L’appel de 700 personnalités sur les dangers de l’intelligence artificielle

24/01/2015 à 8h00

Les progrès de l’intelligence artificielle sont fulgurants, mais peu de recherches sont effectuées sur les conséquences de cette révolution technologique. C’est l’avertissement émis dans une lettre ouverte par un groupe de chercheurs et d’entrepreneurs. Parmi eux, Stephen Hawking et Elon Musk.



« Le financement de Skynet est voté. Le système se met en place le 4 août 1997, la décision humaine est éliminée de la

> ***Défense Stratégique. Skynet commence à apprendre à vitesse géométrique. il devient autonome le 29 août 1997 à 2h14 du matin, heure de Greenwich ». Terminator 2, 1991.***
>
> ***Et si la réalité rattrapait la fiction ? Sept cents personnalités du monde des sciences et de la technologie appellent, dans une lettre ouverte publiée le 10 janvier, à plus de prudence face à l'essor de l'intelligence artificielle, dont les récents progrès ont été fulgurants.***

À leur tête, l'astrophysicien britannique Stephen Hawking, mondialement connu pour ses recherches conduites sur les trous noirs et la gravité quantique, malgré une paralysie physique quasi totale. La lettre a été publiée par le Future of Life Institute (FLI), un *think tank* créé par Jaan Tallinn, cofondateur de Skype, et dédié à « *protéger l'humanité des risques existentiels* ». Parmi ses autres célèbres signataires, Elon Musk, fondateur de SpaceX et de Tesla Motors ou encore Morgan Freeman, qui occupe l'affiche du film *Transcendance* (2014), justement consacré aux dangers de l'intelligence artificielle.

Progrès exponentiels

« *Il existe désormais un large consensus selon lequel les recherches sur l'intelligence artificielle avancent à un rythme soutenu, et leur impact sur la société va probablement aller en augmentant* », affirment les 700 signataires. Ces effets s'observent déjà autour de nous : reconnaissance vocale, interprétation d'images, traduction automatique et bientôt pilotage des véhicules par eux mêmes. Autant d'innovations qui ont entrainé un cycle vertueux d'investissements massifs, de progrès technologiques et de compétition effrénée entre les entreprises. La question, estiment les signataires, n'est plus de savoir « si », mais « quand » et « qui » aura le premier créé une intelligence artificielle véritable.

> *Les bénéfices d'une telle révolution seraient colossaux, reconnaissent les auteurs, qui entrevoient déjà « l'éradication des maladies et de la pauvreté ». Mais ses risques sont tout aussi grands. Et imprévisibles. Que pourrait créer une intelligence différente ou supérieure à la notre ? Pourrait-elle s'autoaméliorer et dépasser notre compréhension ? Et à quelle vitesse ? Et nous considérera-t-elle comme une menace ? En bref, comment appréhender la singularité ?*

Ce concept a été développé dans les années 1990 par l'écrivain et mathématicien Vernor Vinge. Selon lui, l'apparition d'une intelligence artificielle (vers 2035) créerait un point de rupture tel dans l'histoire de l'humanité et de la recherche scientifique qu'il est simplement impossible d'en supposer les conséquences. L'intelligence artificielle pourrait ainsi être la plus grande invention humaine, mais aussi la dernière.

Assurer le progrès humain

« *Étant donné le grand potentiel de l'intelligence artificielle,* estiment les signataires, *il est important d'étudier comment la société peut profiter de ses bienfaits, mais aussi comment éviter ses pièges.* » Ils évoquent, entre autres, la possibilité d'un chômage de masse dû à la robotisation, l'apparition d'armes autonomes, ou la délégation de choix moraux à des machines.
Les signataires appellent donc la communauté scientifique à réfléchir à des garde-fous éthiques et légaux et proposent des pistes de recherche afin que « *nos systèmes d'intelligence artificielle fassent ce que l'on leur demande* ».
http://www.youtube.com/watch?v=AyaOCaI5sDc

> *Dans une **interview accordée à la BBC** en décembre, Stephen Hawking mettait en garde : « Les formes primitives d'intelligence artificielle que nous avons déjà se sont montrées très utiles. Mais je pense que le développement d'une intelligence artificielle complète pourrait mettre fin*

à la race humaine ».

WE DEMAIN