il d'écraser un piéton pour sauver les passagers ou de donner un coup de volant, au risque de précipiter le véhicule contre un mur ? En cas d'erreur de diagnostic assisté par un ordinateur, qui sera responsable, le médecin ou le concepteur du logiciel ?

**Lire aussi :** [Enquête au cœur de l'intelligence artificielle, ses promesses et ses périls](#)

([pixels/article/2017/12/30/l-intelligence-artificielle-ses-promesses-et-ses-perils\\_5236008\\_4408996.html](#))

Bien des entreprises ont anticipé ces questions. Fin 2015, une société à but non lucratif, OpenAI, est créée par le Sud-Africain Elon Musk et Peter Thiel, un entrepreneur américain, fondateur de PayPal et investisseur dans Facebook. OpenAI, soutenue notamment par Microsoft, se fixe pour objectif de « *mettre en œuvre le chemin menant à la sécurité de l'intelligence générale artificielle* ». Un an plus tard, Partnership on AI est lancé par plusieurs géants du Net et des technologies (Amazon, Google, Microsoft, Apple, IBM...). Son but : « *Proposer de bonnes pratiques dans les technologies d'intelligence artificielle.* »

### Treize « problèmes »

D'autres initiatives sont déjà plus avancées. Le 12 décembre, la plus grande société savante du domaine, l'Institut des ingénieurs en électricité et électronique (IEEE), a rendu publique la seconde version de son texte sur le « design éthique », un an après une première publication sur la question. « *Le document identifie une série de problèmes et fait déjà des propositions. Dans certains cas, cela pourrait aboutir à définir des standards industriels, comme le fait régulièrement l'IEEE* », explique Raja Chatila, directeur de l'Institut des systèmes intelligents et de robotique (université Paris-VI) et coordinateur de cette initiative. Treize « problèmes » sont identifiés. On y trouve ainsi des questions sur la sécurité de ces systèmes, la protection de la vie privée, le contrôle des armes autonomes, les conséquences économiques et sociales (notamment sur l'emploi)...

**Lire aussi : Des « juges virtuels » pour désengorger les tribunaux** ([/pixels/article/2018/01/01/des-juges-virtuels-pour-desengorger-les-tribunaux\\_5236479\\_4408996.html](/pixels/article/2018/01/01/des-juges-virtuels-pour-desengorger-les-tribunaux_5236479_4408996.html))

D'autres aspects sont plus inattendus, plus complexes aussi, comme la nécessité d'introduire des valeurs et des normes humaines dans les logiciels et surtout d'y penser avant la conception. Un exemple de cette « éthique par le design » est celui de la start-up française Snips. Ses algorithmes de reconnaissance vocale tournent dans les objets connectés et ne nécessitent donc pas d'envoyer la voix à analyser sur des serveurs extérieurs ; ce qui évite de facto les fuites et les atteintes possibles à la vie privée.

### Sept principes

Le document s'interroge aussi sur la façon de mesurer le bien-être individuel que les machines sont censées améliorer. Ou sur la manière d'éviter que les futures machines dotées d'émotions ne causent plus de mal que de bien aux humains. « *Quels effets les robots compagnons auront-ils sur les capacités cognitives sociales des utilisateurs ?* », s'interroge Raja Chatila. *S'attacher à une machine ne risque-t-il pas de créer un nouvel isolement alors même que ces robots doivent y remédier ?* »

**Lire aussi : Que savent vraiment faire les algorithmes ?** ([/pixels/article/2018/01/01/que-savent-vraiment-faire-les-algorithmes\\_5236426\\_4408996.html](/pixels/article/2018/01/01/que-savent-vraiment-faire-les-algorithmes_5236426_4408996.html))

En novembre 2017, après un colloque organisé par l'université de Montréal, paraissait une déclaration en « sept principes », ouverts à discussion jusqu'en mars 2018. Leur objectif : s'assurer que les développements de ces technologies rendent nos sociétés meilleures. Le texte ne cible pas que les hommes mais tous les êtres à même de ressentir des émotions, de la douleur. Selon ces principes, une intelligence artificielle (IA) devra donc respecter aussi les autres vertébrés, voire certains invertébrés, par exemple le poulpe.

Autre souhait formulé dans cette déclaration : « *L'IA devrait promouvoir la pensée critique et nous prémunir contre la propagande et la manipulation.* » Signe des temps, le laboratoire d'IA de Montréal, leader mondial sur ces sujets, a recruté en juillet un philosophe spécialiste de psychologie morale, le Canadien Martin Gibert.

### Protection des données et transparence des algorithmes

Ces sept principes font écho à une autre déclaration, dite d'Asilomar, formulée en janvier 2017 par l'Institut de l'avenir de la vie (Future of Life Institute), créé par plusieurs scientifiques afin de « *limiter*

*les risques existentiels menaçant l'humanité* ». Elle comporte 23 principes parmi lesquels la volonté que l'IA bénéficie à tous, la demande d'éviter la course aux armes autonomes, le respect des droits de l'homme et des libertés, mais aussi l'appel à anticiper le développement de superintelligences par des mesures de sécurité et de régulation.

Au-delà de ces risques potentiels, des soucis plus terre à terre se posent déjà, notamment sur la protection des données personnelles et la transparence des algorithmes. « *Des algorithmes ou des modèles biaisés peuvent conduire à des traitements inégaux et injustes pour certains groupes sociaux ou ethniques* », s'inquiète ainsi un autre groupe de réflexion américain, AI Now, dans un rapport sorti en octobre 2017.

En France, c'est aussi l'un des constats du rapport de la Commission nationale de l'informatique et des libertés (CNIL) du 15 décembre 2017. L'institution plaide pour rendre intelligibles les algorithmes et recommande l'audit des programmes ou une obligation de communiquer les informations permettant de comprendre la logique de fonctionnement d'un programme. « *Pour que l'homme garde la main* », précise-t-elle en préambule.