

MACHINES

INTELLIGENCE ARTIFICIELLE ▶

Faut-il avoir peur des progrès récents de l'IA? L'autonomie de ces systèmes « intelligents » pose d'ores et déjà des problèmes éthiques. Une réflexion s'est engagée afin d'éviter les dérapages

DAVID LAROUSSE

Monteriez-vous dans une voiture sans pilote, conduite par une intelligence artificielle? Si cela évite des accidents, sans doute. Mais si son programme était prêt à vous sacrifier en donnant un coup de volant afin d'éviter d'écraser dix piétons? A cette perspective, le taux d'adhésion est nettement plus faible... Ce dilemme, détaillé par des psychologues dans la revue *Science* du 24 juin, illustre les innombrables cas de conscience que fait surgir l'usage de l'intelligence artificielle (IA). Un concept où convergent informatique, mathématiques, sciences de l'ingénieur et neurosciences, visant à doter diverses machines de capacités « intelligentes » : reconnaissance d'images ou de voix, analyse de situation, choix entre plusieurs actions.

Depuis sa naissance dans les années 1950, cette discipline est sollicitée pour des tâches de plus en plus performantes : pilotage automatique d'avions ou de voitures, animation de robots plus ou moins humanoïdes, recommandations d'achats en ligne... Avec des hauts et des bas selon les périodes. Ces derniers mois, elle est au sommet de la vague, en matière de promesses comme de peurs. Côté promesses, la victoire d'un programme informatique contre l'un des meilleurs joueurs de go du monde a marqué, en mars 2016, une étape dans la suprématie des machines. Côté peurs, des personnalités célèbres – le physicien Stephen Hawking, les entrepreneurs Elon Musk et Bill Gates, entre autres – s'inquiètent des risques d'une super-intelligence qui finirait par surpasser l'humain – mettant même fin à l'humanité.

Dans une lettre ouverte publiée en juillet 2015, plus d'un millier de personnalités ont ainsi réclamé un moratoire contre les armes autonomes. Une réticence que divers dérapages ont pu alimenter ces derniers mois : un algorithme de reconnaissance d'images de Google qui prend des personnes noires pour des singes, un accrochage entre une voiture autonome et un bus, un logiciel de conversation (Tay, de Microsoft) qui lance des propos racistes... Plus l'IA devient puissante, plus la crainte qu'elle devienne incontrôlable et néfaste à l'homme hante les esprits. Peur justifiée ou simple fantasme?

Scénarios catastrophes

« Beaucoup des scénarios catastrophes sont élaborés par des gens qui ne connaissent pas les limites actuelles du domaine. Or les spécialistes disent qu'ils sont loin de la réalité », estime le Français Yann LeCun, professeur à l'université de New York. Egalement directeur du laboratoire de recherche en IA de Facebook, il ne croit pas à ce moment de bascule où les machines seront supérieures à l'homme. « Elles vont simplement devenir de plus en plus intelligentes et de plus en plus faciles à utiliser. Cela amplifiera notre propre intelligence. » D'autres notent que l'IA dont il est pour l'instant question est une intelligence « étroite », « faible » et non « générale » comme peut l'être la nôtre. Elle n'effectue que des tâches bien cloisonnées. « Aujourd'hui, tous les systèmes intelligents ne comprennent pas ce qu'ils font. Ils peuvent reconnaître une voiture mais ne savent pas ce que c'est », rappelle Raja Chatila, directeur de l'Institut des systèmes intelligents et de robotique à l'université Paris-VI. Et c'est de ces limites, justement, que viennent les problèmes.

« Parler de ces hypothétiques super-intelligences est un faux débat qui détourne d'enjeux concrets et plus urgents », précise Laurence Devillers. Professeure d'informatique à l'université Paris-Sorbonne et

chercheuse au CNRS, elle pilote un groupe de travail sur ce thème au sein de la Commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique (Cerna). Le 13 juin, en introduction d'une journée de réflexion sur l'IA, elle a égrainé les questions que posent déjà les robots compagnons dans les hôpitaux, les véhicules autonomes, les agents conversationnels ou les drones. Qui décide du choix des cibles d'un robot tueur? Qui est responsable en cas d'accident d'une voiture autonome? Si un robot soignant comprend que son maître n'aime pas le sirop, le lui donnera-t-il s'il s'agit d'un médicament? Comment apprendre à un logiciel de conversation à ne pas faire de la surenchère raciste?

« La rapidité à laquelle ce domaine évolue nous oblige à répondre à ces questions », indique Danièle Bourcier, juriste au CNRS et membre de la Cerna. Les cerveaux s'agitent donc. L'Institut des ingénieurs en électronique et électrique – une société savante internationale – a mis sur pied un comité, mené par Raja Chatila, d'une dizaine de groupes de travail autour des « considérations éthiques dans le design des systèmes autonomes ». Il rendra ses conclusions en août, ébauche d'une charte qui abordera plusieurs aspects militaires, juridiques ou économiques de

l'IA. En France, l'Office parlementaire d'évaluation des choix scientifiques et technologiques (Opecst) a débuté des auditions sur ses enjeux économiques et sociétaux. Et les industriels eux-mêmes, tel Google, se dotent de comités d'éthique – lesquels semblent toutefois manquer de transparence.

« L'urgence de cette réflexion éthique est d'autant plus grande que nous sommes aujourd'hui face à une rupture technologique dans l'IA même, qui provient de ce qu'on a appris aux robots à apprendre », rappelle Laurence Devillers. Cette technique d'apprentissage est la clé des derniers succès. Elle consiste à faire progresser les programmes par la répétition. Soit en les entraînant sur des données connues, comme des images, soit en les faisant « jouer » contre eux-mêmes, comme au go, soit – plus compliqué – en faisant en sorte qu'ils apprennent sans supervision, sur des données inconnues. Un peu comme un enfant découvrant le monde en interagissant avec lui. « Sauf que si les parents ne savent pas toujours ce qu'ils font avec leurs enfants, les ingénieurs avec leurs algorithmes doivent le savoir », soulignait Raja Chatila lors des débats du 13 juin.

Mesurer la loyauté des programmes

« Ces systèmes sont vraiment une coconception entre l'homme et la machine », insiste Laurence Devillers. Comme en informatique classique, c'est bien un ingénieur qui écrit le programme. Mais il choisit aussi les données qui vont nourrir le modèle, et peut même y ajouter des règles « morales ». « Tout est très opaque dans ces systèmes numériques. Il est donc important de mettre en place des mesures de vérification de leur loyauté », souligne Laurence Devillers. Cette qualité morale est en effet essentielle au bon usage des machines intelligentes. Exemple : si un ingénieur entraîne son logiciel à reconnaître des visages sans que sa base de données soit représentative de la diversité humaine, il identifiera une personne noire comme étant un singe. Une plus grande transparence permettrait aussi de vérifier s'il n'y a pas d'idées préconçues – sociales, religieuses, ethniques – derrière ces programmes. Ou des biais favorisant des intérêts cachés, comme certains accusent le moteur de recherche de Google de le faire pour privilégier des services au détriment d'autres.

La loyauté des robots implique, enfin, de pouvoir contrôler et vérifier leur fonctionnement au fil de leur « éducation ». « Dans quelques années, ces systèmes s'adapteront à nous en co-évolution, et devront être contrôlés au fur et à mesure de leur apprentissage », explique Laurence Devillers, pour qui « l'éthique ne doit pas arriver à la fin, mais avant et pendant le travail de recherche ». Ce ne sera pas une mince affaire, puisque, par sa programmation même, le comportement de l'objet sorti d'usine se modifiera dans le temps. C'est ainsi tout un pan de recherche qui s'ouvre sur le contrôle, la vérification, l'évaluation de ces nouvelles machines.

Une fois ces défis éthico-techniques relevés, plusieurs solutions pourraient se dégager pour encadrer ce domaine en expansion : des permis pour robots, des autorisations de mise sur le marché ou encore des règles à respecter avant toute recherche, comme dans les laboratoires de biologie. « Nous apportons pour l'instant plus de questions que de réponses », confesse Laurence Devillers. Mais pour Danièle Bourcier, c'est une certitude : « Le numérique posant à la société des problèmes au moins aussi importants que la biologie ou la santé, les réflexions éthiques en cours déboucheront un jour ou l'autre sur des textes réglementaires. » ♦

VU DU WEB

LES JEUX SONT FAITS, RIEN NE VA PLUS

WILLIAM AUDUREAU

« Aïe », commente de manière lacunaire le site spécialisé Gameblog. « Scandaleux », selon son concurrent français Gamekult. « Un imposteur de seconde zone », corrobore le site américain IGN. Depuis sa sortie, le 24 juin, le jeu vidéo *Mighty n° 9* s'attire des critiques assassines et amères. Il n'est pourtant ni le premier ni le dernier navet dont accouche l'industrie vidéoludique. Mais si un projet n'avait pas le droit de décevoir, c'est bien celui-ci, le petit jeu indépendant porté par la notoriété d'une vieille gloire du jeu vidéo japonais, Keiji Inafune, mythique créateur de la série *Megaman* chère au cœur de nombreux trentenaires.

Hors des circuits d'édition traditionnels, ce sont ces nostalgiques qui ont presque entièrement financé le titre sur Kickstarter, plate-forme de financement participatif dont l'activité a explosé en 2012. Le projet y est lancé le 31 août 2013, et joue d'emblée sur la corde sensible : « *Un jeu d'action en 2D japonais à l'ancienne, revisité par Keiji Inafune, une dream team de vétérans de la saga Megaman, etc... VOUS.* » Plus de 65 000 internautes mettent la main à la poche pour donner vie aux aventures du successeur spirituel du petit robot bleu. En seulement deux jours, les 900 000 dollars demandés sont récoltés. Les sommes recueillies montent même jusque à 3,8 millions de dollars – soit le deuxième plus grand succès pour un projet de jeu vidéo sur Kickstarter.

Kickstarter, théâtre des illusions

Mais tout se complique. En dépit du ton assuré de la page de *Mighty n° 9*, qui se targue de compter dans ses rangs des développeurs aguerris, fiables et expérimentés, le nouveau robot de Keiji Inafune se perd dans les méandres du temps. Initialement prévu pour décembre 2013, il est reporté de multiples fois, avant de voir le jour avec deux ans et demi de retard. Pire, l'attente n'en valait pas la chandelle : le jeu est techniquement daté, mollasson, fade. Tout simplement ni fait ni à faire. « C'est toujours mieux que rien », aurait déclaré Keiji Inafune avec nonchalance (la phrase, qui a fait le tour du monde, a en fait été prononcée par son traducteur).

Pour les joueurs ayant financé l'existence du projet, la déconvenue a un goût amer. Plus encore que participatif, Kickstarter est un site de financement partisan. Les internautes qui s'y pressent sont des consommateurs militants, souvent animés par l'envie d'apporter leur contribution à un projet qui leur est cher. Ainsi de *Shenmue 3*, dernier volet d'une saga réputée que plus personne ne voulait financer ; *Yooka-Laylee*, jeu de plate-forme conçu par les créateurs de titres cultes des années 1990, comme *Banjo-Kazooie* ou *Bloodstained*, suite spirituelle des jeux d'action gothique *Castlevania*... Sur la plate-forme américaine, moyennant participation, chacun peut espérer redonner vie à sa madeleine de Proust. Avec les risques que cela comporte : la déception, voire, plus rarement, l'annulation pure et simple du projet.

Dans ce théâtre des illusions, chacun joue un rôle rationnel. Les développeurs indépendants vont chercher les fonds qui leur manquent. La plate-forme américaine se paye à la commission. Les internautes, eux, assument le rôle de banquiers et de producteurs, même s'ils ne bénéficient ni de l'intéressement des premiers ni du pouvoir de contrôle des seconds. A la place, ils se payent d'espoir. Sur Kickstarter, rêver a un prix. Mais c'est toujours mieux que rien. ♦

À PENSER