



HAL
open science

Préface

Gilbert Saporta

► **To cite this version:**

Gilbert Saporta. Préface. Axel Cypel. Voyage au bout de l'IA: Ce qu'il faut savoir sur l'intelligence artificielle, De Boeck Supérieur, pp.5-8, 2023, 9782807360594. hal-04231996

HAL Id: hal-04231996

<https://cnam.hal.science/hal-04231996v1>

Submitted on 7 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Préface

Voici un livre indispensable au moment où nous sommes envahis par d'innombrables articles et ouvrages écrits par des spécialistes autoproclamés de l'IA et où toute méthode informatisée est baptisée abusivement du nom d'IA. Axel Cypel dont le précédent opus¹ avait été remarqué, sait lui de quoi il parle et il le fait avec un bonheur d'expression et un humour qui m'ont réjoui.

Il est rassurant de voir un expert dire sans fard certaines vérités et dégonfler maintes baudruches conçues par un marketing habile. L'utilisation d'expressions ambiguës, dont la première est bien sur celle d'intelligence artificielle, est à l'origine de bien des fantasmes comme « la figure du cerveau en tant que supposé modèle d'inspiration » .

Les nuages de fumées cachent aussi des intérêts économiques bien réels car l'IA est un *business*, et on ne peut que saluer les efforts de l'auteur pour rendre ses lecteurs plus clairvoyants et moins naïfs. J'ai trouvé d'un intérêt tout particulier le chapitre final, empreint d'humanisme, sur les risques de l'IA: on y trouve des réflexions sérieuses sur l'influence réelle ou supposée de l'IA sur l'emploi allant au delà des analyses catastrophistes (destruction d'innombrables emplois) ou d'un optimisme béat (sur la complémentarité homme-machine). Les discours sur l'IA éthique et la constitution de comités d'éthique participent d'un grand bal des hypocrites (*ethics-washing*) pour éviter aux acteurs dominants du numérique d'être soumis à des réglementations², bal auquel sont conviés des chercheurs de bonne foi proposant des dizaines de mesures d'équité des algorithmes, d'ailleurs contradictoires entre elles.

Mais parlons maintenant un peu plus en détail du contenu du livre.

Après un chapitre historique au style alerte, balayant mythologies, héros de cinéma, périodes d'enthousiasme et de déceptions (les *hivers*), et ayant éliminé les lubies de l'IA forte, le chapitre II rappelle que l'on ne peut faire d'IA sans données, établissant une filiation avec la statistique à laquelle je ne peux qu'adhérer.

L'IA actuelle, de l'*après* comme la qualifie Axel Cypel par opposition à celle d'*avant* (celle des systèmes experts encodant des règles), est celle de l'apprentissage machine. A partir de milliers, et même de millions d'exemples, des algorithmes sont entraînés à retrouver la valeur ou la catégorie d'intérêt comme des chats dans des images. Ces algorithmes sont souvent qualifiés de boîtes noires, car leur fonctionnement est trop complexe pour être compris par le commun des mortels : ils dépendent souvent de milliers de paramètres dont les valeurs sont optimisées pendant la phase dite d'apprentissage. Mais l'essentiel n'est pas de prédire correctement ce que l'on connaît déjà mais bien plutôt de prédire de nouveaux cas, ce que les spécialistes appellent la *généralisation*, et qui est plus difficile comme le rappelle cet aphorisme attribué à Pierre Dac : « Les prévisions sont difficiles, surtout lorsqu'elles concernent l'avenir ».

Le chapitre III décrit dans un langage simple mais précis, les principaux algorithmes d'apprentissage, permettant de comprendre sur quoi reposent leurs succès et comment on évalue leurs performances. L'auteur insiste à juste titre sur le caractère fondamentalement probabiliste des prévisions, trop souvent oublié.

La réussite incontestable des algorithmes d'IA repose sur de gigantesques bases de données ; les géants de la tech, les fameux GAFAM, ont ainsi des avantages structurels grâce aux traces et aux données personnelles que nous laissons plus ou moins consciemment ainsi que les gouvernements autoritaires de pays très peuplés.

Le nombre et la qualité des données sont des ingrédients indispensables pour que les algorithmes fonctionnent bien. Le vieil adage « *garbage in, garbage out* » reste toujours vrai et un travail souvent considérable est indispensable pour passer des données brutes à des données informatives. Pour

¹ Au cœur de l'intelligence artificielle: Des algorithmes à l'IA forte, De Boeck, 2020

² On consultera avec profit le chapitre 11 « The battle for data ethics » de Wiggins, C., & Jones, M. L. (2023) *How Data Happened: A History from the Age of Reason to the Age of Algorithms*. W. W. Norton and Company

pouvoir reconnaître des chats dans des images, encore faut-il que l'on dispose au départ de fichiers d'images classées comme des images de chats. Or qui effectue cette *labélisation*, sinon des humains ? Certains payés des misères comme ceux que l'on appelle les travailleurs du clic, d'autres bénévolement car les utilisateurs travaillent souvent gratuitement et inconsciemment pour le plus grand profit des entreprises du net en notant, annotant, commentant. Voilà un paradoxe méconnu de l'IA qui est de reposer sur de l'intelligence pas artificielle du tout.

Le chapitre IV est consacré à ChatGPT, l'agent conversationnel introduit en novembre 2022 dont tout le monde parle. ChatGPT est l'emblème d'une nouvelle catégorie d'IA : l'IA générative qui produit des textes ou des images en réponse à des requêtes. Qui n'a pas été bluffé par la qualité des textes produits ? Le revers de la médaille est que ces textes comportent souvent des erreurs ou des absurdités, dénommées judicieusement *hallucinations*. ChatGPT n'est pas un moteur de recherche, ne cite pas ses sources, et pour cause : « ChatGPT n'a pas été nourri avec des connaissances, seulement avec des textes » et procède par une approche statistique produisant des textes probables, ne prétendant pas à la vérité. Comme le dit bien l'auteur « ce phénomène d'hallucination, s'il est recherché dans les IA génératives d'images en tant que garant d'inventivité ou d'originalité, s'avère assez gênant pour les textes ».

Le chapitre décrypte le fonctionnement de ces IA génératives en dévoilant entre autres un de leurs secrets : *l'apprentissage auto supervisé*, qui consiste, telle une variante des exercices à trous, à deviner des mots cachés.

On retrouve ici le fait que l'apprentissage machine n'a que peu à voir avec l'apprentissage humain: on a souvent dit qu'un enfant n'avait pas besoin de voir des centaines de chats, mais juste quelques uns, pour pouvoir en identifier un autre. De manière similaire, l'auteur note: « nous ne nous entraînons pas à comprendre autrui en masquant des mots dans les phrases et en ayant, de plus, connaissance des mots qui viendront après celui qui est en train d'être énoncé ». Comme l'annonçait l'introduction « Rien n'est donc plus éloigné de l'intelligence que l'intelligence artificielle » .

Au delà des succès médiatiques, Axel Cypel nous entraîne au chapitre V du côté des réussites en entreprise et des conditions de ces réussites.

Le chapitre VI décrit bien les limites de l'IA et leurs natures diverses. J'y ai retrouvé un écho des débats suscités par l'article de Leo Breiman³ il y a déjà plus de 20 ans sur la distinction entre modèles génératifs (causaux ou explicatifs) et modèles prédictifs ainsi que par la tribune polémique de Chris Anderson⁴ pronostiquant la fin de la théorie puisque l'on dispose de données massives.

La complexité de certains algorithmes les rend vulnérables à des modifications infimes des données : changer la couleur d'un seul pixel peut modifier une prévision. Un retour à la simplicité est-il possible, peut-on aller vers une IA explicable ou interprétable? Voilà bien des sujets de réflexion.

La promesse de donner des clés de lecture est donc bien tenue et je souhaite au lecteur de trouver autant d'intérêt à lire ce livre que j'en ai eu. Il est clair qu'un tel ouvrage n'aurait pas pu être écrit par une IA générative !

Paris, le 31 juillet 2023

Gilbert Saporta

Professeur émérite de statistique appliquée

Conservatoire national des arts et métiers, Paris

³ Breiman, L. (2001). Statistical Modeling: The Two Cultures, *Statistical Science*, 16, 3, 199–231

⁴ Anderson, C. (2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete, *Wired magazine* <http://www.wired.com/2008/06/pb-theory/>