

## BOOK REVIEWS

**ANDLER, Daniel (2023), *Intelligence artificielle, intelligence humaine : la double énigme*, Paris : Gallimard, 432 p.**

Dans cet ouvrage, Daniel Andler s'interroge sur la question de l'intelligence artificielle dans son rapport avec l'intelligence humaine, afin de déterminer dans quelle mesure l'IA se rapproche de l'intelligence humaine pour la reproduire ou même pour la surpasser. La particularité de la réflexion de l'auteur réside dans le fait qu'elle se situe à l'intersection de deux de ses domaines de spécialité, à savoir la philosophie et les sciences cognitives. C'est ce qui rend son approche atypique et originale par rapport à d'autres recherches se rapportant au même champ disciplinaire. Au début de son livre, Andler présente sa problématique en insistant sur le fait suivant : « l'intelligence artificielle constitue une énigme et [...] l'intelligence humaine en est une autre, [...] ces deux énigmes sont étroitement liées, et [...] elles ne sont pas des mystères – de fait je prétends les résoudre » (p.12).

Dans cette perspective, Andler s'efforce de délimiter les différentes acceptations auxquelles réfère le sigle « IA » :

L'IA désigne d'abord un objet qu'on recherche à créer : un système doté d'une certaine propriété, mais l'expression désigne aussi la discipline, et plus largement l'institution qui se donne pour but de concevoir, c'est-à-dire de caractériser et de construire, cet objet. Pour parler de la discipline et de l'institution qui l'abrite – en un mot, de l'entreprise – j'utiliserais le sigle courant IA. Pour ce que l'IA vise et ce qu'elle produit, j'écrirai intelligence artificielle en toutes lettres. En troisième lieu, un système produit par l'IA est souvent appelé une IA. (p. 22)

À la suite de ces trois dénominations, Andler choisit d'utiliser le terme de « système artificiel intelligent (SAI) » pour désigner les outils et les applications produits par l'IA. C'est ainsi qu'il précise que « tout système produit par un laboratoire d'intelligence artificielle est par stipulation un SAI » (p. 23). Après avoir mis l'accent sur les acceptations épistémiques de ce concept, Andler retrace l'évolution historique qu'a connue l'IA depuis son émergence dans les années 50, en passant par l'IA symbolique, le neuro-calcul ou le connexionnisme pour arriver à son état actuel, le Deep Learning.

Dans la deuxième partie de son livre, Andler propose une réflexion sur l'intelligence et l'avenir de l'IA et tente de répondre à la question suivante : peut-on comparer IA et intelligence humaine ? Bien qu'elles soient énigmatiques et divergentes, ces

deux intelligences sont reliées entre elles. C'est ainsi qu'Andler procède à une **COMPARAISON** – un terme qu'il écrit en lettres capitales – en vue d'identifier et d'analyser les différences entre ces deux intelligences. L'auteur affirme toutefois, après avoir parcouru quelques exemples, qu'une telle question « n'a pas de sens ou du moins qu'elle est mal posée » (p. 209). Développant sa réflexion, il compare ces deux intelligences en termes de résolution de problèmes pour mieux identifier les caractéristiques de chacune d'elles :

Nous devons nous contenter de constater que l'IA est actuellement impuissante devant nombre de types de problèmes accessibles à l'Intelligence humaine, inversement, elle résout certains types de problèmes inaccessibles à l'Intelligence humaine du fait que celle-ci ne peut pas traiter les données suffisamment vite ou en nombre suffisant. (p. 219)

À cet égard, Andler avoue qu'il ne détient pas de réponse définitive à la question de l'impuissance de l'IA face à des tâches non algorithmiques et qu'il reste tout de même sceptique à l'idée qu'elle puisse surmonter ses défis actuels pour résoudre tous les types de problèmes. Comme il le note :

Deux voies s'ouvrent à nous : soit chercher s'il existe une ressource essentielle dont on pourrait doter l'IA pour qu'elle puisse résoudre les types de problèmes qui l'arrêtent aujourd'hui, soit nous demander si l'intelligence est en réalité davantage ou autre chose que la capacité de résoudre des problèmes. (p. 220)

En s'interrogeant sur la ressource qui manque à l'IA, Andler creuse davantage sa réflexion afin de mieux comprendre ce qui sépare l'IA de l'intelligence humaine et l'empêche, par conséquent, d'atteindre le stade d'une IA intelligente au sens humain du terme. Il écrit :

Ces questions sont de savoir si les productions de l'IA participent d'une essence particulière ou partagent une propriété spécifique ; si cette essence est identique ou du moins apparentée à son homologue humain ; et enfin si l'écart entre les réalisations de l'IA d'aujourd'hui et les performances de l'intelligence humaine est dû à la difficulté de la tâche technoscientifique et à la lenteur de l'entreprise, ou bien à une différence entre les deux essences ou propriétés caractéristiques de l'intelligence artificielle et de l'intelligence humaine. (p. 213)

Il arrive à l'idée selon laquelle l'intelligence humaine s'étend à d'autres aspects ou, pour reprendre son expression, à *des ingrédients* qui la distinguent et que les SAI n'ont pas pu développer, du moins pour le moment.

Pour le premier ingrédient, la conscience humaine, Andler soutient qu'il est difficile à cerner et à délimiter, d'où l'absence d'une idée claire quant à cette spécificité que seul l'Humain détient. La complexité de la définition de la conscience humaine rend de plus en plus difficile la possibilité d'en doter les SAI, sachant qu'un SAI conscient serait certainement plus proche de l'intelligence humaine.

En ce qui concerne le deuxième ingrédient, le sens, Andler souligne que c'est ce qui manque aux systèmes d'IA: le sens de ce qu'ils font ou qu'ils attribuent à ce qu'ils font reste une capacité qui est propre aux humains. Cette question du sens semble déterminante et même décisive dans certains domaines sensibles tels que l'aéronautique et l'automobile, où les SAI peuvent causer des dégâts importants.

Quant au troisième ingrédient, le sens commun, il semble que les SAI soient dans l'incapacité de comprendre la manière dont des choses qui peuvent nous paraître familières, en tant qu'humains, fonctionnent, notamment lorsqu'elles s'inscrivent dans un contexte plus global. Andler explicite son propos en donnant l'exemple des modèles de SAI de traduction automatique, qui se révèlent parfois incapables d'identifier le terme exact adapté au contexte sémantique de la phrase.

La métacognition constitue le quatrième ingrédient des attributs de l'intelligence humaine et se réfère à la manière dont l'être humain pense et réfléchit sur ses propres processus mentaux et cognitifs. Ces processus permettent à l'être humain d'évoluer, de progresser et de favoriser ainsi son apprentissage.

Le dernier ingrédient renvoie aux sentiments et aux émotions, ce que l'auteur « regroupe sous l'appellation générique des affects auxquels on peut joindre les humains » (p. 236). Si l'Humain dispose d'un corps et donc a automatiquement des affects qui lui permettent de réfléchir et d'analyser les choses autour de lui, ceci n'est pas du tout valable pour les SAI. Andler considère, à ce niveau, que l'IA semble désincarnée. Cette absence du corps implique une absence de Soi. En effet, s'il est possible aujourd'hui de doter les SAI, par exemple un Chatbot, de tâches qui leur donnent la possibilité de reconnaître les affects et de les exprimer pour interagir avec leurs usagers, il est, par contre, difficile d'amener les SAI à éprouver des affects – ce qui pourrait contribuer à améliorer leur capacité à résoudre des problèmes de tous ordres selon l'approche adoptée par Minsky.

Après avoir parcouru les spécificités de l'intelligence humaine par rapport à l'IA, Andler aboutit à la conclusion suivante :

L'intelligence humaine comporte deux grands compartiments, l'un consacré à la résolution de problèmes au sens strict, l'autre à la gestion d'autres tâches intellectuelles ou mentales – appelons-les « générales » par commodité, sachant qu'elles n'ont peut-être en commun que de ne pas consister en problèmes. (p. 243)

Une conclusion qui ne semble pas concluante puisqu'Andler est conduit à creuser davantage dans les travaux de ses prédécesseurs pour définir précisément ce qu'est une intelligence. Cette quête d'une définition plus exhaustive l'amène à faire un tour d'horizon de plusieurs définitions issues d'autres auteurs, tels que Shane Legg et Marcus Hutter. En s'appuyant sur ces derniers, Andler définit l'intelligence comme étant « la capacité de raisonner, de faire des plans, de résoudre des problèmes, de développer une pensée abstraite, de saisir des idées complexes, d'apprendre rapidement et de tirer profit de l'expérience » (p. 251).

Outre ces processus mentaux développés par l'Humain pour comprendre, décider et agir, Andler met l'accent sur un autre élément qui complexifie davantage le concept d'intelligence, à savoir « la normativité », qui consiste à porter un jugement ou une appréciation sur un comportement ou un acte. Il cite ainsi l'exemple de la cruauté et du courage, c'est-à-dire, la capacité de juger d'une personne qu'elle est cruelle et d'un acte qu'il est courageux, qui est à la fois décisive et normative. Cependant, Andler nuance ce concept de « normativité » en précisant qu'il est réparti en deux niveaux : une « normativité forte » et une « normativité faible ». Il considère

que les notions de cruauté et de courage sont fortement normatives étant donné que l'appréciation et le jugement restent subjectifs, tandis que d'autres termes sont normatifs en un sens plus faible.

Il s'avère donc que le concept d'intelligence est « insaisissable », bien qu'il soit d'usage et presque évident, dans la mesure où il n'est pas soumis à une évaluation objective.

À l'issue de la comparaison entre les deux intelligences, humaine et artificielle, Andler aboutit à l'idée qu'une IA forte « prométhéenne », qui serait à l'image de l'intelligence humaine ou qui servirait à s'y substituer, semble vaine. C'est ainsi qu'il emprunte une nouvelle voie en interrogeant, cette fois-ci, les possibilités d'une IA puissante qui aurait pour but de dépasser, voire de surpasser, l'intelligence humaine. Cette super-intelligence n'aurait pas pour objectif de développer uniquement la capacité des SAI à résoudre des problèmes, mais viserait aussi d'autres aspects cognitifs, en empruntant une voie différente de celle qui consiste à vouloir imiter l'intelligence humaine :

Cette super-intelligence artificielle serait alors à l'image de l'intelligence humaine, en beaucoup plus puissante, quelque chose comme l'intelligence d'un super-génie. Mais la super-intelligence pourrait aussi ne pas être humanoïde : sauter la case en question pour accéder à des performances surhumaines. (p. 303)

C'est ainsi qu'Andler conclut sa réflexion autour de l'intelligence artificielle rêvée en précisant que :

Cette ingéniosité humaine qui semble surpasser les capacités scientifiques de l'intelligence artificielle pourrait après tout, contre toute vraisemblance, accoucher d'une idée révolutionnaire pour créer une intelligence artificielle profondément différente de celle que nous connaissons, reposant sur des mécanismes sans rapport direct avec les procédés intellectuels humains. (p. 327-328)

Certes, cette super-intelligence n'est pour le moment qu'une hypothèse, mais Andler se projette dans l'avenir et considère qu'une fois que ce projet sera réalisé, plusieurs questions se poseront quant à savoir ce que la super-intelligence serait capable de faire, notamment si elle arriverait à se développer sur le plan cognitif.

L'auteur termine son livre par une réflexion éthique autour de l'IA – une question qui revêt aujourd'hui une importance cruciale. En effet, l'IA devrait contribuer à ce que son impact sur l'Humanité et la société soit bénéfique en renforçant ses bienfaits et en évitant ses risques. À l'instar d'autres domaines tels que la médecine, les SAI devraient être soumis à des chartes éthiques ayant pour objectif de réduire leurs risques ainsi que leur capacité à prendre le dessus. Ils devraient respecter certaines valeurs éthiques comme la transparence, la justice ou l'équité, l'innocuité (la non malversation), la responsabilité, la protection de la vie privée, le progrès (la bienfaisance), l'autonomie et le respect des obligations générales des populations et de l'environnement. Le respect de ces chartes éthiques par l'IA semble toutefois assez complexe. Prenons l'exemple du principe de l'innocuité. La question qui se pose est la suivante : « Comment garantir qu'un SAI conçu pour accomplir une fonction utile et sans risque d'être dommageable ne soit pas détourné de cette fonction, alors que des dispositifs infiniment plus simples le sont constamment ? » (p. 342). Plus encore,

une IA éthique qui ne ferait que du bien n'est-elle pas une illusion ? Les systèmes développés par l'Humain ne pourraient-ils pas à un moment donné échapper au contrôle de ce dernier ? Andler considère à propos d'« une éthique incapable de faire le mal quelles que soient les mains entre lesquelles elle tombe » que « c'est là un rêve, ou peut-être le nom d'un problème, pas une solution » (p. 342). L'auteur nuance ici ses propos en s'appuyant sur le principe de « modération » qui devrait être adopté par toutes les parties prenantes, et qui consiste à « n'utiliser l'IA que lorsque les risques sont réduits et les bénéfices importants » (p. 361).

Au terme de l'ouvrage, il semble qu'Andler ait interrogé l'IA sous plusieurs angles, tout en interpellant les recherches de ses prédécesseurs ayant également adopté une approche réflexive et critique de l'IA. La particularité de la démarche d'Andler réside dans son interdisciplinarité, à la croisée notamment de la philosophie et des sciences cognitives.

L'auteur fait preuve de générosité et de rigueur scientifique dans son analyse. En témoignent les développements épistémiques et historiques détaillés auxquels il a recours, rappelant non seulement l'histoire de l'IA et ses évolutions actuelles, mais aussi les appréhensions et les craintes quant à ses potentiels progrès dans le futur et le manque de contrôle de l'Humain à son égard – d'où la question de l'éthique qui semble fondamentale aujourd'hui en vue d'accompagner les progrès spectaculaires et accélérés de l'IA durant les années à venir.

Au-delà de l'aspect éthique, il convient de mettre l'accent sur une autre dimension qui est inhérente à l'IA et qui n'est pas évoquée par l'auteur, à savoir l'inclusion. Dans quelle mesure l'IA favoriserait-elle cette dernière et diminuerait-elle les risques de la fracture numérique ? Celle-ci ne va-t-elle pas s'accentuer davantage entre les pays qui déploient des investissements colossaux pour contribuer à faire progresser l'IA et les autres, et les premiers ne seront-ils pas dès lors mieux outillés pour contrôler les seconds, qui ne pourront pas suivre ces évolutions ? Jusqu'à quel niveau l'Humain parviendra-t-il à contrôler l'IA pour qu'elle soit bénéfique et puisse correspondre aux valeurs éthiques qu'il défend ?

Autant de questions qui méritent d'être discutées, à l'ère de l'IA et de ses évolutions incessantes, par des chercheurs, des experts, des philosophes et des futurologues, et qui ouvrent de nouveaux horizons de réflexion allant au-delà de la relation dialectique entre les deux intelligences, humaine et artificielle, pour interroger d'autres aspects, d'ordre socio-économique, politique et humain.

Hanaa Ait Kaikai  
(Université Hassan 1er Settat)

**ESTEBAN, Ángel - APARICIO, Yannelys (2024), «Introducción», en GARRO, E.,  
*Los recuerdos del porvenir*, Madrid: Ediciones Cátedra, 11-125**

El año 2026 marcará el centésimo aniversario del estallido de la Guerra Cristera en México. Al mismo tiempo, se cumplirán ciento diez años del nacimiento de la escritora mexicana Elena Garro. Ambas circunstancias invitan a la lectura de la más importante obra de la autora, *Los recuerdos del porvenir*, que recibió una nueva edición