

Manuel d'introduction à l'IA

Un compagnon pour développer les
connaissances en matière d'IA

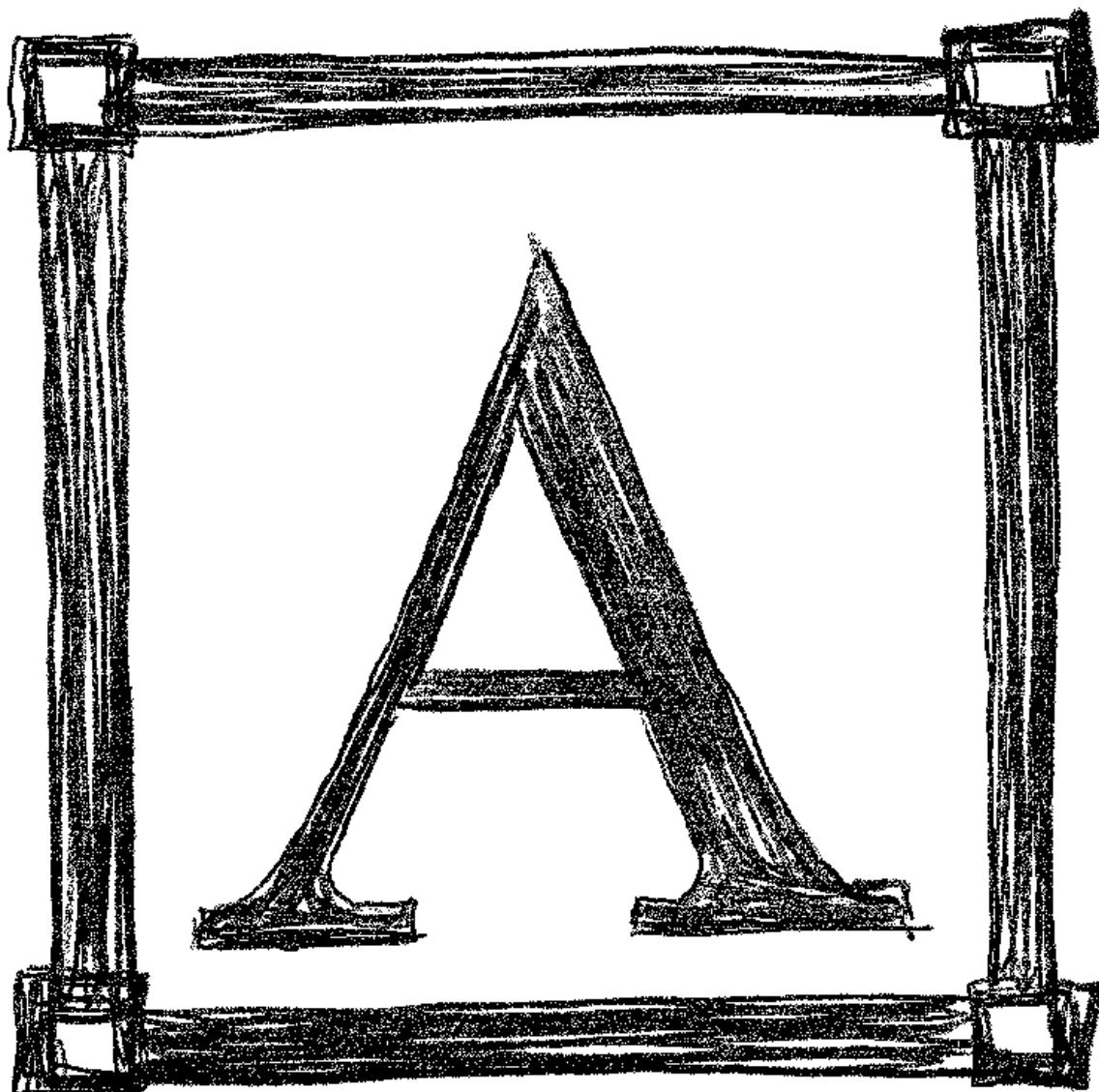
Par Juliet Waters

Illustrations de Danielle Grisnich

2023



Le b.a.-ba de la citoyenneté à l'ère de l'IA



Attitude

Que pensez-vous de l'intelligence artificielle? Êtes-vous curieux-se ou hésitant-e? Enthousiaste ou inquiet-ète? Avez-vous hâte de découvrir les avantages que présentent les étonnants agents conversationnels et les voitures autonomes? Vous imaginez-vous confier des tâches banales à l'intelligence artificielle (IA) et disposer ainsi de plus de temps pour vous consacrer à votre famille, à vos ami-e-s et à des activités stimulantes? Ou craignez-vous que l'IA ne remplace les compétences et les talents que vous offrez au monde?

Sachez que, quoi que vous ressentiez, vous n'êtes pas seul-e.

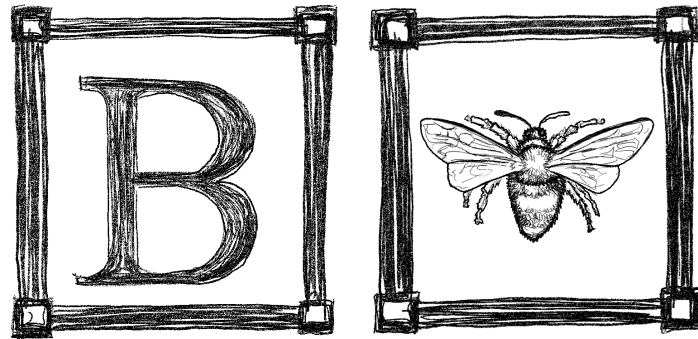
Une enquête récente menée dans [120 pays](#) indique que l'IA n'est pas perçue de la même manière dans toutes les régions du monde. En Europe, en Amérique du Nord et en Amérique latine, les attitudes négatives prédominent. Au Moyen-Orient, en Asie centrale et en Asie du Sud, les attitudes positives l'emportent légèrement sur les attitudes négatives. L'écart est encore plus grand en Afrique et en Asie du Sud-Est, où les opinions à l'égard de l'IA sont plus positives. En Chine, seuls 11 % des personnes interrogées ont une attitude négative à l'égard de l'IA.

Le clivage régional des opinions sur l'IA reflète probablement les différences culturelles et économiques à travers le monde, ce qui soulève une autre question : nos opinions sur l'IA sont-elles fondées sur nos connaissances et notre expérience, ou sur des préjugés personnels et culturels?

Quel que soit notre point de vue, l'IA n'est pas près de disparaître. Son influence sur nos sociétés s'accélère. Les changements qu'elle apporte affecteront la manière dont nous interagissons avec notre économie, ainsi que les relations sociales et les responsabilités citoyennes. En effet, tous les changements technologiques, y compris l'invention de l'écriture, de l'imprimerie, du téléphone, d'internet ou de l'IA, finissent

par avoir un impact sur la manière dont nous comprenons, utilisons et communiquons l'information. Les changements dans le traitement de l'information modifient également la société. Nous nous devons d'avoir une opinion éclairée sur l'IA ainsi qu'une compréhension de base de son fonctionnement. Cela nous aidera à travailler, à enseigner et à vivre aux côtés de machines intelligentes.

Il convient de trouver un équilibre entre une saine dose de scepticisme et de prudence, une ouverture d'esprit déterminée et une vigilance constante, pour réévaluer nos propres idées préconçues.

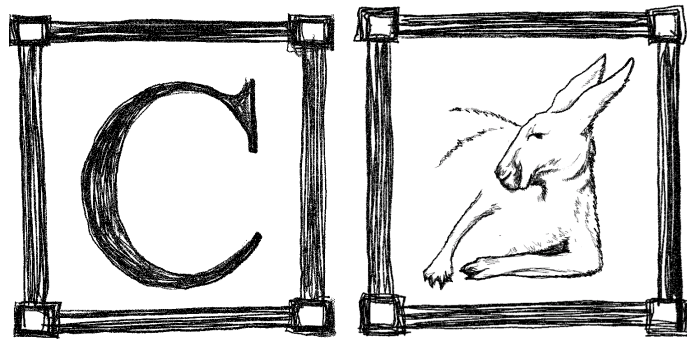


Sauvons les abeilles, de même que notre intelligence collective

En 2017, les Nations unies ont déclaré le 20 mai Journée mondiale de l'abeille. Trente-cinq pour cent des abeilles meurent chaque année, en grande partie à cause de l'accélération des tendances du changement climatique, telles que la fluctuation des schémas saisonniers et l'augmentation de la fréquence des catastrophes naturelles. Les apiculteur-riche-s traditionnel-le-s ne sont tout simplement pas en mesure de faire face à l'énorme quantité de travail nécessaire pour rétablir cet équilibre. C'est une très mauvaise nouvelle, mais pas seulement pour les abeilles. 90 % des plantes sauvages et 75 % des cultures dépendent de la pollinisation, ce qui signifie qu'un tiers des aliments que nous consommons nécessite leur survie. Sans les abeilles, le tiers de notre approvisionnement mondial en alimentation pourrait disparaître!

Mais voici une bonne nouvelle : grâce aux innovateur·rice·s de l'IA, nous disposons aujourd'hui de ruches autonomes alimentées par l'énergie solaire, dotées de systèmes de vision par ordinateur et de robotique qui aident les apiculteur·rice·s à surveiller et à assurer la thermorégulation de leurs ruches, à réduire les parasites de manière organique, à augmenter le taux de survie des abeilles et à diminuer la quantité de travail humain nécessaire pour s'occuper des abeilles.

Si, d'ici 2030, les abeilles prospèrent et que l'alimentation se développe au rythme dont nous avons besoin pour continuer à croître, cela pourrait en grande partie s'expliquer par la capacité de l'IA à accélérer les changements positifs. Mais cela démontrera également notre intelligence collective en développant des solutions ingénieuses qui mettent l'IA au service des écosystèmes et du bien-être de l'humain.



Protéger les kangourous et toutes les espèces (y compris les humains) affectées par l'IA...

Lorsque Volvo a commencé à tester [ses premières voitures autonomes en Australie](#), elle s'est heurtée à un problème. Les véhicules pouvaient identifier et éviter les animaux d'Europe du Nord tels que les cerfs et les élans. Mais pour une IA, qui ne dispose pas d'une perception intuitive de la profondeur, les kangourous posaient un problème

différent. Lorsque les kangourous sont en plein saut, ils semblent plus éloignés qu'ils ne sont en réalité. Lorsqu'ils atterrissent, ils semblent plus proches. L'IA n'avait aucun moyen d'expliquer ou de corriger cette confusion. C'est à l'humain de trouver la solution et de fournir suffisamment de données sur les kangourous pour remettre l'IA dans le droit chemin.

Mais tous les défis ne sont pas aussi évidents pour l'œil humain que le problème d'identification des kangourous. L'IA recueille beaucoup de données et est capable d'y reconnaître des schémas complexes, mais elle ne comprend pas véritablement les informations et n'est pas en mesure de fournir un contexte.

Par exemple, l'assistant intelligent Alexa peut apprendre à reconnaître une commande telle que « commande-moi une maison de poupée ». Mais il n'est pas nécessairement capable de distinguer si c'est un parent ou un enfant qui commande la maison de poupée. C'est ce qui s'est passé dans une famille, où l'on a découvert que la petite fille avait réussi à commander une maison de poupée très coûteuse. Lorsque l'histoire a été racontée un jour à la radio locale, le commentateur a dit « J'adore l'histoire de la petite fille qui a dit Alexa, commande une maison de poupée », *une autre* Alexa dans une autre maison a entendu la commande et a commandé *une autre* maison de poupée.

Dans le passé, les ordinateurs suivaient des règles créées par les êtres humains. Mais aujourd'hui, grâce aux progrès de l'« apprentissage automatique », l'IA est guidée par les données. Dans une large mesure, elle établit ses propres règles en fonction de ce qu'elle déchiffre à partir de ces données. Nous en apprendrons davantage sur ce processus ultérieurement.

La valeur de l'IA réside dans sa capacité à traiter d'énormes quantités de données. Cependant, cette même capacité pose de nouveaux problèmes. À mesure que le domaine de l'IA progresse et que les algorithmes d'apprentissage deviennent plus complexes, l'IA développe des stratégies qu'il est impossible pour un humain, et même pour un système d'IA, de suivre.

C'est ce qu'on appelle le problème de la « boîte noire ».

Vous avez peut-être déjà entendu parler de la boîte noire des avions ou des voitures à laquelle on accède après un accident. Ces boîtes enregistrent toutes les activités et les commandes qui ont précédé l'accident afin que nous puissions tirer des leçons de ce qui s'est passé.

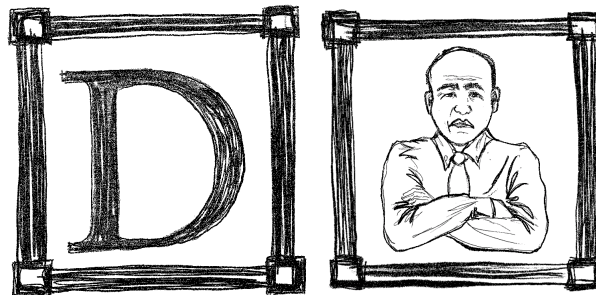
Si seulement l'IA disposait d'une boîte similaire! Malheureusement, si l'on considère la quantité de données créées par les algorithmes d'apprentissage automatique, il est probable que la boîte noire d'un système d'IA que l'on pourrait ouvrir n'émettrait qu'une masse indéchiffrable de bruits blancs.

Le problème avec la boîte noire dans l'IA est qu'il n'y a, en réalité, pas de boîte noire.

Les algorithmes d'apprentissage automatique traitent tellement d'informations que l'IA revêt un caractère quelque peu mystérieux pour nous... mais aussi pour elle-même. De notre point de vue, tout ce dont nous pouvons être sûrs, c'est des données qui entrent dans les algorithmes et de celles qui en sortent. Cela signifie que l'IA a toujours besoin d'une intervention humaine pour vérifier et équilibrer son travail.

L'IA peut faire beaucoup pour rendre le monde plus sûr et plus efficace. Elle peut aussi faire beaucoup pour rendre le monde encore plus dangereux, confus et inéquitable qu'il ne l'est déjà.

Ce qui nous amène à un autre point...



Données, décisions et diversité

En fin de compte, l'IA ne sera véritablement efficace ou digne de confiance que si elle bénéficie de données de qualité, diversifiées et explicables sur lesquelles s'entraîner. En outre, sa valeur dépendra de notre capacité à prendre de bonnes décisions sur la manière dont elle devrait, ou ne devrait pas, être utilisée.

Nous partageons tous le désir d'être heureux et de mettre notre intelligence collective au service de la construction d'un monde qui favorise la sécurité, le bien-être et le respect mutuel. La manière dont l'IA s'inscrit dans cet objectif dépend de nous!

Pour commencer

Qu'est-ce qu'un manuel d'introduction?

Commençons par ce que ce n'est pas : un manuel d'introduction n'est pas un guide exhaustif. Il ne s'agit pas non plus d'un manuel qui vous indiquera le contenu de l'examen. Ce n'est pas non plus l'endroit où vous trouverez un glossaire des termes (bien que ce manuel vous dirige vers d'autres ressources qui contiennent des glossaires).

Un bon manuel d'introduction est un compagnon convivial qui facilite les premiers pas. Il s'appuie sur les connaissances existantes et fournit juste assez de nouvelles connaissances pour vous mettre sur la bonne voie et vous permettre d'en acquérir encore plus. Imaginez à un manuel classique qui utilise la banque d'images et la mémoire orale d'un enfant et qui associe des mots simples à ces images afin que l'enfant puisse commencer à décoder de nouvelles connaissances.

Mais un manuel d'introduction destiné à des personnes qui ont déjà passé un certain temps à vivre et à apprendre dans le monde doit également accomplir une autre chose : il doit aussi identifier les connaissances et les idées existantes susceptibles d'entrer en conflit avec les nouvelles façons de voir les choses. Voyez ce manuel comme une sous-couche de peinture, que vous devez appliquer pour que la nouvelle couleur de peinture apparaisse comme vous le souhaitez et ne soit pas déformée par les couleurs ou les motifs qui existaient auparavant.

Dans le monde de l'IA en particulier, nous devons être conscients des idées préconçues que nous formulons. Nos idées sur la société et la réalité imprègnent la manière dont nous créons et utilisons l'IA et peuvent devenir des préjugés qui affaiblissent à la fois l'utilité et l'équité de l'IA.

Joy Buolamwini, chercheuse au MIT, [l'a découvert](#) lorsque les algorithmes de reconnaissance faciale avec lesquels elle travaillait n'ont pas réussi à reconnaître son visage. Les algorithmes n'avaient pas été suffisamment entraînés avec des photos de femmes noires. Les programmeur·euse·s qui ont formé l'IA ont supposé que les énormes banques de données de photos qui avaient alimenté leur IA reflétaient la diversité du monde. Mais ce n'était pas le cas. Aucun système automatisé n'est intrinsèquement neutre. Comme toutes les choses faites par les humains, chaque programme automatisé reflète les connaissances et les idées préconçues de la personne qui l'a créé. Si les créateur·rice·s d'une IA ont inconsciemment de fausses idées sur les gens, l'IA peut ancrer ces erreurs dans des systèmes plus vastes qui ont un impact sur des milliers, voire des millions de personnes.

La citoyenneté à l'ère de l'IA commence par un état d'esprit qui consiste à confronter et à atténuer intentionnellement nos préjugés, en particulier lorsque les systèmes qui nous entourent risquent de les amplifier.

À qui s'adresse ce manuel d'introduction?

Ce manuel d'introduction s'adresse aux adolescent·e·s, aux éducateur·rice·s et aux adultes qui s'intéressent à l'IA, dans un cadre formel ou informel. Il les aidera à :

1. reconnaître ce qu'ils savent déjà;
2. acquérir une bonne compréhension conceptuelle de ce qu'il leur reste à apprendre, voire à désapprendre;
3. essayer certaines choses afin de constater comment fonctionne l'IA;
4. se sentir à même de partager avec d'autres ce qu'ils ont appris;
5. s'amuser!

Une fois que ce manuel d'introduction vous aura aidé à conceptualiser le fonctionnement de l'IA, il vous orientera vers des activités pratiques ludiques et d'autres pistes à explorer, non seulement pour apprendre, mais aussi pour jouer.

L'IA peut enrichir nos vies de toutes sortes de façons, au-delà de la simple résolution de problèmes. Elle est utilisée dans le monde entier pour créer des œuvres d'art, des médias et d'innombrables autres projets inventifs et collaboratifs. L'IA peut rendre la vie plus colorée, plus inspirante et plus pratique. Elle peut nous donner plus de temps pour être de bon·ne·s ami·e·s, parents, enfants et citoyen·ne·s. L'IA a le potentiel de devenir un outil de cocréation. Mais ce résultat dépend de l'équilibre entre la confiance et la collaboration, d'une part, et le respect des perspectives d'autrui, de la vie privée et du niveau de confort avec la technologie, d'autre part.

Avant de commencer : n'oubliez pas la fameuse boîte noire!

Quel que soit le degré d'amusement ou de réussite que vous procure l'IA, n'oubliez jamais l'exemple des kangourous! Gardez à l'esprit que, peu importe ce que nous

apprenons, un problème demeure : l'IA ne pourra jamais expliquer entièrement comment elle résout les problèmes. Ses ensembles de données sont trop volumineux et ses algorithmes trop complexes. C'est pourquoi il est si important que les humains conservent un droit de regard à l'heure où nous évoluons vers une société de plus en plus dépendante des machines intelligentes.



L'IA peut fournir des réponses, mais contrairement à vous, elle ne peut pas les comprendre.

Demandez à l'IA de vous expliquer cela :



Does ChatGPT understand questions?





Yes, ChatGPT is able to understand and respond to questions. It is a language model that has been trained on a large dataset of text and can generate human-like responses to a wide variety of prompts. If you have any questions, feel free to ask.  



Does ChatGPT understand what it's writing?



ChatGPT, like other language models, does not have the ability to understand the meaning of the text it generates. It is trained to predict the next word in a sentence based on the patterns it has learned from the text data it was trained on. However, it can still generate coherent and often human-like responses to prompts, but it doesn't have any understanding of what it's writing.  

Comment puis-je en apprendre plus sur l'IA si je ne peux jamais retracer le processus que celle-ci utilise pour parvenir à ses réponses?

Là encore, il s'agit d'une question d'attitude. Ce n'est pas parce que vous ne pouvez pas tout savoir que vous ne pouvez pas commencer et continuer à apprendre. Vous ne comprenez peut-être pas exactement comment les plantes poussent, mais cela ne doit pas vous empêcher de jardiner. Vous ne comprenez peut-être pas entièrement comment le cerveau apprend, mais cela ne vous empêchera pas de devenir enseignant.

Au fur et à mesure que vous apprendrez et travaillerez avec l'IA, vous développerez une compréhension des concepts qui régissent son fonctionnement. Vous découvrirez les meilleurs types de questions à poser à l'IA que vous utilisez. Apprendre à poser les bonnes questions est essentiel pour chaque citoyen·ne, qu'ils et elles soient étudiant·e·s, expert·e·s ou non-expert·e·s, car cela nous aide à saisir l'ampleur des problèmes que l'IA peut résoudre. Cela nous aide à évaluer les risques que pose l'intelligence artificielle. Sachant cela, il nous est possible de décider quels problèmes pourraient être mieux résolus en utilisant l'intelligence humaine naturelle, la pensée critique, le dialogue interpersonnel et d'autres méthodes traditionnelles de résolution de problèmes.

Nous nous informons sur l'IA afin de pouvoir prendre les meilleures décisions quant aux activités et aux conditions dans lesquelles il convient de l'utiliser. Il se peut que nous découvriions que certaines activités et certaines conditions ne conviennent pas. Si l'IA se trompe en recommandant un film qu'elle pense pouvoir nous plaire, il est peu probable que nous décidions de ne plus l'utiliser, dans la mesure où elle est généralement fiable. Si un robot de soins aux personnes âgées conçu pour transférer des patients de 90 ans d'un fauteuil roulant à un lit ne fait tomber qu'un patient sur cinq, ou même un patient sur mille, il ne s'agit pas d'une IA que nous pourrions utiliser de manière éthique sans continuer à garantir une réglementation et une surveillance humaines significatives et permanentes.

Commençons donc par le premier concept à comprendre lorsque nous parlons d'IA dans le monde dans lequel nous vivons : l'apprentissage automatique.

Comment maîtriser l'apprentissage automatique?

Vous avez probablement déjà appris comment fonctionne l'apprentissage automatique, et cela d'une manière très significative!

Si vous avez déjà développé une préférence pour un certain type de nourriture, vous l'avez probablement fait de la même manière que les machines développent leurs capacités d'apprentissage. La grande différence est que les machines n'éprouvent pas les sensations qui vous permettent de développer vos préférences alimentaires. Les machines peuvent apprendre à partir d'images, de sons, de nombres et de nombreux autres types de données, mais cela ne constitue pas pour autant ce que nous décrivons comme des « sensations ».

Pour l'instant, voyons ce que nous avons en commun avec les ordinateurs : un « langage source » binaire qui nous aide à interpréter le monde. Le terme « **binaire** » désigne tout système reposant sur deux valeurs seulement. Pour les ordinateurs, ce langage source binaire est le modèle composé de 1 et de 0 à partir duquel tous les autres langages et codes informatiques sont construits. Les êtres humains disposent d'un langage très similaire à un langage source binaire dans la manière dont le cerveau classe les sensations comme agréables ou désagréables. Par le simple fait d'aimer ou de ne pas aimer des sensations, et de suivre le degré auquel nous aimons ou n'aimons pas ces sensations, notre cerveau a créé des millions de réponses au monde qui nous entoure. Le fait d'aimer ou de ne pas aimer nous aide à faire des choix dans nos environnements complexes.

Au niveau fondamental, c'est ainsi que fonctionne l'apprentissage automatique. C'est plutôt simple. Un ordinateur recueille des ensembles de données, leur attribue des étiquettes et tente ensuite de créer des règles à partir des modèles qu'il découvre.

Création d'un modèle

Un autre processus que nous avons en commun avec les machines est que notre apprentissage crée des modèles, que nous utilisons ensuite pour prédire des choses.

Imaginez un aliment que vous aimez, un bol de crème glacée ou une fraise biologique bien charnue. Il vous suffit de l'imaginer et votre langue commencera déjà à saliver, que vous ayez ou non de la crème glacée ou des fraises à portée de main.

Il est peu probable que les mots « **analyse prédictive** » suscitent le même niveau de salivation, mais c'est exactement ce que votre cerveau est en train de faire. Votre cerveau prédit et pressent le plaisir sur la base d'un souvenir emmagasiné. Il agit selon un modèle que vous avez stocké pour un certain aliment. Ce modèle classe les données et notre cerveau prédit nos réponses à partir des schémas contenus dans ces données. À partir de ces modèles, le cerveau commence à élaborer des règles. Les règles deviennent la base des comportements et des réponses automatiques de votre cerveau.

Données

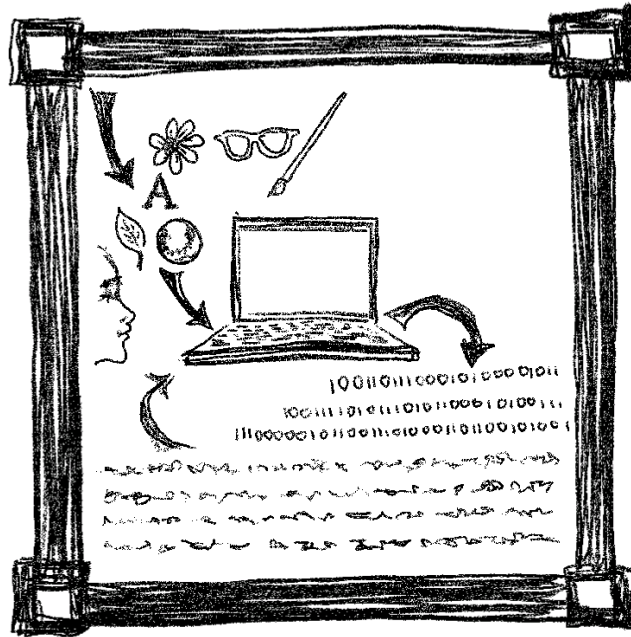
De la même manière que vous avez fait des prédictions sur les choses que vous aimez, les ordinateurs du monde entier font des prédictions sur votre comportement en se basant sur votre activité en ligne. En moyenne (en 2023), chaque jour, les humains du monde entier :

- utilisent Google 4 millions de fois;
- regardent 4,5 millions de vidéos sur YouTube;
- échangent 188 millions de courriels.

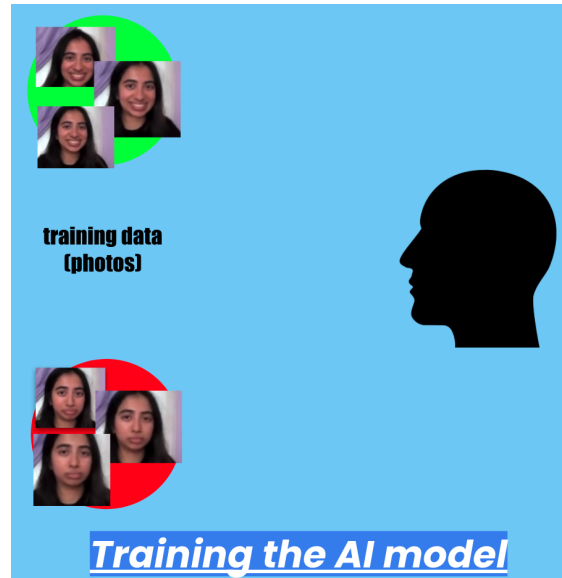
Ces échanges produisent beaucoup de données. Les progrès réalisés dans le domaine des puces et des semi-conducteurs ont considérablement accéléré la facilité et la rapidité avec lesquelles les ordinateurs peuvent désormais travailler avec de grands ensembles de données, ce qui a largement contribué au développement de l'apprentissage automatique. Les ordinateurs passent au crible les mégadonnées pour faire des prédictions sur ce que vous aimez ou n'aimez pas, et répondent souvent à ces prédictions en vous suggérant des choses que vous aimeriez acheter, regarder ou consommer.

Examinons ce processus de plus près :

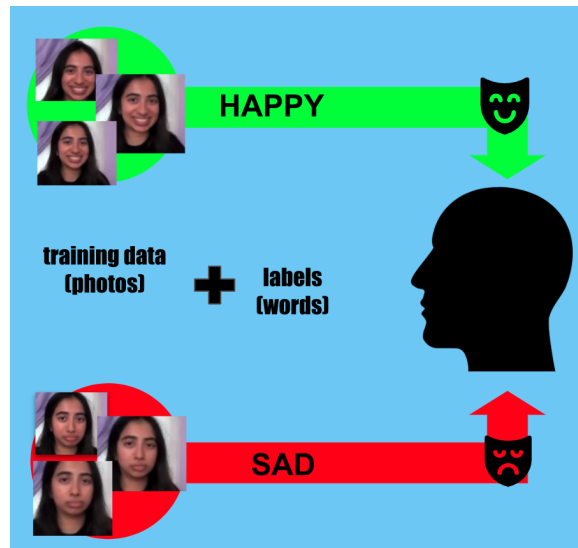
Entraînement des algorithmes d'IA



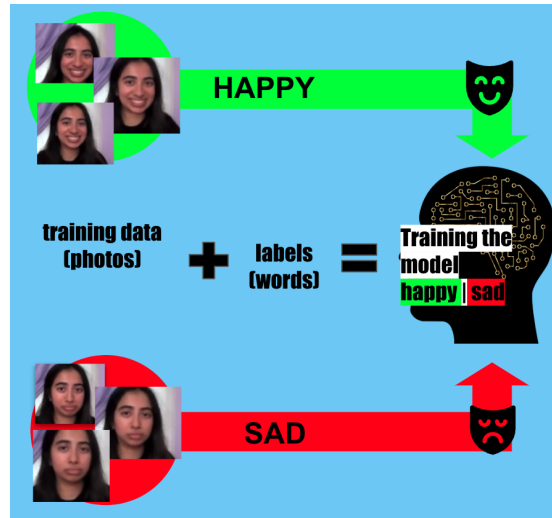
Pour mieux comprendre comment l'IA peut prédire ce que vous aimerez ou n'aimerez pas, examinons le processus par lequel une IA peut apprendre quels types de choses vous rendent heureux ou triste.



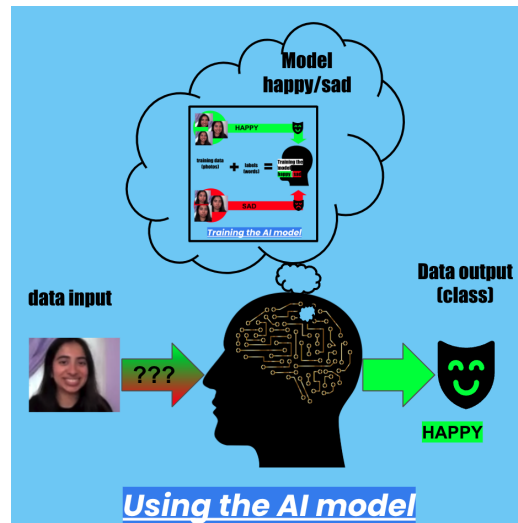
Avant que l'IA ne puisse commencer à prédire vos réponses, elle doit étiqueter les données qu'elle collecte. Dans ce cas, l'IA est entraînée à partir de captures d'écran de visages pour distinguer les visages heureux des visages tristes. Elle aura besoin d'un bon nombre de captures, disons un minimum de 30, pour identifier les caractéristiques communes d'un visage « heureux ».



L'IA n'a pas de compréhension inhérente du bonheur ou de la tristesse. Elle ne peut que diviser les choses en catégories qui ont été étiquetées avec des variables, comme les mots « heureux » et « triste ». L'IA ne se soucie pas des étiquettes (parce qu'elle n'est pas une personne, dotée de la capacité de se soucier de quoi que ce soit), elle a juste besoin de suffisamment de données pour pouvoir trier les objets dans les modèles d'ensembles de données que nous lui donnons. Avec suffisamment d'exemples de visages souriants étiquetés « heureux » et de visages abattus étiquetés « tristes », l'IA peut créer un modèle fonctionnel pour tester son efficacité avec des ensembles de données beaucoup plus vastes. Bien qu'elle soit en mesure de créer un modèle préliminaire avec seulement 30 exemples, pour renforcer et affiner le modèle, elle aura besoin d'exemples beaucoup plus nombreux et diversifiés.



À l'aide de ce modèle, l'IA peut alors « émettre » sa supposition sur l'émotion véhiculée par le visage et tester si sa supposition est correcte ou non. Au fur et à mesure que l'IA recueille davantage de données et teste ses propres prédictions, elle devient de plus en plus précise, car ses choix sont fondés sur un nombre de plus en plus important d'exemples humains réels. L'algorithme qui permet à l'IA de construire son propre ensemble de règles est appelé « réseau neuronal ».

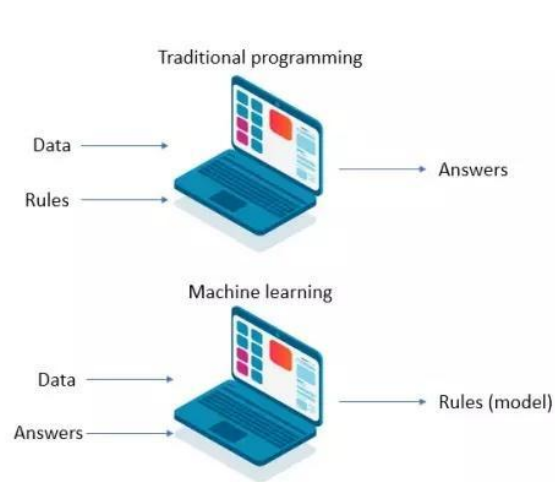


La différence importante entre les algorithmes d'apprentissage automatique, ou réseaux neuronaux, et les algorithmes traditionnels utilisés avant les progrès de l'IA, est que dans le passé, les humains établissaient les règles que les machines suivaient.

Un algorithme informatique traditionnel pourrait être un programme fournissant des images qu'un humain aurait déjà étiquetées, puis l'ordinateur produirait des réponses basées sur ces données pré-étiquetées.

Dans le cadre de l'apprentissage automatique, nous pouvons étiqueter de grands ensembles d'images. Nous laissons ensuite l'algorithme déduire les similitudes, puis nous vérifions s'il étiquette correctement les nouvelles données.

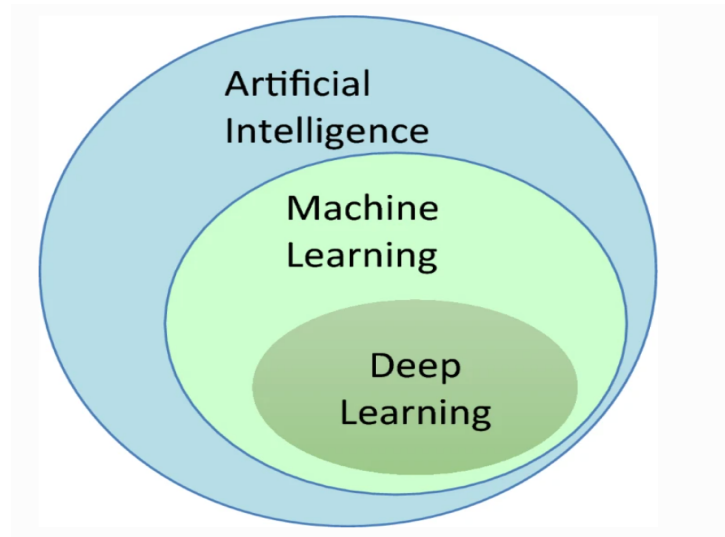
Plus l'ordinateur a été entraîné sur des données et a appris à partir de celles-ci, plus la machine elle-même établit des règles à partir du modèle qu'elle a créé.



C'est important! Il convient de toujours garder à l'esprit que ce n'est pas parce qu'un algorithme informatique *est capable* d'établir un grand nombre de règles qu'il *doit* les établir toutes. Cela ne signifie surtout pas que toutes les règles sont toujours justes!

Le type d'apprentissage effectué par l'IA est encore ce que nous appelons un apprentissage « étroit ». Contrairement aux humains, l'IA n'est pas équipée pour mettre son apprentissage en contexte comme le font les humains.

Plus les algorithmes d'IA sont axés sur les données, plus ils entrent dans l'état appelé « apprentissage profond ». Dans ce cas, les couches de réseaux neuronaux sont nombreuses et complexes. Elles sont moins transparentes et plus difficiles à expliquer. Si une IA se retrouve dans une situation tragique avec un kangourou, c'est probablement parce qu'il n'y a pas eu suffisamment d'implication humaine dans la formation et le façonnage de l'algorithme pour toutes les situations différentes qu'un être humain pourrait rencontrer.



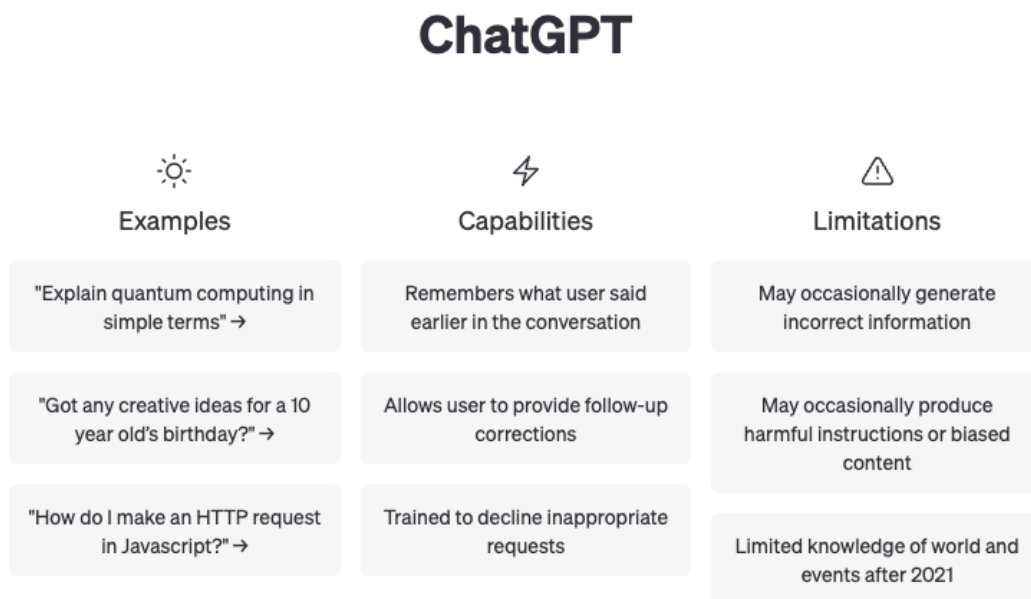
Mais revenons à notre modèle « heureux/triste ». Il existe de nombreuses raisons pour lesquelles l'IA pourrait encore se tromper ou établir des corrélations qui ne sont pas pertinentes. Il se peut, par exemple, qu'elle ait été entraînée sur des milliers de visages, mais que ces visages ne soient que caucasiens. Ou encore, ses utilisateur·rice·s étaient issu·e·s d'une culture plus susceptible de réagir avec mécontentement à quelque chose qui produisait du bonheur dans une culture dominante.

Les humains peuvent et doivent continuer à influencer les règles qui régissent l'apprentissage automatique. L'astuce pour une utilisation optimale de l'apprentissage automatique réside dans le fait de savoir quand nous voulons que les machines établissent les règles, et quand nous ne le voulons pas. Nous pourrions, par exemple, bénéficier de systèmes hybrides qui utilisent l'apprentissage automatique pour l'apprentissage automatisé à faible risque, mais qui font appel à des humains pour s'assurer que les données sont à la fois diversifiées et bien équilibrées pour les décisions importantes où les risques sont plus élevés.

L'utilisation de « cartes modèle » est un moyen de clarifier les décisions de ce type. Il s'agit de documents que les créateurs et créatrices d'IA peuvent utiliser pour

apporter davantage de transparence sur la manière dont un modèle d'IA est destiné à être utilisé, sur ses limites et ses avertissements, sur les données à partir desquelles il a été entraîné et sur les étiquettes et les caractéristiques utilisées pour entraîner les données.

Voici, par exemple, une carte modèle créée pour ChatGPT, l'agent conversationnel aux compétences impressionnantes. Cette carte modèle a été introduite quelques mois après que la première vague d'utilisateur-riche-s se soit rendue compte que ChatGPT n'était pas toujours aussi performant qu'il y paraissait :



Les obstacles invisibles

Les cartes modèles sont l'un des moyens par lesquels un-e créateur-riche d'IA peut aider un-e utilisateur-riche à déterminer si cette IA est adaptée au problème que l'utilisateur-riche souhaite résoudre ou à l'objet qu'il ou elle tente de créer.

Mais tous les systèmes pilotés par l'IA qu'un-e citoyen-ne vivant à l'ère du numérique rencontrera ne sont pas accompagnés d'une carte modèle. Dans notre vie à l'ère de l'IA, l'une des principales difficultés est de prendre conscience des modes de fonctionnement de l'IA qui sont invisibles ou obscurs pour nous. Lorsqu'une plateforme de diffusion en continu vous fait des recommandations ou que des publicités pour quelque chose que vous avez récemment mentionné dans un courriel ou sur les réseaux sociaux apparaissent comme par magie, vous savez que l'IA a suivi et prédit votre comportement.

L'IA s'améliore sans cesse pour prédire ce que les humains vont faire ou dire. Si vous avez essayé un outil d'IA conversationnelle comme ChatGPT, vous avez vu avec quelle efficacité l'IA peut couper et coller les phrases qu'elle pense être de bonnes réponses à vos questions.

Vous avez peut-être aussi déjà constaté par vous-même que ChatGPT peut échouer. Il peut fournir une réponse qui semble très crédible, mais qui est totalement ou partiellement fausse. Souvent, les plus grosses erreurs ne sont pas créées par l'IA, mais par les humains qui utilisent l'IA. La plus grande erreur que nous puissions commettre est de supposer que parce que l'IA fait un excellent travail d'imitation des humains, elle est en réalité humaine.

L'une de nos tendances humaines les plus universelles est l'anthropomorphisme, c'est-à-dire la tendance à projeter notre intelligence humaine sur d'autres espèces et même sur des objets. Nous croyons souvent que nos animaux de compagnie sont

plus intelligents qu'ils ne le sont en réalité, et il est facile pour nous de supposer que l'IA dans notre vie a des pouvoirs humains, voire magiques.

Mais en réalité, il y aura toujours des choses que les humains feront mieux que l'IA, en particulier lorsqu'il s'agit de comprendre le contexte et ce que l'on appelle l'apprentissage « large ». Nous restons tout de même plus aptes à établir les connexions visuelles et spatiales qui nous permettent de juger de la proximité ou de la distance d'un kangourou par rapport à notre voiture, et ce même si nous ne sommes pas australiens!

Les agents conversationnels devenant de plus en plus courants dans le monde, il sera de plus en plus important de savoir quand vous interagissez avec l'un d'entre eux.

Nous ne pouvons pas prédire l'avenir de l'IA. Mais on peut certainement affirmer que l'IA met les humains au défi de devenir plus intelligents et encore plus compétents en les aidant à s'améliorer dans les tâches d'apprentissage répétitives pour lesquelles nous ne sommes pas aussi bien équipés.

Cela nous pousse également à démontrer notre capacité supérieure à effectuer le type de travail que l'IA ne peut pas faire, tel que réfléchir de manière critique à des situations, prendre des décisions sur ce à quoi nous tenons, comprendre pourquoi nous tenons à ces choses et nous soucier du bien-être des autres.

Ce que l'apprentissage automatique nous apprend sur notre propre apprentissage

L'une des plus grandes opportunités offertes par l'IA est qu'elle nous pousse à mieux comprendre les différences importantes entre la façon dont les machines

apprennent et la façon dont les humains apprennent. En apprenant à mieux connaître et juger ces différences, nous serons mieux à même de prendre des décisions sur la manière d'utiliser l'IA de manière éthique et efficace.

L'apprentissage de l'IA peut apprendre à l'humain plusieurs choses sur sa propre intelligence :

1. **Limites** : L'une des leçons les plus importantes que nous pouvons tirer de l'IA est que notre propre intelligence a des limites inhérentes. Les modèles d'IA peuvent effectuer des calculs complexes à la vitesse de l'éclair et mémoriser facilement de grandes quantités de données, ce qui met en évidence les limites de nos propres capacités cognitives. Cela peut nous aider à identifier les domaines dans lesquels nous devons nous améliorer et ceux dans lesquels nous devons compter sur la technologie pour nous aider.
2. **Résolution de problèmes** : Les algorithmes de l'IA sont conçus pour résoudre des problèmes spécifiques, et l'étude de la manière dont ils y parviennent peut nous apprendre de nouvelles façons de résoudre les problèmes. En comprenant comment l'IA aborde les problèmes, nous pouvons appliquer des techniques similaires pour résoudre nos propres problèmes, que ce soit dans le domaine des affaires, des sciences ou de la vie quotidienne.
3. **Partialité** : Les algorithmes d'IA sont aussi impartiaux que les données qui les alimentent. Si les données sont biaisées, les résultats s'en ressentiront. En étudiant l'IA, nous pouvons apprendre à connaître les risques de partialité et à les atténuer dans nos propres décisions.
4. **Créativité** : Bien que les algorithmes d'IA deviennent de plus en plus sophistiqués, ils ne possèdent toujours pas la créativité et l'imagination propres aux humains. Cela peut nous rappeler les qualités uniques qui font

de nous des êtres humains, telles que la créativité, l'intuition et l'intelligence émotionnelle.

Dans l'ensemble, l'étude de l'IA peut nous aider à comprendre les forces et les limites de notre propre intelligence et nous inciter à trouver des moyens innovants de résoudre des problèmes et de créer de la valeur.

Fin de ce manuel d'introduction : le début d'un parcours exaltant

Alors que les progrès de l'IA semblent s'accélérer, le chemin qui mène à son utilisation optimale sera probablement jalonné de débats sur ses utilisations et ses dangers, ainsi que d'une prise de conscience croissante des compétences que nous devons maîtriser pour mieux la comprendre. Il nous faudra dépasser la simple connaissance de l'interaction avec l'IA et acquérir des compétences de base dans des domaines émergents tels que la maîtrise des algorithmes et des données, la modélisation et l'apprentissage automatique.

L'enthousiasme suscité par l'IA peut souvent se transformer en anxiété, voire en désillusion et en désespoir. L'évolution de l'IA se fera en partie sous notre contrôle et en partie hors de notre contrôle. C'est pourquoi il est important d'ancrer notre enthousiasme dans la joie d'apprendre.

Dans les années à venir, il y aura beaucoup à apprendre sur l'IA, sur ce qu'elle peut faire et sur ce qu'elle ne peut pas faire. Il y aura également beaucoup à apprendre sur l'espèce humaine, sur ce qu'elle ne peut pas faire, mais aussi sur des choses que nous pouvons désormais faire alors que nous n'aurions pas pu les imaginer.

Nous nous tournons vers l'avenir et il peut être utile, afin de maintenir notre enthousiasme, de tenir compte du passé et de penser au principe de la septième génération utilisé par la nation Haudenosaunee (Iroquois) pour soutenir l'une des plus anciennes démocraties du monde.

Si nous concevons et utilisons l'IA avec une intention qui respecte tout ce que nous devons aux actions, aux talents et aux compétences des générations qui nous ont précédé, et tout ce à quoi nous aspirons pour les générations qui viendront après nous, nous aurons au moins la certitude d'avoir fait ce que nous pouvons pour que l'IA devienne et reste un élément vital d'une planète prospère dans laquelle toutes les espèces ont de la valeur.

Ressources

[Décode les algorithmes et les données](#)

["AI and Creative Learning: Concerns, Opportunities, and Choices"](#) (L'IA et l'apprentissage créatif : préoccupations, opportunités et choix) Mitch Resnick

Apps for good [AfG Machine Learning in a Day Student Workbook 1920 v4](#)
(Apprentissage automatique en un jour - Manuel de l'élève)

[Manuel éducatif d'Actua sur l'IA](#)

[Danser avec l'IA](#)

[A quick reference guide to artificial intelligence for teachers](#) (Un guide de référence rapide sur l'intelligence artificielle pour les enseignants)