

A X E L C Y P E L

PRÉFACE DE GILBERT SAPORTA

# VOYAGE AU BOUT DE L'IA

**CE QU'IL FAUT SAVOIR**  
SUR L'INTELLIGENCE ARTIFICIELLE

deboeck **B**  
SUPÉRIEUR

The bottom half of the cover features a vibrant, futuristic cityscape at night. The scene is dominated by a dense array of vertical light trails in shades of blue, cyan, and magenta, which appear to be data streams or digital connections. These trails are set against a backdrop of a city skyline with illuminated skyscrapers. The overall aesthetic is high-tech and digital, suggesting themes of artificial intelligence and data processing.



Axel Cypel

# Voyage au bout de l'IA

*Ce qu'il faut savoir sur l'intelligence artificielle*

## Du même auteur :

*Au cœur de l'intelligence artificielle. Des algorithmes à l'IA forte*, De Boeck Supérieur, 2020.

Pour toute information sur notre fonds et les nouveautés dans votre domaine de spécialisation, consultez notre site web :  
**[www.deboecksuperieur.com](http://www.deboecksuperieur.com)**

© De Boeck Supérieur s.a., 2023  
Rue du Bosquet 7, B - 1348 Louvain-la-Neuve

*Il est interdit, sauf accord préalable et écrit de l'éditeur, de reproduire (notamment par photocopie) partiellement ou totalement le présent ouvrage, de le stocker dans une banque de données ou de le communiquer au public, sous quelque forme et de quelque manière que ce soit.*

Dépôt légal :  
Bibliothèque Nationale, Paris : octobre 2023  
Bibliothèque royale de Belgique, Bruxelles : 2023/13647/147

ISBN 978-2-8073-6059-4

# Sommaire

<b>Préface</b>	<b>5</b>
<b>Introduction</b>	<b>9</b>
<b>I Petite histoire subjective et définition de l'IA</b>	<b>13</b>
1 Une dénomination mythique . . . . .	13
2 Mythologies: le pépin du fruit défendu . . . . .	14
3 <i>Winter is coming</i> : apprendre à crier au loup . . . . .	17
4 L'IA e(s)t ses doubles . . . . .	18
5 L'IA c'est... Tentative de définition . . . . .	22
6 Les promesses de l'IA . . . . .	25
7 Un changement de paradigme . . . . .	27
8 IA faible/IA forte . . . . .	29
<b>II Comment fait-on de l'IA ?</b>	<b>31</b>
1 <i>Data teaser</i> . . . . .	31
2 <i>Datascience</i> : les ingrédients . . . . .	31
3 Les acteurs de l'IA . . . . .	32
4 Les outils de l'IA . . . . .	34
5 « La <i>data</i> » dans tous ses états . . . . .	37
6 <i>Big data</i> ? . . . . .	40
7 Le moteur mathématique du ML . . . . .	41
8 Retour sur le changement de paradigme . . . . .	43
9 À comprendre avec l'IA moderne: l'inférence . . . . .	44
<b>III Focus sur l'apprentissage machine</b>	<b>47</b>
1 Situer l'apprentissage dans l'IA . . . . .	47
2 L'apprentissage supervisé . . . . .	48
3 L'apprentissage non supervisé . . . . .	54
4 L'apprentissage par renforcement . . . . .	56
5 Catalogue d'algorithmes . . . . .	58

## Voyage au bout de l'IA

6	La généralisation . . . . .	66
7	Évaluation des modèles . . . . .	70
8	Une application pas si facile que ça . . . . .	74
9	D'où vient l'intelligence ? . . . . .	76
<b>IV</b>	<b>Concevoir ChatGPT</b>	<b>79</b>
1	Le NLP: un peu d'Histoire . . . . .	79
2	Apprentissage d'un modèle de langage par RN . . . . .	86
3	La spécificité de ChatGPT . . . . .	90
4	Gloire et déboires de ChatGPT . . . . .	91
5	Génération « IA génératives » . . . . .	105
<b>V</b>	<b>L'IA opérationnelle</b>	<b>109</b>
1	Les succès médiatiques de l'IA . . . . .	109
2	Les applications quotidiennes de l'IA . . . . .	112
3	Implémenter l'IA en entreprise . . . . .	117
<b>VI</b>	<b>Les limites de l'IA</b>	<b>131</b>
1	Les limites de l'apprentissage machine . . . . .	131
2	Les limites opérationnelles . . . . .	136
3	Les limites héritées de la Statistique . . . . .	141
4	Les limites de la modélisation . . . . .	146
<b>VII</b>	<b>Les risques de l'IA et leur traitement</b>	<b>151</b>
1	Les capitalisations de l'IA . . . . .	151
2	Inquiétudes sur l'emploi . . . . .	155
3	Surveillance généralisée . . . . .	164
4	Déresponsabilisation . . . . .	166
5	Le transhumanisme et son IA costaude . . . . .	169
6	Réguler l'IA . . . . .	174
7	De l'éthique en IA . . . . .	176
	<b>Conclusion</b>	<b>185</b>
	<b>Remerciements</b>	<b>191</b>

# Préface

Voici un livre indispensable alors que 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 opus<sup>1</sup> 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 sûr 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 régulations<sup>2</sup>, 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

---

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

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

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 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 Breiman<sup>3</sup> il y a déjà plus de 20 ans sur la distinction entre modèles génératifs (causaux ou

---

3. L. Breiman, « Statistical Modeling: The Two Cultures », *Statistical Science*, 16, 3, 199-231, 2001.

explicatifs) et modèles prédictifs ainsi que par la tribune polémique de Chris Anderson<sup>4</sup> 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*

---

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

# Introduction

**Qui est-elle ?** L'intelligence artificielle vit dans un monde empli de paradoxes qui saturent l'agora. Du côté des poncifs : « l'IA est partout ! » Mais être de tous les lieux, c'est finalement n'en être d'aucun en particulier. Ou encore : « en IA, tout va très vite ! » Cela donne certes l'impression de faire partie du mouvement, mais fournit une raison de plus pour s'accorder le temps d'en comprendre les fondamentaux. Plus déroutant : avec la reconnaissance faciale, l'IA est censée pouvoir débusquer, dans un support vidéo quelconque, un individu à partir d'un simple cliché, mais elle produit dans le même temps des images si parfaites (les fameux *deep fakes*) que ces contrefaçons ne se laissent pas distinguer d'une image réelle. Identification et falsification partagent mêmes techniques et supports. Que l'IA reconnaît-elle donc ?

Fréquemment annoncée comme une révolution, l'IA resterait cependant à l'informatique ce que la modernité serait à l'époque, n'occasionnant au final pas de bifurcation notable. Du déjà-vu en mesure de vous mettre au chômage, en somme. Les machines de demain deviendraient plus intelligentes que leurs humains de propriétaires, qui en resteraient pourtant les maîtres... Toutes ces élégantes antithèses ne forment que des instances particulières du père des paradoxes qu'il convient maintenant d'introduire : l'IA est présentée comme vecteur de progrès majeurs justifiant ainsi la nécessité d'une adoption très large, en même temps qu'elle est dite porteuse de risques immenses, au point qu'il serait souhaitable d'en circonvenir la propagation. Que faire alors ?

Tout cela sans évoquer la dénomination même de la discipline, de laquelle s'exhale un parfum de contradiction propre à stimuler une attitude spéculative qui masque peut-être les vraies questions : comment fait-elle pour apprendre ? Qu'est-ce qui distingue l'IA des autres technologies dans cette prétendue compétition avec le vivant ? Faut-il, enfin, mettre sur le compte d'une dérive narrative l'usurpation, et donc la déchéance, du rang de l'espèce humaine, ou bien accrédi-ter la possibilité réelle, et par là tragique, d'une concurrence directe dans l'expression de ce qui, entre autres, nous fait Hommes, à savoir apprentissage et réflexion ?

**De l'artifice** L'intelligence artificielle est dorénavant présente dans de nombreuses applications grand public, mais, en tant qu'ensemble de techniques mathématiques et informatiques, elle reste difficile à appréhender pour ce même public. Comment aller au-delà des fantasmes cinématographiques et saisir quel en sera l'impact sur la manière dont pourraient être organisées nos vies, au gré de son intervention? Car plus la technique s'améliore, plus elle devient invisible; et il est vrai qu'une bonne IA est une IA qui ne se voit pas, c'est-à-dire qui se trouve insérée dans un processus faisant passer son action comme allant de soi.

Cette incorporation feutrée offre prise à des présentations où l'IA serait une forme de magie, ce qui octroie à l'exposant ayant choisi cette approche le double avantage de n'avoir pas à expliquer son fonctionnement et d'éviter l'embarras de devoir se justifier lorsque l'application ressemble trop à de la pseudo-science. Et puisque l'on aborde ici les communications équivoques en IA, ajoutons qu'il est une arnaque commune consistant à faire passer une modélisation souvent efficace en tant qu'entité compréhensible et agissante. En maintes occasions le sensationnalisme est ainsi de sortie, quand il serait plus avisé d'offrir au public des clefs de lecture.

**D'où vient-elle ?** Le principe premier de l'IA réside dans le numérique, entendu comme l'ensemble des fonctionnalités permises par le traitement informatique, c'est-à-dire le traitement automatisé de l'information. On pourrait dire poétiquement que le numérique possède la force de l'eau, c'est-à-dire une capacité d'infiltration et d'imprégnation. De sorte que, dans une société reposant toujours davantage sur l'accès internet, la connectivité permise entre les objets et la captation puis l'échange de données, le numérique devient de plus en plus difficile à contrer. En sus, il ne vient pas seul, mais accompagné de son bras armé qu'il permet tout autant qu'il appelle, l'immanquable intelligence artificielle.

L'industrie du numérique, représentée par des sociétés dont la réussite à la fois insolente et remarquable les a constituées en quasi-monopoles de fait, a tout intérêt à faire accroire qu'aucun pan du réel n'échappe au digital, au numérique, partant aux données. Nous verrons qu'un certain nombre de limitations entachent ce *credo*.

**Une ressemblance douteuse** Il serait d'un intérêt indéniable, si l'on entend traiter d'intelligence artificielle, de disposer de quelques points d'accroche sur son pendant « naturel ». L'intelligence, cette qualité que tout un chacun assurera avoir en partage, se laisse pourtant malaisément définir. Au nombre de ses caractéristiques, et sans prétention à l'exhaustivité aucune, produisons toutefois deux propositions. L'intelligence consiste à faire des liens entre les objets (et aussi les concepts), et ce de manière autonome. C'est le niveau étymologique, que nous pourrions raffiner en disant que l'intelligence est la faculté d'expliquer le monde « causalement », c'est-à-dire en repérant les chaînes de causalité en son sein. Mais qu'est-ce que la causalité? Quand – et pourquoi – peut-on dire avec certitude qu'un événement est la conséquence d'un autre?

De fait, l'intelligence, moteur de la réflexion, est dynamique et adaptative; elle se remet en question en permanence. Rien n'est donc plus éloigné de l'intelligence que l'intelligence artificielle! Mais alors pourquoi ce vocable? Essentiellement, en vue de servir le nécessaire marketing de la science qui n'échappe pas à la loi commune: pour rentrer dans une entreprise ou dans un foyer, il faut qu'elle se vende.

**L'enjeu actuel** Nous l'avons dit plus haut, ce qui paraît le plus urgent, à l'endroit de l'intelligence artificielle, est de donner des clefs de lecture afin que chacun, qu'importe sa proximité au domaine, puisse se saisir de cette révolution technologique en cours, quelle qu'en soit la dénomination utilisée, satisfaisante ou pas. Comprendre les enjeux, décoder les intérêts, en un mot voir au-delà du mirage à nos yeux présenté, voici les éléments nécessaires à une correcte interprétation des techniques et de leurs implications. Afin de doter chacun de ces capacités d'analyse, des bases sont nécessaires pour comprendre de quoi il retourne et si les discours entendus au sujet de l'IA sont appropriés. C'est souvent là que le bas blesse, car il n'est pas dans l'intérêt de ceux que conforte une position dominante chèrement acquise d'instruire leurs interlocuteurs. Les forces en présence ne poussent guère à un éclaircissement de la situation.

Ce livre se veut une réponse à cette problématique: conférer à son lecteur suffisamment de compréhension pour qu'il soit en mesure de former sa propre position sur ce sujet d'actualité dont la jeunesse semble se régénérer d'elle-même, un peu aidée en cela néanmoins par une R&D foisonnante et

un narratif marketing qui en font un thème continuellement commenté. Car cette technologie, qui s'est répandue partout et dont l'accélération va sans aucun doute s'amplifier au cours de la décennie à venir, demeure en recherche d'une acceptabilité sociale propre à réunir, pour le compte de ses détenteurs, les conditions pérennes d'une exploitation commerciale massive dont il serait erroné de croire qu'elle n'a pas déjà débuté.

**Plan de route** Il est ainsi temps de donner une définition de l'IA en brossant, au départ, un panorama à grands traits qui nous servira de carte pour la suite des opérations (chapitre I). Prenant appui sur des questions très concrètes, nous verrons quels ingrédients l'IA réunit pour exister et en quoi elle se distingue de l'informatique traditionnelle (chapitre II). Nous serons alors munis des notions qui permettront d'entrer plus avant dans les explications techniques que nous avons souhaité maintenir en vulgarisant les grandes notions de cette discipline issue des mathématiques appliquées (chapitre III). Cela nous permettra d'aborder rien de moins que la conception d'une IA qui a vraiment su faire parler d'elle, le *chatbot* au nom infâme, mais à la redoutable maîtrise syntaxique, « ChatGPT » (chapitre IV). Et puisque nous aurons construit (en pensée) un outil de pointe de l'intelligence artificielle, il sera pertinent de revenir aux usages courants puis à une vision opérationnelle de l'IA, telle qu'on peut la pratiquer en entreprise (chapitre V). Rattrapés par les problèmes qu'elle ne manque pas de susciter, un passage par les rapides des limites de l'IA (chapitre VI) nous ouvrira la voie vers les terres australes et risquées d'une certaine conception de l'IA, au pinacle desquelles nous aurons beau jeu de nous questionner sur l'issue éthique à laquelle elle pourra prétendre (chapitre VII), avant d'aboutir.

# I

## Petite histoire subjective et définition de l'IA

### 1 Une dénomination mythique

L'intelligence artificielle a dorénavant suffisamment d'années d'existence pour s'être permis le luxe de se créer sa propre légende, légende qui n'est pas sans rappeler une autre bien plus ancienne: celle associée à la ville de Rome. L'iconographie classique évoque le bronze montrant les petits Rémus et Romulus, ses mythiques fondateurs, allaités par la Louve. « Louve », en latin, se dit « *lupa* », et ce mot a, dans ce langage, une double signification puisqu'il veut dire également « prostituée »<sup>1</sup>. Bref, de dieux vivants, les valeureux fondateurs de Rome pourraient tout aussi bien être des fils de p... Comme dans tous les mythes, il y a plusieurs interprétations, et c'est précisément cette ambivalence qui fait les mythes.

À cette fondation légendaire de Rome répond celle, moins poétique mais non moins imaginative, de l'IA par le séminaire d'été de Dartmouth de 1956 dont le legs pour la postérité fixe, sur argentique, les pères fondateurs. On y retrouve des gens très bien, pour n'en citer qu'un: Claude Shannon, le grand théoricien du traitement du signal. Comme d'habitude, c'est-à-dire à l'époque comme de nos jours, ces brillants esprits cherchent à attirer des pourvoyeurs de fonds pour leurs recherches. Tout bon universitaire vous le confirmera: l'essentiel du temps consacré à la recherche se passe à rechercher des financements. La différence avec la quête des œufs à Pâques, c'est qu'ici la chasse dure toute l'année.

Quand on cherche des financements, mieux vaut avoir une belle histoire à raconter et donc un nom de projet bien marketé. Ainsi, plutôt que de vendre à vos industriels que vous allez faire de la recherche algorithmique en

---

1. On trouve trace de cette étymologie dans « lupanar ».

mathématiques sur la théorie des variétés à 42 dimensions, mieux vaut dire qu'avec la technologie naissante, vous allez recréer l'intelligence: vous allez concevoir une *intelligence artificielle*. Et là, tous les Rockefeller vous signent le chèque dont vous aviez besoin. Pour rester un minimum crédible, il vous faudra tout de même veiller à afficher des ambitions, garantes de la rentabilité de l'argent versé: on émulera un être humain d'ici 20 ans. Disons 30... Et ce qui est amusant, c'est que 70 ans après, rien n'a changé: il se trouve toujours des gens pour prédire l'IA forte dans 20 ans. Disons 30...!

## 2 Mythologies : le pépin du fruit défendu

Les plus versés dans la mythologie pourront peut-être faire remonter la fondation de Rome bien avant les compères Rémus et Romulus. Il nous faut revenir à la guerre de Troie, ce conflit qui eut lieu entre les Troyens et les Grecs suite à l'enlèvement de la belle Hélène par le berger Pâris. Mais pourquoi donc est-il parti avec elle? Tout remonte à l'histoire d'une pomme, fruit jouant un rôle crucial jusque dans la genèse du sujet de notre voyage. Et puisque l'on vient de parler de fondation, si l'on veut que ce voyage soit bien fondé et nous conduite sans danger à son terme, mieux vaut partir du bon bout!

**Au mont I(D)A, trois déesses...** Nous commencerons curieusement avec quelqu'un dont on évite d'ordinaire la compagnie: Eris, la déesse de la discorde, n'avait déjà à l'époque pas bonne presse et, lors de fêtes divines ou de banquets royaux, on omettait généralement de lui envoyer un faire-part, ce qui avait pour effet de la rendre plus rancunière encore. Elle eut donc l'idée d'un stratagème. Un jour que se célébrait un mariage important où à peu près tous les dieux de l'Olympe étaient réunis, elle se pointe à la fête, à laquelle on ne l'avait, bien entendu, encore une fois pas invitée, et balance dans la salle du festin une pomme en or – comme il en pousse dans le jardin des Hespérides – avec une petite étiquette dessus, sauf que là il n'y avait pas marqué « pink lady », mais « à la plus belle ».

C'est la pomme de discorde. Effet immédiat: une grande pagaille, car toutes les déesses veulent la récupérer, et personne n'est vraiment d'accord sur celle qui peut véritablement y prétendre d'après l'*étiquette*<sup>2</sup>. Ayant une fois de plus réussi à semer la zizanie, Eris, là-dessus, se retire satisfaite. Mais

---

2. Nous verrons plus tard que l'IA est assez souvent une affaire d'étiquettes.

## I. Petite histoire subjective et définition de l'IA

imaginait-elle elle-même les conséquences de son acte? Quant aux déesses qui sont restées, comme on n'est pas fous chez les dieux, le choix se réduit finalement à un « *top trois* »: Héra, l'épouse de Zeus, Athéna, sa fille (déesse de la stratégie militaire et de la sagesse, soit en un mot, déesse de l'*intelligence* – n'est-elle pas sortie toute casquée du cerveau de son père?), et Aphrodite, la déesse de l'amour, qu'on ne présente plus. La concurrence est donc très rude. Ce sont des déesses, elles ne sont pas belles, elles sont infiniment belles! Alors comment départager ces infinis entre eux (en absence de la théorie des grands cardinaux)?

Elles vont voir le patron, Zeus, qui, très sagement, préfère ne pas trop se mêler de cette affaire. En revanche, il leur indique un homme au goût très sûr en la matière, un Prince berger nommé Pâris, passant des jours paisibles sur le mont Ida, non loin de Troie. Ni une ni deux, les trois déesses se téléportent devant lui, et l'on peut imaginer la tête qu'il dut faire quand on lui expliqua qu'il devrait donner la pomme d'or, qu'il se vit prestement remettre, à celle qu'il jugeait la plus belle des trois.

Il n'eut pas vraiment le temps de réfléchir, la stratégie de chaque déesse ayant consisté immédiatement à le soudoyer pour assurer sa propre élection. C'est ainsi que Pâris se voit promettre les biens les plus fous: Héra lui propose la souveraineté sur le monde connu, Athéna la victoire écrasante de Troie sur la Grèce, et Aphrodite, l'amour de la plus belle femme vivante sur Terre à ce moment. Pâris lui remet la pomme. La plus belle femme vivante sur Terre à ce moment-là a pour nom Hélène, et c'est l'épouse de Ménélas, le roi de Sparte. Pâris la lui enlève et, comme Hélène c'était vraiment quelqu'un, les princes de Grèce, unis par serment, partent en guerre contre Troie.

Ulysse, le cheval de Troie<sup>3</sup>, la prise de la ville, la débandade, le départ précipité d'Enée avec son père Anchise sur les épaules, l'accostage à Carthage chez la reine Didon, ses amours malheureux avec le bel Enée, voué par Zeus à fonder la future Rome et qui donc la plante là, et le suicide de la reine... Troie est perdue, vive Rome!

**Problèmes de mœurs (entre autres) au xx<sup>e</sup>** Retour à notre époque où se poursuit notre parallèle. À l'évocation des pères fondateurs de l'IA,

---

3. Où l'on constate encore l'influence de ces mythes: un cheval de Troie est aussi un virus informatique, rentrant camouflé dans la place.

comment ne pas citer le nom de Turing, éminent mathématicien britannique, décodeur d'Enigma et qui s'est intéressé, dans un article dorénavant célèbre (« Computing machinery and intelligence »), au problème des machines pensantes ?

Enigma, vous le savez peut-être, est le nom de cet abominable engin de cryptographie utilisé par les Allemands pendant la Seconde Guerre mondiale pour s'envoyer des messages codés. Si cet engin était infâme, c'est qu'il était très efficace : décoder un message à la main sans avoir la clef prenait beaucoup trop de temps, il y avait des dizaines de millions de possibilités. Turing inventa alors une sorte d'ordinateur (« la bombe ») qui permettait de casser le code en moins d'une journée (les clefs changeaient tous les jours et un message militaire vieux d'il y a un mois ne vous renseigne guère sur les intentions actuelles de l'ennemi). Cela fonctionna aussi, comme souvent en cryptographie, par l'exploitation de failles humaines : carnets de codes dérobés dans des sous-marins – il faut les codes à l'avance quand on reste sous l'eau longtemps –, ou encore certaines régularités, genre le « Heil Hitler » qui ponctuait systématiquement les messages nazis. Tout cela est relaté dans le film *Imitation Game* où l'acteur Benedict Cumberbatch joue Alan Turing.

Alan Turing appartient à une génération de logiciens exceptionnels qui ont sévi à la même époque que lui (Gödel, Church, Kleene). C'est un théoricien de l'informatique<sup>4</sup> qui a formalisé la notion de calcul en créant un algorithme appelé « la machine de Turing », une machine de papier lui servant à prouver des théorèmes sur le calcul et qui est maintenant utilisée comme étalon pour définir la complexité algorithmique. Il est aussi l'auteur du fameux test de Turing, censé servir de critère pour repérer une vraie IA. Celle-ci passe le test si vous ne faites pas la différence avec un humain dans les réponses qu'elle vous fournit. Sans doute dépassé, il nous pousse néanmoins à nous interroger sur la question de l'illusion et de ses implications.

De manière amusante, la figure de Turing se retrouve défendue par deux chapelles rigoureusement opposées : la première, naïve, qui brandit généreusement des succès au test de Turing, est celle d'une intelligence artificielle dont la prétention est de pouvoir répondre à tout, l'ordinateur énonçant la vérité

---

4. Indécidabilité de l'arrêt des programmes...

sur n'importe quelle question ; alors qu'aux yeux de la seconde, bien entendu dans le vrai, Turing apporte à cela même une réfutation catégorique.

De manière moins amusante, comme ils disent dans *Certains l'aiment chaud*, « nobody's perfect » : Turing était *gay*. Et ce n'était pas bien vu dans la prude Angleterre. Il est donc jugé et condamné à la peine de son choix : emprisonnement ou castration chimique. Il opte pour cette dernière. En gros, il doit prendre des comprimés – on n'ose pas appeler ça des « médicaments » – qui sont probablement des hormones censées l'aider à contrôler ses pulsions... Et donc cela ne doit pas favoriser le bien-être. Il se suicidera un an ou deux plus tard, je crois, la légende voulant qu'il ait croqué, à dessein, dans une pomme empoisonnée au cyanure : *la pomme*.

**Le symbole récurrent** Quand on y pense, c'est incroyable comme ce fruit si goûteux se retrouve mêlé à des imbroglios le dépassant de beaucoup. Outre la guerre de Troie, on lui doit, de manière plus connue, la chute biblique : il avait pourtant été indiqué formellement au couple primitif (et premier) de ne pas croquer de fruit poussant sur l'Arbre du savoir, fruit que la tradition a associé à la pomme. Face à l'éviction du jardin d'Eden, à la défaite de Troie et à la fin tragique de Turing, le pouvoir supposé de la pomme, de reconnaître beauté et intelligence à son détenteur, pourrait-il favoriser l'IA ? L'intelligence artificielle peut-elle déceimment prétendre au savoir ? Au pouvoir ?

### 3 *Winter is coming* : apprendre à crier au loup

La jeune histoire de l'IA commence donc sur les chapeaux de roue avec des débuts quelque peu délirants (puisqu'à l'époque on voulait reproduire les fonctions cognitives humaines sur des calculateurs pas vraiment performants), mais tout de même prometteurs, car les premiers ordinateurs sont en train d'être construits et l'on en pressent déjà tout le potentiel. Cela dit, eu égard aux fins, les moyens ne suivent pas vraiment et les limites scientifiques se font sournoisement sentir. L'IA arrête de faire parler d'elle pour hiberner un peu : c'est le premier hiver de l'IA.

Résurgence dans les années 80, à l'ombre non pas des jeunes filles en fleurs, mais des systèmes experts : l'IA devient de l'informatique intelligente, car elle suit les règles intelligentes que des experts (humains) lui ont données. Machines intelligentes peut-être, mais avant tout machines à gaz. Tout cela

**L**a démocratisation sans pareille des IA génératives d'images et de textes a rappelé, dans un troublant mélange d'engouement et d'inquiétude, la réalité de l'intelligence artificielle – ou plutôt que l'intelligence artificielle était une réalité. Ces merveilles technologiques appellent à coup sûr des lendemains qui chantent pour les géants du numérique, mais peut-être aussi des lendemains qui déchantent pour des professionnels en voie de déclassement intellectuel face à des dispositifs de plus en plus performants. C'est là qu'il faut raison garder, car ces IA ne forment, en réalité, ni une rupture ni un saut conceptuel, simplement la suite naturelle de la numérisation de nos vies, engagée à grande vitesse.

Qu'il ambitionne de la mettre à son service ou qu'il s'agisse, pour le lecteur, de mieux se familiariser avec ses principaux concepts (data, modèles, algorithmes, outillage en entreprise, éthique), ce livre lui propose de dégager l'IA de sa gangue de mythes pour le faire parvenir à une vision techniquement juste et informée des impacts pluriels de cette technologie. Car plus que jamais, il devient urgent de comprendre ce que sont ces outils, c'est-à-dire ce qu'est l'IA qui les a formés, son mode de fonctionnement, ses limites, ses enjeux... afin de garder le contrôle.

Ingénieur de l'Ecole des Mines de Paris, Axel Cypel possède une expérience variée dans la direction de projets d'IA. Auteur d'un premier essai de vulgarisation scientifique (« Au cœur de l'intelligence artificielle. Des algorithmes à l'IA forte », paru aux éditions De Boeck Supérieur), et ayant à cœur d'expliquer l'IA, il mène également une activité de conférencier dans ce domaine.

17,90 €

ISBN : 978-2-8073-6059-4



9 782807 360594

deboeck **B**  
SUPÉRIEUR

[www.deboecksuperieur.com](http://www.deboecksuperieur.com)