

Jean-Gabriel Ganascia

Intelligence artificielle vers une domination programmée ?



Le Cavalier Bleu
EDITIONS

Intelligence artificielle

vers une domination programmée ?

Intelligence artificielle

vers une domination programmée ?

Jean-Gabriel Ganascia

2^e édition revue et augmentée

Issues de la tradition ou de l'air du temps, mêlant souvent vrai et faux, les idées reçues sont dans toutes les têtes. Les auteur-e-s les prennent pour point de départ et apportent ici un éclairage distancié et approfondi sur ce que l'on sait ou croit savoir.

Jean-Gabriel Ganascia

Suite à des études de physique et de philosophie, Jean-Gabriel Ganascia s'est spécialisé d'abord en intelligence artificielle et en apprentissage machine puis en modélisation cognitive. Aujourd'hui il est professeur à l'université Pierre et Marie Curie, *EurAI Fellow* et membre de l'Institut universitaire de France. Il a dirigé pendant 12 ans le diplôme d'études approfondies IARFA (Intelligence artificielle, reconnaissance des formes et applications). Il a aussi été chargé de mission à la direction du CNRS avant de diriger le programme de recherches coordonnées « Sciences cognitives » et d'animer le groupement d'intérêt scientifique « Sciences de la cognition ». Actuellement, il pilote l'équipe ACASA (Agents cognitifs et apprentissage symbolique automatique) au sein du laboratoire d'informatique de Paris-VI. Il est aussi le directeur adjoint du Labex OBVIL où il poursuit, en collaboration avec les équipes de littérature de l'université Paris-Sorbonne, des recherches sur le versant littéraire des humanités numériques. Enfin, il est président du comité d'éthique du CNRS (COMETS) et membre de la CERNA (Commission de réflexion sur l'éthique de la recherche dans les sciences du numérique d'Allistène).

Du même auteur

- *L'Âme machine*, Le Seuil, 1990.
- *L'Intelligence artificielle*, Flammarion, 1993.
- *2001, l'odyssée de l'esprit*, Flammarion, 1999.
- *Gédéon, ou les aventures extravagantes d'un expérimentateur en chambre*, Éditions du Pommier, 2002.
- *Sciences cognitives*, Éditions du Pommier, 2006.
- *Voir et pouvoir : qui nous surveille ?*, Éditions du Pommier, 2009.
- *Le Mythe de la Singularité : faut-il craindre l'intelligence artificielle ?*, Le Seuil, 2017.

sommaire

Introduction. 11

L'intelligence artificielle, qui et quoi ?

- « Alan Turing est l'inventeur de l'intelligence artificielle. » . . . 19
- « Désormais des machines passent le test de Turing. » 25
- « L'intelligence artificielle n'est pas une science. » 31
- « L'intelligence artificielle est une idée neuve. » 37
- « Les Japonais sont les champions
de l'intelligence artificielle. » 43
- « La recherche en intelligence artificielle est menée
par les GAFAs. » 49
- « L'intelligence artificielle pallie les défaillances
de notre intelligence. » 55
- « Nous passerons bientôt de l'intelligence artificielle faible
à l'intelligence artificielle forte. » 59

L'intelligence artificielle, comment ça fonctionne ?

- « Il n'y a rien à craindre avec les ordinateurs,
il suffit de les débrancher. » 69
- « L'intelligence artificielle reproduit l'activité
de notre cerveau. » 75
- « L'intelligence artificielle n'est pas naturelle. » 81
- « Les ordinateurs raisonnent de façon binaire. » 85
- « Les ordinateurs ne se trompent jamais. » 89
- « Les ordinateurs sont invincibles aux échecs et au go. » 93
- « Le "*Deep Learning*" révolutionne
l'intelligence artificielle. » 101

Des machines et des hommes

- « Une machine ne peut pas être créative. » 109
- « Les machines n'ont pas d'émotions ni de conscience. » . . . 117
- « Les machines n'ont pas d'intuition. » 123
- « Avec l'intelligence artificielle émotive,
nous confierons bientôt les personnes âgées aux robots. » . . 129
- « Les voitures autonomes sont programmées
pour tuer leurs passagers. » 133
- « Le “*Big Data*” menace la démocratie. » 137
- « Les “robots tueurs” remplaceront bientôt les soldats. » . . . 143
- « Il faut donner des droits aux robots. » 151

L'avenir de l'intelligence artificielle

- « Nous ne sommes pas prêts pour le tsunami technologique
qui advient. » 159
- « L'intelligence artificielle n'a pas tenu ses promesses. » . . . 163
- « Les robots nous mettront tous au chômage. » 169
- « Demain, nous serons les esclaves des machines. » 175
- « Il n'y a pas ou plus de débouchés professionnels
en intelligence artificielle. » 181
- « L'intelligence artificielle constitue un danger existentiel
majeur et inéluctable pour l'humanité. » 187
- « Grâce à l'intelligence artificielle, nous téléchargerons
nos consciences et deviendrons immortels ! » 193
- « La machine est l'avenir de l'homme. » 199

Conclusion. 203

Annexes

- Glossaire 209
- Pour aller plus loin 213

définition

Intelligence artificielle n. f.

La discipline est officiellement née en 1956, au « Dartmouth College », Hanover, New Hampshire aux États-Unis, lors d'une école d'été organisée par quatre chercheurs : John McCarthy, Marvin Minsky, Nicolas Rochester et Claude Shannon. L'année précédente, en 1955, le plus jeune d'entre eux, John McCarthy, avait adressé une demande de subvention en leurs quatre noms à la fondation Rockefeller ; lui, et son comparse Marvin Minsky, étaient âgés de moins de trente ans lorsqu'ils inventèrent le terme d'intelligence artificielle pour frapper les esprits. Et ce terme fit fortune puisque leur projet fut financé et qu'on l'emploie encore pour désigner la discipline informatique qui vise à fabriquer des machines simulant une à une les différentes fonctions de l'intelligence.

Pour John McCarthy et Marvin Minsky comme pour les autres promoteurs de l'école d'été du Dartmouth College, l'intelligence artificielle reposait sur la conjecture selon laquelle toutes les facultés cognitives, en particulier le raisonnement, le calcul, la perception, la mémorisation, voire même la découverte scientifique ou la créativité artistique, pourraient être décrites avec une précision telle qu'il devrait être possible de les reproduire à l'aide d'un ordinateur. Soixante ans plus tard, beaucoup de progrès ont été réalisés dans cette perspective. Désormais, les ordinateurs se sont disséminés partout, dans toutes les activités quotidiennes. Une machine a vaincu, à plusieurs reprises, le champion du monde en titre au jeu d'échecs et même, plus récemment, l'un des meilleurs joueurs au monde au jeu de go ; d'autres démontrent ou aident à démontrer des théorèmes mathématiques ; on construit automatiquement des connaissances à partir de masses immenses de données (*Big Data*). Grâce à cela,

des automates reconnaissent la parole articulée et comprennent des textes écrits en langage naturel ; des voitures se conduisent seules ; des robots font la guerre à la place des hommes ; certains scientifiques cherchent même à vaincre la mort en déterminant les mécanismes du vieillissement... Non seulement, la plupart des dimensions de l'intelligence – sauf peut-être l'humour – font l'objet d'analyses et de reconstructions rationnelles avec des ordinateurs, mais de plus les machines outrepassent nos facultés cognitives dans la plupart des domaines, ce qui fait craindre à certains un risque pour le futur de l'humanité. En dépit des progrès époustouflants enregistrés ces dernières années, l'étude de l'intelligence artificielle repose toujours sur la même conjecture que rien, jusqu'à présent, n'a permis ni de démentir, ni de démontrer irréfutablement.

introduction

Telle une étincelle, l'esprit jaillit parfois de l'entrechoquement de mots contraires. Les ardents et subtils rhétoriciens de l'âge baroque le savaient et en usaient à merveille ; les informaticiens contemporains, nourris de syllogismes et de hamburgers, ne l'ont pas oublié... Les succès qu'alimentèrent les controverses nées autour de l'intelligence artificielle leur ont donné raison : l'accolement des deux mots « intelligence » et « artificielle » fait toujours scandale, à défaut de faire recette.

Et pourtant, à bien y réfléchir, les termes sont tout à fait appropriés : au sens étymologique, l'intelligence artificielle est bien un « artifice », c'est-à-dire un art consommé qui fait illusion en produisant des leurres fabriqués tout exprès pour nous tromper, en faisant accroire que les machines seraient effectivement intelligentes.

Qu'une machine interprète les ordres que nous lui donnons en langue ordinaire, par écrit ou oralement, pour s'exécuter conformément à nos souhaits ; qu'elle pose quelques questions pertinentes avant de suggérer un diagnostic médical ; qu'elle localise d'elle-même les pièces défectueuses d'une voiture atteinte de hoquet ; qu'elle démontre quelques théorèmes mathématiques inédits ; qu'elle reconnaisse des cellules malignes dans une coupe biologique grossie plusieurs centaines de fois au microscope ; qu'elle conduise une voiture au milieu de la circulation ; qu'elle repère, dans le flux des questions posées par les internautes à un moteur de recherche, la manifestation

des symptômes propres à la naissance d'une épidémie ; qu'elle joue au jeu de go et gagne une partie contre l'un des meilleurs joueurs au monde... l'intelligence artificielle est là, et tous s'exclament : « Ô prodige ! ». Serait-ce que les machines, au terme de longs calculs, seraient vraiment devenues intelligentes, et qu'elles posséderaient un esprit, voire que les ingénieurs, à force de les instruire, les auraient dotées d'une conscience ? Point n'est besoin d'aller jusque-là, et d'ailleurs, personne n'est vraiment dupe. Il suffit qu'un ensemble de techniques mises au point par des informaticiens simule des capacités cognitives ordinaires de compréhension du langage naturel, de reconnaissance de la parole, de raisonnement, de résolution de problèmes, de vision, de planification, de jeux, d'apprentissage...

Ces techniques font toutes intervenir des opérations informatiques ordinaires sur des chaînes de caractères, c'est-à-dire sur des mots ou, plus exactement, sur des textes qui symbolisent des sons, des images, des sensations, des états d'esprit... Comme les autres techniques de l'informatique, elles font appel à la logique, aux mathématiques discrètes*, à l'algorithmique, à l'optimisation, à la programmation...

Parfois, le rapprochement des mots « intelligence » et « artificielle » fait resurgir d'un lointain passé de vieux mythes, tels ceux du Golem de Prague, des automates d'Artus de Bretagne ou de l'Ève future de Villiers de l'Isle-Adam, avec leur cortège de légendes et de maléfices. Or, en dépit des craintes que beaucoup nourrissent et des déclarations enflammées d'hommes investis qui abusent de leur autorité, comme Stephen Hawking, Bill Gates ou Elon Musk, il n'en est rien. Les machines fabriquées par l'intelligence artificielle ne possèdent pas, par elles-mêmes, la capacité de prendre le pouvoir sur l'espèce humaine et de la réduire à l'esclavage ;

* Les mots signalés par un astérisque renvoient au glossaire en fin d'ouvrage.

d'ailleurs, pour se prémunir de leurs dangers, il suffit de les débrancher.

De plus, l'intelligence artificielle ne vise aucunement à destituer l'homme de son privilège de penser, pour lui substituer une machine pensante. Elle ne bâtit que des théâtres imaginaires, où se meuvent des personnages chimériques dotés d'aptitudes partielles. Elle n'est qu'une intelligence fabriquée au moyen de techniques informatiques ; autrement dit, elle n'est qu'une « intelligence artificielle »...



Le Golem

La tradition cabalistique juive rapporte l'existence d'une statue d'argile fabriquée par le rabbin Loew, plus connu sous le nom de « Maharl de Prague », vers la fin du ^{xvi}^e siècle. À l'instar des ordinateurs contemporains, cette machine s'animait lorsqu'on passait un message derrière ses dents. Usuellement, elle vaquait aux occupations domestiques quotidiennes, comme un serviteur zélé et assidu.

Beaucoup de légendes ont couru autour de cette statue extraordinaire. Selon l'une d'entre elles, le rabbin Loew aurait oublié, un samedi, jour de prière, d'enlever le message derrière les dents du Golem et celui-ci aurait commencé à s'agiter, à crier et à effrayer tous les voisins pendant que son maître remplissait ses devoirs saints à la synagogue. De retour chez lui, le rabbin Loew aurait détruit son œuvre, de peur qu'elle ne recommence à prendre de fâcheuses initiatives.

Selon une autre légende, sur le front du Golem était écrit le mot *emeth*, qui signifie « vérité » en hébreu ; or, on dit qu'un jour celui-ci aurait pris un couteau pour effacer la première lettre du mot *emeth*, ce qui aurait donné le mot *meth*, soit « mort » en hébreu...

Il résulte de toutes ces mythologies une ambivalence du Golem, qui annonce celle de la technique contemporaine. D'un côté, le rabbin Loew, capable, par son savoir, de fabriquer un objet si perfectionné, fut grandement loué, et même vénéré, au point que le fauteuil sur lequel il s'asseyait est toujours visible dans la synagogue Vieille-Nouvelle de Prague. D'un autre côté, un tel Golem risque parfois d'échapper à ses

maîtres et créateurs, lesquels doivent toujours se garder d'une telle éventualité.

À cet égard, il n'est pas anodin que le père de la cybernétique, Norbert Wiener ait intitulé *God and Golem, Inc.: A Comment on Certain Points Where Cybernetics Impinges on Religion* (« Dieu et le Golem : Un commentaire sur certains points où la cybernétique empiète sur la religion) son dernier ouvrage publié en 1964, l'année de sa mort.

Pour plus d'informations sur le Golem, on pourra lire *Le Golem*, par Moshe Idel, Henri Atlan, et Cyril Aslanof, Le Cerf éditions, 1992.



Étapes du développement de l'intelligence artificielle

Née en 1956, l'intelligence artificielle a connu de nombreuses évolutions au cours de sa courte existence. On peut les résumer en six étapes.

i. Le temps des prophètes

Tout d'abord, dans l'euphorie des origines, les chercheurs se sont laissés aller à des déclarations un peu inconsidérées qu'on leur a beaucoup reprochées par la suite. C'est ainsi qu'en 1958, Herbert Simon, qui deviendra par la suite prix Nobel d'économie, a déclaré que d'ici dix ans les machines seraient championnes du monde aux échecs, si elles n'étaient pas exclues des compétitions internationales. Et que, toujours d'ici dix ans, elles démontreraient des théorèmes originaux en mathématiques, qu'elles composeraient de la musique douée d'une indéniable valeur esthétique, que les théories psychologiques prendraient toutes la forme de programmes informatiques, etc.

ii. Les années sombres

Au milieu des années 1960, les progrès tardèrent ; un enfant de dix ans a battu un ordinateur au jeu d'échecs en 1965 ; un rapport commandé par le sénat américain fit état, en 1966, des limitations intrinsèques de la traduction automatique. L'intelligence artificielle eut alors mauvaise presse, et c'est ainsi que commencèrent quelques années sombres.

iii. L'intelligence artificielle sémantique

Les travaux ne s'interrompirent pas pour autant, mais on axa les recherches dans de nouvelles directions. On s'intéressa à la mémoire, aux mécanismes de compréhension, que l'on chercha à simuler sur un ordinateur, et au rôle de la connaissance dans le raisonnement. C'est ce qui donna naissance aux techniques de représentation des connaissances, qui se développèrent considérablement dans le milieu des années 1970, avec entre autre les réseaux sémantiques et les « cadres de données ». Cela conduisit aussi à développer des systèmes dits experts*, parce qu'ils recourraient au savoir d'hommes de métiers pour reproduire leurs raisonnements. Ces derniers suscitèrent d'énormes espoirs au début des années 1980.

iv. Néo-connexionnisme* et apprentissage machine

Parallèlement à l'essor de l'intelligence artificielle au début des années 1980, les techniques issues de la cybernétique* et du connexionnisme se perfectionnèrent, s'affranchirent de leurs limitations initiales et firent l'objet de multiples formalisations mathématiques. Cela donna naissance à de nombreux développements théoriques puis à des applications industrielles, où les approches se combinèrent pour donner des systèmes hybrides, faisant côtoyer des techniques issues de l'intelligence artificielle, de la recherche opérationnelle, de la cybernétique, de la théorie des systèmes, de la vie artificielle, de l'apprentissage statistique ou de la programmation dynamique.

v. De l'intelligence artificielle à l'informatique animiste...

À partir de la fin des années 1990, on coupla l'intelligence artificielle à la robotique et aux interfaces homme-machine, de façon à produire des agents intelligents qui suggèrent la présence d'un autre. Plus généralement, les réactions des machines usuelles sont désormais calculées de façon à provoquer en nous, à leur contact, l'illusion d'une intelligence les animant, c'est-à-dire d'une âme au sens aristotélicien de « souffle qui anime ». Cette tendance actuelle de l'intelligence artificielle peut éventuellement se caractériser comme une forme d'animisme informatique en cela qu'elle s'emploie à susciter la projection d'un souffle de vie sur les objets quotidiens de notre environnement.

vi. Renaissance de l'intelligence artificielle

Depuis environ 2010, la puissance des machines permet d'exploiter des grandes masses de données (ce que l'on appelle couramment les *Big Data*) avec des techniques d'apprentissage machine qui se fondent sur le recours à de l'apprentissage par renforcement ou à des réseaux de neurones formels, c'est-à-dire à des techniques relativement anciennes que l'on déploie aujourd'hui sur des architectures de dimensions beaucoup plus conséquentes qu'auparavant. Les applications très fructueuses de ces techniques à tous les domaines de l'intelligence artificielle (reconnaissance de la parole, vision, compréhension du langage naturelle, pilotage automatique de voiture, etc.) conduisent à parler d'une renaissance de l'intelligence artificielle qui bouleverse désormais tous les secteurs d'activités (commerce, industrie, banque, assurances, robotique, etc.) en modifiant les métiers, les rôles et les pouvoirs.

L'

INTELLIGENCE
ARTIFICIELLE,
QUI ET QUOI ?

« Alan Turing est l'inventeur de l'intelligence artificielle. »

Dans un article qui a eu une immense influence, Alan Turing a soutenu qu'il était possible de concevoir une expérience prouvant que l'intelligence de l'ordinateur ne pouvait pas être distinguée de celle d'un être humain. Le pari de Turing a éveillé l'ambition de l'intelligence artificielle.

« L'Ordinateur et l'intelligence », site de Michel Volle

Dès 1936, à l'âge de 24 ans, avec les machines dites de Turing, Alan Turing jette les fondements théoriques de l'informatique en établissant un pont entre une formalisation mathématique du calcul et les automates à états finis, autrement dit, les ordinateurs. Il démontre alors qu'une machine très simple est à même de simuler le comportement de n'importe quel ordinateur. Quelques années plus tard, pendant la Seconde Guerre mondiale, il rentre dans les services de renseignement anglais où il emploie ses talents de mathématicien au décryptage des messages ennemis interceptés sur les ondes. Il fait alors appel aux techniques de l'électronique naissante pour fabriquer des calculateurs rapides. Après la guerre, il contribue à la construction d'un des premiers ordinateurs électroniques, puis il poursuit des travaux plus spéculatifs sur les capacités des machines futures à penser, et il préfigure ainsi ce que sera l'intelligence artificielle. Il travaille ensuite sur des simulations informatiques de la croissance des cellules biologiques pour apporter une

contribution originale à la compréhension de la morphogénèse des organismes vivants. Aujourd'hui, ses travaux font toujours l'objet de bien des discussions et alimentent des débats scientifiques enflammés dans les communautés de l'intelligence artificielle et de la biologie. C'est tout particulièrement le cas du test de Turing, qui tente d'apporter une réponse expérimentale à une question souvent rebattue et un peu académique, mais toujours stimulante : « Une machine peut-elle penser ? » Turing imagine une mise en scène, le jeu de l'imitation, dans lequel un interrogateur tente de discerner une femme d'un homme qui travestit ses réponses pour ressembler à une femme. Tout l'attrait tient au dispositif télématique par l'intermédiaire duquel les messages transitent, les personnages ne communiquant que par l'écrit, sans accéder ni à la voix, ni au visage de leurs interlocuteurs. Hormis l'apparence physique, existe-t-il une différence entre l'homme et la femme dans l'ordre de l'intelligence ? Cette question subsiste certainement dans l'esprit d'Alan Turing et le jeu de l'imitation y apportera peut-être une réponse. Mais dans ses articles scientifiques, il la double d'une autre question : existe-t-il, entre l'homme et l'ordinateur, une différence dans l'ordre de l'intelligence, en dépit de leur différence de constitution physique ? Et, pour tenter d'y répondre, il superpose à la première simulation de la femme par l'homme, une seconde simulation en remplaçant, à l'insu de l'interrogateur, l'homme qui imite la femme par un ordinateur qui imite l'homme qui imite la femme. Turing prédit, en 1950, que d'ici 50 ans – c'est-à-dire en l'an 2000 – l'interrogateur n'aura pas plus de 70 % de chance de percer la supercherie en jouant au jeu de l'imitation contre un ordinateur pendant cinq minutes.

philosophie de l'esprit à la lumière des neurosciences, collection « Pour demain », Vrin, 1993.

Un article sur le projet européen *Human Brain Project* de simulation du cerveau à l'aide de techniques d'intelligence artificielle : Lee Gomes, « Facebook AI Director Yann LeCun on His Quest to Unleash Deep Learning and Make Machines Smarter », spectrum.ieee.org.

Un livre sur la modélisation des joueurs d'échecs avec des techniques d'intelligence artificielle : Fernand Gobet, *Les Mémoires d'un joueur d'échecs*, Éditions Universitaires (Fribourg), 1993.

Un livre sur la construction du programme Deep Blue qui l'a emporté sur le champion du monde en titre aux échecs, Garry Kasparov, en 1997 : Hsu, Feng-Hsiung, *Behind Deep Blue: Building the Computer that Defeated the World Chess Champion*, Princeton University Press, 2002.

Enfin, il existe aussi une jolie bande dessinée sur l'histoire de l'intelligence artificielle : Jean-Noël Lafargue, Marion Montaigne, *L'Intelligence artificielle. Fantômes et réalités*, La Petite Bédéthèque des Savoirs, 2016.

Éditeur : Marie-Laurence Dubray
Remerciements de l'éditeur à : Anne-Laure Marsaleix

© Le Cavalier Bleu - 5, avenue de la République - 75011 Paris.
www.lecavalierbleu.com
« idées reçues » est une marque protégée.

Couverture : © Mademoiselle
Imprimé en France par CPI Firmin Didot en avril 2017. 140172
ISBN 979-10-318-0213-8 / Dépôt légal : avril 2017
ISSN 1964-700X