

La guerre automatique

Le Premier ministre et ex-ministre des Armées Sébastien Lecornu nous a prévenus : l'économie de guerre dans laquelle nous entrons doit permettre d'investir massivement dans « les ruptures technologiques récentes autour de la robotique, de l'intelligence artificielle (IA), de la numérisation du combat », les machines autonomes ayant « vocation à secondar, voire remplacer, le soldat dans nombre de domaines ». Cette automatisation de la guerre est en cours en Ukraine ou en Iran. Avec son association Pause IA, Maxime Fournes, ingénieur qui travaillait au développement de l'intelligence artificielle, alerte sur les menaces qu'elle fait peser sur l'humanité.

La Déroissance : L'intelligence artificielle – notamment le Maven Smart System de la société Palantir – est utilisée par les États-Unis et Israël pour mener leur guerre d'agression contre l'Iran. Peut-on résister à un ennemi qui dispose de télébots si l'on n'en est pas soi-même équipé ?

Maxime Fournes : La réponse est simple : non, on ne peut pas résister. C'est le problème clé quand on développe des IA de plus en plus puissantes. C'est évident dans le cadre de la guerre. Un des camps a entièrement équipé son armée de systèmes d'IA. Avec des entreprises comme Palantir, il a réussi à automatiser son commandement. Celui-ci a une vision globale du champ de bataille. Il peut alors réagir plus rapidement et plus intelligemment que des humains. Ses adversaires ne peuvent plus faire grand-chose.

C'est un défaut des films de science-fiction qui parlent d'IA : comme dans *Terminator*, il en ressort l'impression que face à des intelligences artificielles avancées, il demeurerait une espèce d'égalité entre les humains et les machines, que les humains pouvaient encore les vaincre. La réalité est tout autre. Dès qu'un des belligérants dispose d'une IA plus performante que les humains, son adversaire n'a aucune chance.

C'est pour cela que la seule façon d'arrêter cette dynamique, c'est de mettre en place un moratoire international pour empêcher les développeurs de créer ces systèmes avancés. Car une fois qu'ils existent, la boîte de Pandore est ouverte, et ceux qui les utilisent ont un

avantage immense sur les autres.

Le droit international n'a plus cours, nous voyons bien qu'à l'heure actuelle la loi du plus fort triomphe...

Si nous ne trouvons pas les moyens d'une coordination à l'échelle mondiale face à cette dynamique de compétition entre les protagonistes, c'est-à-dire les États et les entreprises qui développent les systèmes d'IA, nous n'aurons aucune chance d'arrêter leur propagation. Actuellement, chaque acteur a un intérêt écono-

Au début des bombardements des États-Unis et d'Israël sur l'Iran, on serait l'IA qui aurait ciblé une école de filles, causant la mort de 165 enfants et 26 enseignant. L'IA n'est-elle pas aussi un moyen pour les dirigeants d'échapper à leurs responsabilités ?

L'IA entraîne une déresponsabilisation des humains dans le sens où de moins en moins d'humains se retrouvent dans la chaîne de commandement. De fait, un nombre toujours plus réduit de personnes assument les responsabilités, notamment quand un « accident » se produit. « Accident » entre guillemets, car j'observe que ce qui se passe aux États-Unis relève le plus souvent de la logique qu'ils nomment le « vice signalling » : il est désormais bien vu de montrer qu'on est des durs, des méchants, qu'on n'a peur de rien et qu'on fera tout ce qu'il faut pour arriver à notre but.

La décision de déclencher une guerre peut-elle être imposée par l'IA aux gouvernements ?

Nous ne passons pas brutalement d'un état où les humains auraient la maîtrise des conflits à un autre où ils auraient transmis le déclenchement des guerres aux IA. C'est plutôt un continuum, dans le sens où les systèmes d'IA sont partout intégrés. Les dirigeants aux États-

tements ce pays pour ne pas voir la situation s'aggraver... –, alors ils déléguent de plus en plus à ces systèmes. C'est ce qui est en train de se produire.

L'IA n'est-elle pas aussi une arme inouïe aux mains d'un tyran pour faire taire toute opposition à l'intérieur du son propre pays ?

De nombreuses trajectoires sont possibles dans les prochaines années, et je ne suis pas capable de lire le futur. Mais ce qu'il faut considérer, c'est que nos positions publiques sont accessibles dans le système informatique mondial, et donc accessibles aux intelligences artificielles, qui sont entrainées sur internet et qui auront dans leur mémoire une connaissance de la plupart des citoyens de cette planète. Nous nous dirigeons vers une forme d'IA omniscente.

Je m'inquiète d'autres situations plus péjoratives pour neutraliser des opposants. Récemment, des juges de la Cour internationale critique envers les États-Unis y ont été mis sur une liste de terroristes. Cela veut dire que les entreprises basées aux États-Unis n'ont plus le droit de leur fournir leurs services. Résultat, ils ont perdu leur accès à tout leur système informatique, les e-mails, leurs cartes bancaires, etc.

Au vu de votre description des dangers de l'IA, votre association Pause IA ne devrait-elle pas avoir comme objectif de renoncer à l'IA plutôt que de demander un moratoire ?

Nous avons eu cette réflexion au début. Nous nous sommes demandé si nous ne devrions pas plutôt nous appeler StopIA. Nous avons décidé de nous focaliser sur les risques majeurs à assez court terme. C'est une décision stratégique pour avoir le plus de chances de succès. Notre demande est extrêmement raisonnable : elle consiste à mettre en place le développement des systèmes d'intelligence artificielle les plus dangereux, jusqu'à ce qu'on ait obtenu des garanties sur leur sécurité. En pratique, cela revient à un arrêt quasiment total car, aujourd'hui, les labs qui créent ces systèmes sont complètement incapables de les développer de façon fiable, sécurisée ou robuste. Les systèmes d'IA actuels, ce sont des boîtes noires totales. On crée des cerveaux artificiels géants

et on n'a aucune idée de comment cela fonctionne.

En tant qu'expert en « apprentissage machine », n'avez-vous pas l'impression d'avoir contribué à créer le monstre ?

Je n'ai pas partie des gens qui travaillent sur cette technologie, je pense donc que j'ai une part de responsabilité. Mais cela ne sert à rien de s'apitoyer sur soi-même.

Tout ce qui compte pour moi, c'est le moment présent et le futur. Je vois très clairement qu'il faut arrêter cette technologie. Vu la trajectoire actuelle et la rapidité des progrès en IA, il reste probablement moins de trois ans avant une catastrophe mondiale.

Je ne suis pas le seul qui ait vécu ce parcours, mais il y a vraiment un déni total de ce qui est en train de se dérouler chez la plupart de mes collègues. Il faut dire qu'il est assez facile de se retrouver dans ce domaine et de se convaincre qu'en accroissant la puissance de l'IA, on est en train de faire quelque chose pour le bien de l'humanité, que l'on va contribuer à résoudre tous les problèmes du monde.

Que se passera-t-il si nous ne faisons rien ?

Si les choses suivent leur cours actuels, nous passerons alors un point de non-retour. Désormais, les IA commencent à automatiser de plus en plus le travail. Cela peut provoquer une révolution sociale. L'autre péril, c'est la cybersécurité. L'automatisation du piratage informatique se développe à une échelle sans précédent. Globalement, la meilleure façon de le visualiser, c'est de comprendre que nous sommes en train de créer une forme de destruction massive dans le domaine virtuel. Ceci alors que notre société est entièrement basée sur les systèmes d'information. Nous risquons de faire face à d'énormes dégâts. Un autre danger possible est une escalade militaire entre, par exemple, la Chine et les États-Unis, pays qui automatisent de plus en plus leurs systèmes avec des intelligences artificielles. Il est donc urgent de faire comprendre qu'il faut créer une force politique pour arrêter ça.

1 - Sébastien Lecornu, Vers la guerre ?

2 - La vie de Nicolas Guillot, juge fran-

çais de la CPI sous sanctions des États-Unis : « Vous êtes interdits bancaire sur une bonne partie de la planète ». Le Monde, 19 novembre 2023.

« Vu la trajectoire actuelle de l'IA, il reste probablement moins de trois ans avant une catastrophe. »

mique et militaire à construire des IA de plus en plus puissantes, dans le contexte d'une montée des nationalismes. Peut-être qu'il faut un système nouveau, qui prenne une forme différente de ce qui a existé, ou peut-être pas : nous pourrions nous inspirer de ce qui s'est fait contre la prolifération des armes nucléaires.

Cependant, la clé reste d'informer les gens. Tant que la majorité de la population n'est pas consciente des dangers que cette technologie représente, alors il y a peu de chances que nous soyons capables de mettre en place un mécanisme de coordination international. C'est un prérequis.

Unis utilisent des IA pour prendre leurs décisions : leur propre raisonnement commence donc à être influencé par celui des IA avec lesquels ils interagissent. Nous nous dirigeons vers un monde où l'utilisation massive de l'IA entraîne la délégation progressive des décisions, y compris stratégiques, à ces systèmes. Par exemple, si des généraux se rendent compte que l'IA prend de meilleures décisions qu'eux en termes macro-stratégiques, et bien plus rapidement – imaginez une IA capable de réaliser une analyse d'un pays adverse, de présenter un rapport et d'expliquer dans un laps de temps très court qu'il faut attaquer immédia-

