

L'amorce sur l'IA

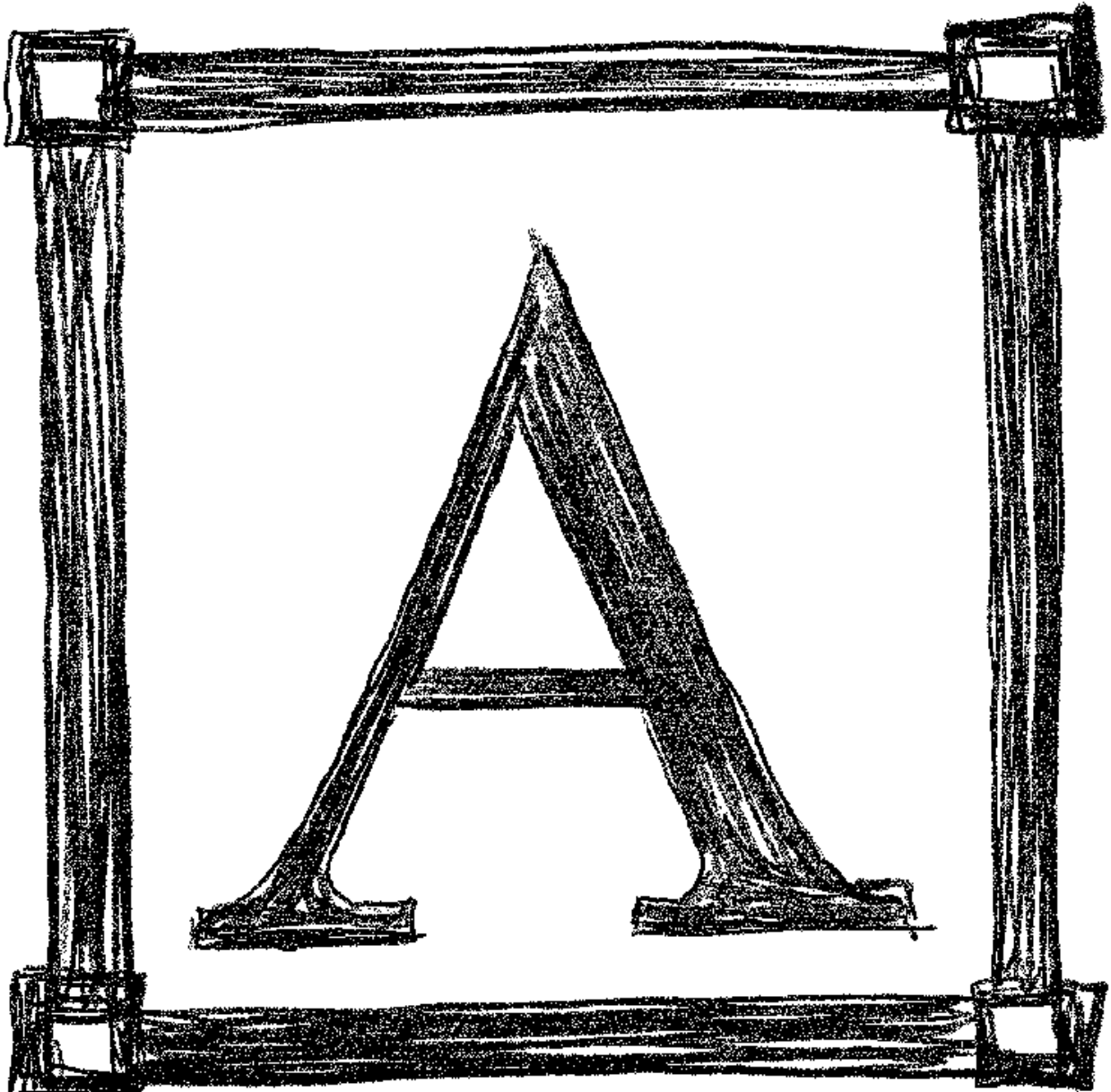
Un guide pour développer la littératie numérique en IA

Par Juliet Waters et illustrations par Danielle Grismich



L'ABC de la citoyenneté à l'ère de l'IA	3
A comme attitude	4
B comme abeilles, et notre intelligence collective humaine	5
C comme considération pour les kangourous et toutes les espèces (y compris les humains) touchées par l'IA	7
D comme données, décisions et diversité	10
Pour commencer	11
À qui s'adresse cette amorce ?	12
Le secret pour apprendre l'apprentissage automatique	15
L'entraînement des algorithmes d'IA	17
Ce que l'apprentissage automatique nous apprend sur notre propre apprentissage	26
E comme enthousiasme et la fin de cette amorce (mais seulement le début de votre parcours)	27
Ressources	28

L'ABC de la citoyenneté à l'ère de l'IA



A COMME ATTITUDE

Que ressentez-vous à l'idée de l'intelligence artificielle? Êtes-vous curieux ou hésitant? Enthousiasmé ou inquiet? Attendez-vous avec impatience la commodité des agents conversationnels étonnants et des voitures autonomes? Vous imaginez-vous confier les tâches banales à l'intelligence artificielle (IA) pour avoir plus de temps et d'attention à consacrer à votre famille, à vos amis et à des activités stimulantes? Ou craignez-vous que l'IA ne remplace les compétences et les talents que vous apportez au monde?

Soyez assuré que, quelle que soit votre réaction, vous n'êtes pas seul.

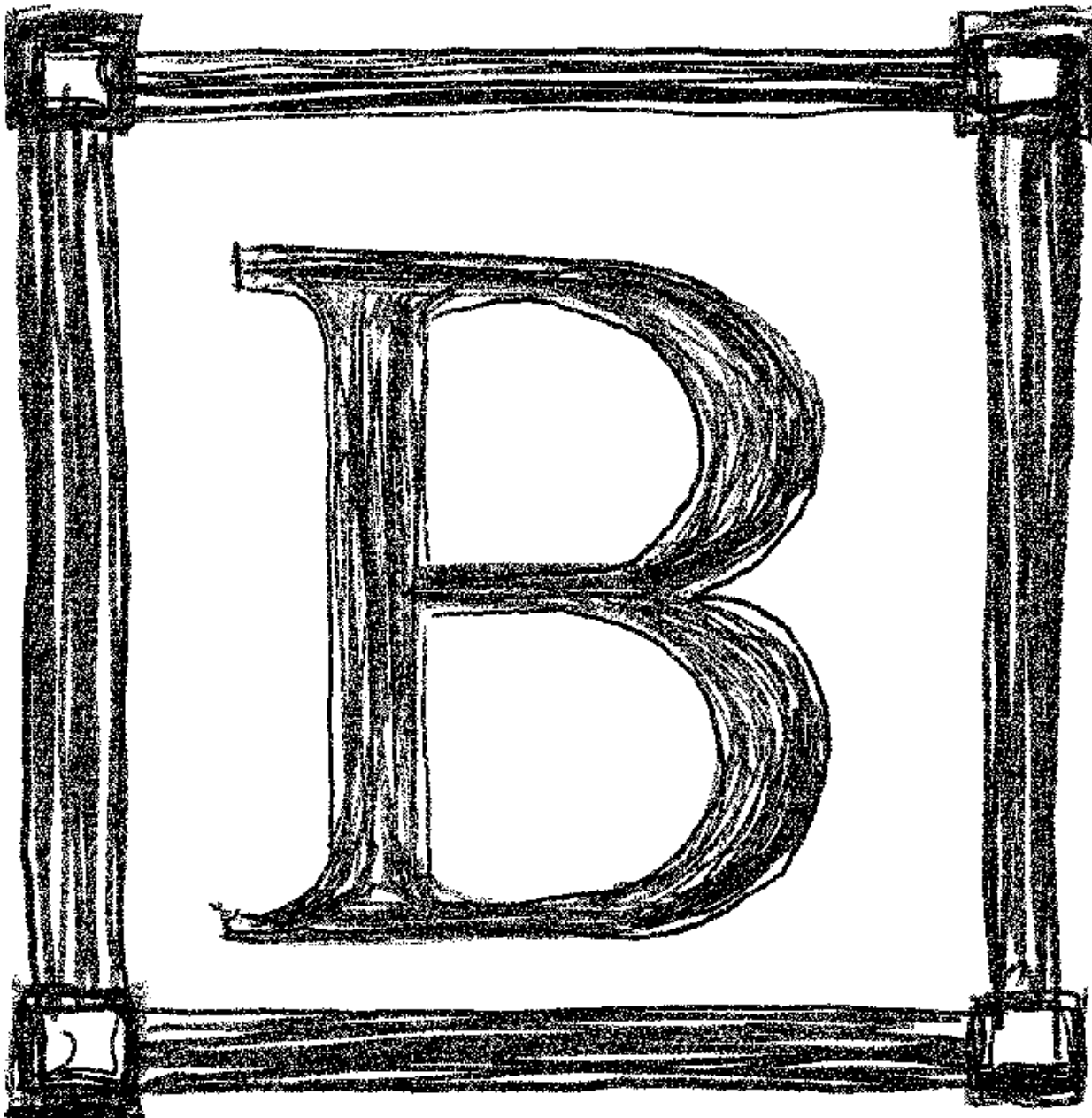
Un récent sondage mené dans 120 pays indique que les différentes régions du monde ont des perceptions différentes de l'IA. En Europe, en Amérique du Nord et en Amérique latine, les attitudes négatives dominent. Au Moyen-Orient, en Asie centrale et en Asie du Sud, les attitudes positives l'emportent légèrement sur les négatives. Un écart encore plus marqué se manifeste en Afrique et en Asie du Sud-Est, où les opinions sur l'IA sont plus favorables. En Chine, seulement 11% des répondants avaient des attitudes négatives envers l'IA.

Cette divergence régionale des opinions sur l'IA reflète probablement des différences culturelles et économiques à l'échelle mondiale, ce qui soulève une question supplémentaire: forgeons-nous nos opinions sur l'IA en fonction de nos connaissances et de notre expérience, ou de nos préjugés personnels et culturels?

Quelle que soit notre perspective, l'IA est là pour rester. Son impact sur notre société s'accélère. Les changements qu'elle entraîne influenceront sur la façon dont nous interagissons avec notre économie, ainsi que sur les relations sociales et les responsabilités citoyennes.

Cela s'explique par le fait que tous les changements technologiques, y compris l'invention de l'écriture, de l'imprimerie, du téléphone, d'Internet ou de l'IA, finissent par modifier la façon 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 des machines intelligentes. Équilibrons le scepticisme sain et la prudence avec une ouverture d'esprit déterminée et une vigilance constante pour réévaluer nos propres préconceptions.

B COMME ABEILLES, ET NOTRE INTELLIGENCE COLLECTIVE HUMAINE

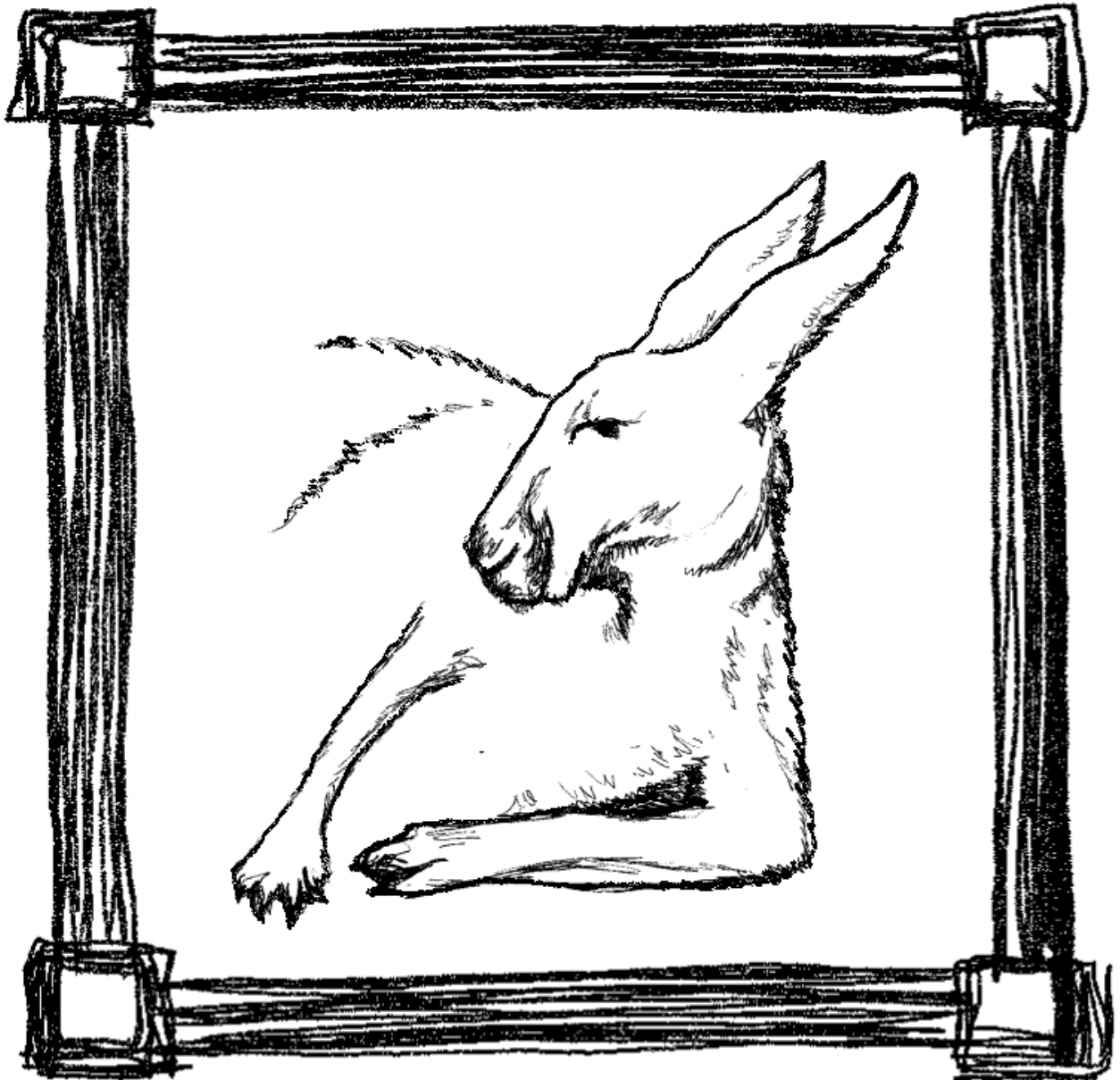


En 2017, les Nations Unies ont déclaré le 20 mai Journée mondiale des abeilles. Trente-cinq pour cent des abeilles meurent chaque année, en grande partie à cause de l'accélération des tendances liées aux changements climatiques, notamment les fluctuations des régimes saisonniers et la multiplication des catastrophes naturelles. Les apiculteurs traditionnels ne sont tout simplement pas en mesure de faire face à l'immense quantité de travail nécessaire pour rétablir cet équilibre. C'est une très mauvaise nouvelle, mais pas seulement pour les abeilles. Étant donné que 90 % des plantes sauvages et 75 % des cultures dépendent de la pollinisation, une bouchée sur trois de la nourriture que nous consommons dépend de leur survie. Sans les abeilles, un tiers de l'approvisionnement alimentaire mondial pourrait disparaître !

Voici cependant une bonne nouvelle. Grâce aux innovateurs en IA, nous disposons désormais de ruches autonomes alimentées à l'énergie solaire, dotées de vision par ordinateur et de robotique, qui aident les apiculteurs à surveiller et à thermoréguler leurs ruches, à réduire les parasites de façon biologique, à augmenter les taux de survie des abeilles et à diminuer la quantité de main-d'œuvre humaine nécessaire pour prendre soin des abeilles.

Si d'ici 2030 les abeilles sont florissantes et que la production alimentaire croît au rythme nécessaire, cela sera peut-être en grande partie attribuable au pouvoir de l'IA d'accélérer les changements positifs. Mais cela démontrera également notre intelligence collective humaine, alors que nous développons des solutions ingénieuses qui mettent l'IA au service des écosystèmes et du bien-être humain.

C COMME CONSIDÉRATION POUR LES KANGOUROUS ET TOUTES LES ESPÈCES (Y COMPRIS LES HUMAINS) TOUCHÉES PAR L'IA



Lorsque Volvo a commencé à tester ses premières voitures autonomes en Australie elles ont rencontré 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 dépourvue de perception intuitive de la profondeur, les kangourous posaient un problème différent. Lorsque les kangourous sont en plein saut, ils semblent plus loin qu'ils ne le sont réellement. Lorsqu'ils atterrissent, ils paraissent plus proches. L'IA n'avait aucun moyen d'expliquer ou de corriger cette confusion. C'est aux humains qu'il incombait de trouver une solution et de fournir suffisamment de données sur les kangourous pour remettre l'IA sur la bonne voie.

Mais tous les défis ne sont pas aussi évidents pour l'œil humain que le problème d'identification des kangourous. L'IA ingère de nombreuses données et reconnaît des schémas complexes dans ces données, mais elle ne comprend pas vraiment l'information et ne peut pas fournir de contexte.

Par exemple, l'assistante IA Alexa peut apprendre à reconnaître une commande telle que "commande-moi une maison de poupée." Mais elle pourrait ne pas être en mesure de distinguer si c'est un parent ou un enfant qui commande la maison de poupée. Cela s'est effectivement produit dans une famille, lorsqu'il a été découvert que leur petite fille avait pu commander pour elle-même une maison de poupée très coûteuse. Quand l'histoire a été racontée un jour à la radio locale par un animateur qui 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.

Autrefois, les ordinateurs suivaient des règles créées par des êtres humains. Mais aujourd'hui, grâce aux progrès de l'"apprentissage automatique," l'IA est pilotée par les données. Dans une large mesure, elle crée ses propres règles en fonction de ce qu'elle déchiffre dans ces données. Nous en apprendrons davantage sur ce processus plus tard.

La valeur de l'IA réside dans sa capacité à traiter d'énormes quantités de données. Cependant, cette capacité même présente 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 impossibles à suivre pour un humain, et même pour un système d'IA.

On appelle cela le problème de la "boîte noire."

Vous avez peut-être entendu parler de la boîte noire dans les avions ou les voitures, consultée après un accident. Ces boîtes enregistrent toutes les activités et commandes précédant un accident afin que nous puissions tirer des leçons de ce qui s'est mal passé.

Si seulement l'IA avait une boîte similaire! Malheureusement, la quantité de données que les algorithmes d'apprentissage automatique génèrent signifie que, s'il existait réellement une boîte noire dans un système d'IA que nous pourrions ouvrir, elle n'émettrait probablement qu'un bruit blanc indéchiffrable.

Le problème de la boîte noire en IA est qu'il n'y a, effectivement, pas de boîte noire.

Les algorithmes d'apprentissage automatique traitent tellement d'informations que l'IA est quelque peu mystérieuse pour nous et même pour elle-même. De notre point de vue, tout ce dont nous pouvons être vraiment sûrs, c'est quelles données entrent dans les algorithmes et quelles données en ressortent. Cela signifie que l'IA nécessite toujours 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 également faire beaucoup pour rendre le monde encore plus dangereux, déroutant et inéquitable qu'il ne l'est déjà.

Ce qui nous amène à une autre lettre....

D COMME DONNÉES, DÉCISIONS ET DIVERSITÉ



En fin de compte, l'IA ne sera jamais meilleure ni plus digne de confiance que la qualité, la diversité et l'explicabilité des données sur lesquelles elle est entraînée. De plus, elle n'aura de valeur qu'en fonction de notre capacité à prendre de bonnes décisions quant à la façon dont elle devrait, ou peut-être 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. C'est à nous de décider comment l'IA s'inscrit dans cet objectif !

POUR COMMENCER

Qu'est-ce qu'une amorce ?

D'abord, ce qu'elle n'est pas : une amorce n'est pas un guide complet. Ce n'est pas non plus un manuel scolaire qui vous dira ce qui sera à l'examen. Ce n'est pas ici que vous trouverez un glossaire de termes (bien que cette amorce vous oriente vers d'autres ressources qui en contiennent).

Une bonne amorce est un guide convivial qui facilite les premiers pas. Elle s'appuie sur les connaissances existantes et fournit juste assez de nouvelles connaissances pour vous mettre sur la bonne voie. Pensez à une amorce classique de littérature qui utilise le répertoire d'images et la mémoire orale d'un enfant pour associer des mots simples à ces images, afin que l'enfant puisse commencer le processus de décodage de nouvelles connaissances.

Mais une amorce destinée aux personnes qui ont déjà passé du temps à vivre et à apprendre dans le monde doit faire autre chose. Elle doit également identifier les connaissances et les suppositions existantes qui pourraient entrer en conflit avec de nouvelles façons de voir les choses. Pensez à un autre sens du mot amorce : une couche de peinture que vous devez appliquer pour qu'une nouvelle couleur apparaisse comme vous le souhaitez et ne soit pas déformée par les couleurs ou les motifs qui s'y trouvaient auparavant.

Surtout dans le monde de l'IA, nous devons être conscients de nos suppositions. Nos suppositions sur la société et la réalité s'infiltreront dans la façon dont nous créons et utilisons l'IA, et peuvent devenir des biais qui diminuent à la fois l'utilité et l'équité de l'IA.

La chercheuse du MIT Joy Buolamwini a découvert cela lorsque les algorithmes de reconnaissance faciale par IA avec lesquels elle travaillait n'ont pas réussi à reconnaître son visage. Les algorithmes n'avaient pas été entraînés sur suffisamment de photos de femmes noires. Les programmeurs qui ont entraîné l'IA ont supposé que les immenses banques de données de photos avec lesquelles ils avaient entraîné leur IA reflétaient la diversité du monde. Mais ils avaient tort. Aucun système automatisé n'est intrinsèquement neutre. Tout programme automatisé, comme toutes les choses que les gens créent, reflète les connaissances et les suppositions de la personne qui le crée. Si les créateurs d'une IA ont des suppositions fausses et inconscientes sur les gens, l'IA peut ancrer ces erreurs dans des systèmes plus vastes qui touchent des milliers, voire des millions de personnes.

La citoyenneté à l'ère de l'IA commence par un état d'esprit qui consiste à affronter et à atténuer intentionnellement nos biais, surtout lorsque les systèmes qui nous entourent pourraient les amplifier.

À QUI S'ADRESSE CETTE AMORCE ?

Cette amorce est rédigée pour les adolescents, les éducateur-trice-s et les adultes curieux à propos de l'IA, dans des contextes formels et informels. Elle les aidera à :

1. Reconnaître ce qu'ils savent déjà.
2. Acquérir une bonne compréhension conceptuelle de ce qu'ils ont encore à apprendre, et peut-être à désapprendre.
3. Essayer certaines choses pour voir comment l'IA fonctionne de première main.
4. Se sentir habilités à partager ce qu'ils ont appris avec les autres.
5. S'amuser !

Une fois que cette amorce vous aura aidé à conceptualiser le fonctionnement de l'IA, elle vous orientera vers des activités pratiques amusantes et d'autres voies 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. L'IA est utilisée partout dans le monde pour créer de l'art, des médias et d'innombrables autres projets inventifs et collaboratifs. L'IA peut rendre la vie plus colorée, inspirante et pratique. Elle peut nous donner plus de temps pour être de meilleurs amis, parents, enfants et citoyens. L'IA a le potentiel d'être un outil de co-création. Mais ce résultat dépend de l'équilibre entre la confiance et la collaboration, dans le respect des perspectives, de la vie privée et du niveau de confort technologique des autres.

Avant de commencer : N'oubliez pas la boîte noire !

Peu importe à quel point vous vous amusez ou réussissez avec l'IA, n'oubliez jamais ces kangourous ! Utilisez-les pour vous rappeler que, peu importe ce que nous apprenons, un problème demeure : l'IA ne peut 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 leur surveillance alors que nous transitons vers une société de plus en plus dépendante des machines intelligentes.

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

Laissons l'IA vous expliquer cela :



ChatGPT comprend-il les questions ?



Oui, ChatGPT est capable de comprendre et de répondre aux questions. Il s'agit d'un modèle de langage qui a été entraîné sur un vaste ensemble de données textuelles et qui peut générer des réponses semblables à celles d'un humain pour une grande variété de requêtes. Si vous avez des questions, n'hésitez pas à les poser.



ChatGPT comprend-il ce qu'il écrit ?



ChatGPT, comme les autres modèles de langage, n'a pas la capacité de comprendre le sens du texte qu'il génère. Il est entraîné à prédire le mot suivant dans une phrase en se basant sur les modèles et régularités qu'il a appris à partir des données textuelles sur lesquelles il a été entraîné. Il peut donc produire des réponses cohérentes et souvent proches de celles d'un humain, mais il n'a pas de compréhension réelle de ce qu'il écrit.



Mais comment puis-je apprendre sur l'IA si je ne peux peut-être jamais retracer le processus qu'elle utilise pour arriver à ses réponses?

Tout revient à l'attitude. Le fait que vous ne puissiez pas tout savoir ne signifie pas que vous ne pouvez pas commencer à apprendre et à continuer. Vous ne comprenez peut-être pas exactement comment les plantes poussent, mais cela ne devrait 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 conceptuelle de son fonctionnement. Vous découvrirez les meilleures questions à poser à l'IA que vous utilisez. Apprendre à poser les bonnes questions est essentiel pour tous les citoyens, qu'ils soient étudiants, non-experts ou même experts, car cela nous aide à saisir la portée des problèmes que l'IA peut résoudre. Cela nous aide à évaluer les risques que pose l'IA. Sachant cela, nous pouvons décider quels problèmes pourraient encore mieux être résolus à l'aide de l'intelligence humaine naturelle, de la pensée critique, du dialogue interpersonnel et d'autres méthodes traditionnelles de résolution de problèmes.

Nous apprenons sur l'IA afin de prendre les meilleures décisions quant aux activités et aux conditions dans lesquelles il est approprié de l'utiliser. Et nous pourrions bien constater que certaines activités et certaines conditions ne sont pas adaptées. Si l'IA fait des erreurs dans ses recommandations de films, il est peu probable que nous décidions de ne pas l'utiliser, tant qu'elle est généralement fiable. En revanche, 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 sur mille, ce n'est pas une IA que nous pourrions utiliser de manière éthique sans veiller à ce qu'il y ait une réglementation et une surveillance humaines importantes et continues.

Commençons donc par le premier concept que nous devons comprendre lorsque l'on parle d'IA dans le monde où nous vivons: l'apprentissage automatique.

Le secret pour apprendre l'apprentissage automatique

D'une façon très importante, vous savez probablement déjà comment fonctionne l'apprentissage automatique!

Si vous avez déjà développé une préférence pour un type de nourriture particulier, vous l'avez probablement développée à peu près de la même façon que les machines développent leurs compétences d'apprentissage. La grande différence est que les machines ne ressentent pas les sensations par lesquelles vous construisez vos préférences alimentaires. Les machines peuvent apprendre à partir d'images, de sons, de chiffres et de nombreux autres types de données, mais ce n'est pas ce que nous entendons par "sensation."

Pour l'instant, examinons ce que nous avons en commun avec les ordinateurs: un "langage source" binaire qui nous aide à interpréter le monde. **Binaire** désigne tout système construit sur seulement deux valeurs. Pour les ordinateurs, ce langage source binaire est le schéma de 1 et de 0 à partir duquel tous les autres langages informatiques et codes sont construits. Les humains ont quelque chose de très similaire à un langage source binaire dans la façon dont le cerveau catégorise les sensations comme agréables ou désagréables. Simplement en aimant et en n'aimant pas des sensations, et en suivant le degré auquel nous aimons ou n'aimons pas les sensations, nos cerveaux ont créé des millions de réponses à notre monde. Aimer et ne pas aimer nous aide à naviguer dans les choix de nos environnements complexes.

À un niveau fondamental, c'est ainsi que fonctionne l'apprentissage automatique. C'est assez simple. Un ordinateur ingère des ensembles de données, étiquette les données, puis essaie de créer des règles à partir des schémas qu'il découvre.

Créer un modèle

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

Imaginez un aliment que vous aimez, un bol de crème glacée ou une fraise biologique bien mûre. 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 **“analytique prédictive”** ne provoquent pas le même niveau de salivation, mais c’est exactement ce que fait votre cerveau. Votre cerveau prédit et préressent le plaisir en fonction d’un souvenir stocké. 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 dans ces données. À partir de ces schémas, votre cerveau commence à développer des règles. Ces règles deviennent la base des comportements et des réponses automatiques de votre cerveau.

Données

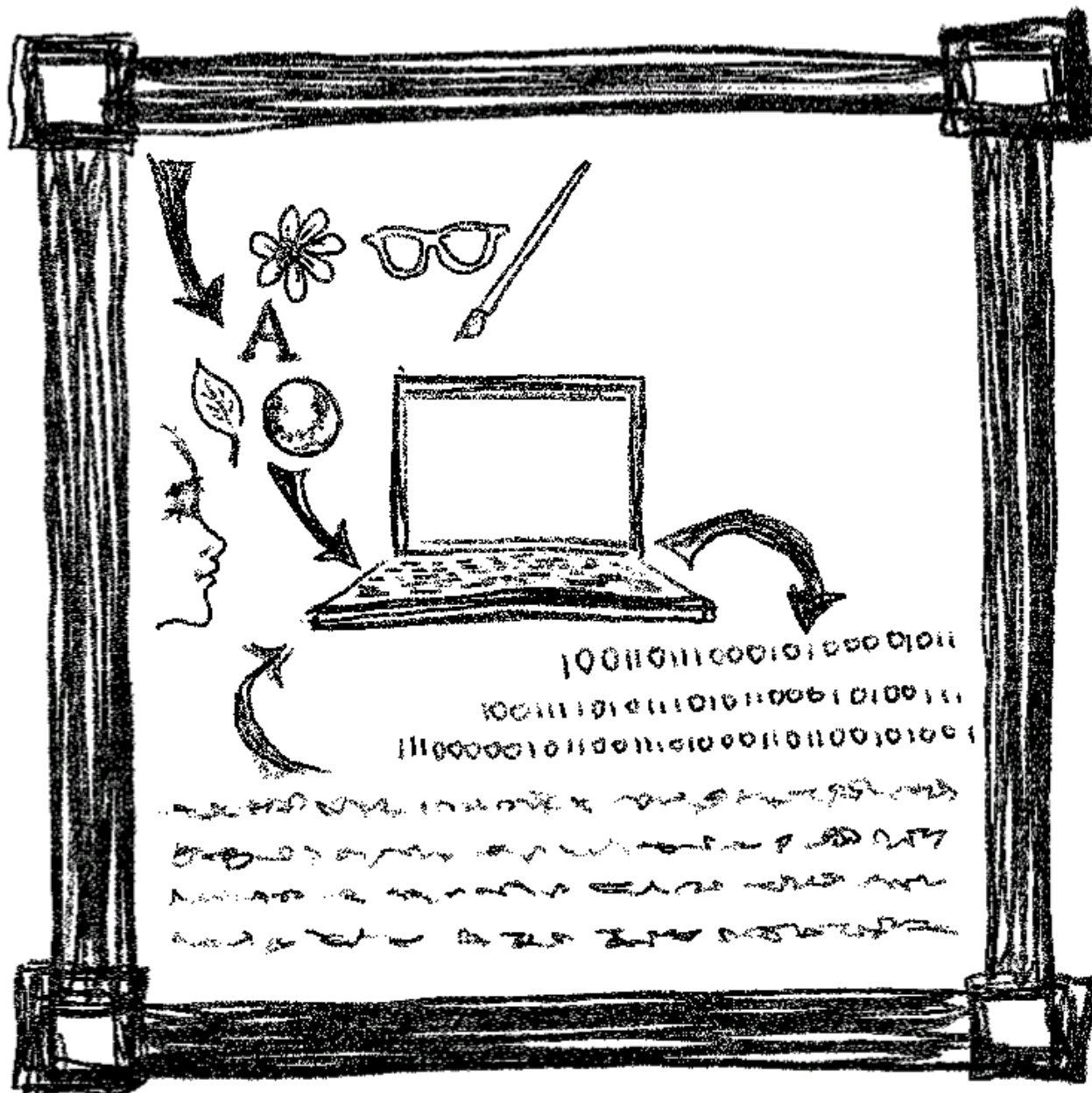
De la même façon que vous avez construit des prédictions sur ce que vous aimez tout au long de votre vie, les ordinateurs du monde entier font des prédictions sur votre comportement en fonction de votre activité en ligne. En moyenne, chaque jour, les humains dans le monde :

- Utilisent Google 4 millions de fois.
- Visionnent 4,5 millions de vidéos sur YouTube.
- Échangent 188 millions de courriels.

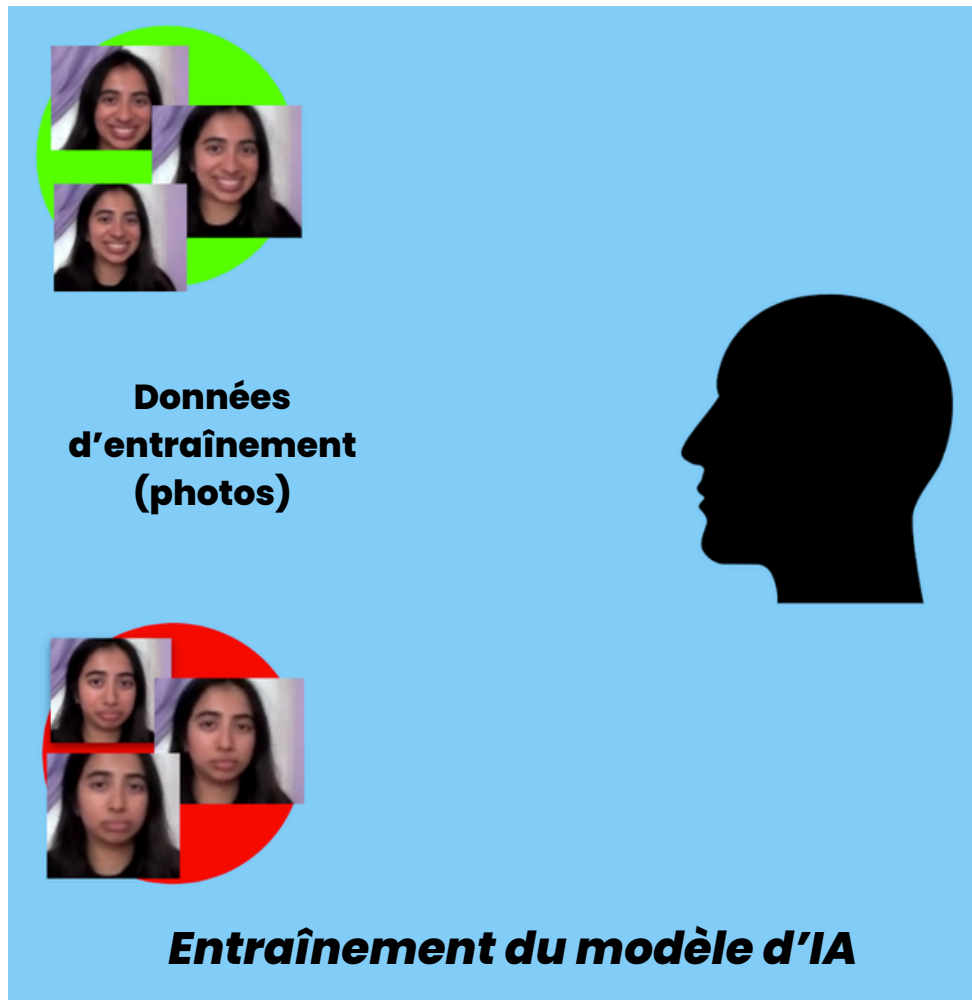
These exchanges produce a lot of data. Advances in chips and semi-conductors have greatly accelerated how easily and how fast computers can now work with large sets of data, which has been an important contributor to the development of machine learning. Computers sift through Big Données to make predictions about what you like or don’t like, and often will respond to those predictions with suggestions for stuff you might like to buy, watch, or consume.

Examinons ce processus de plus près :

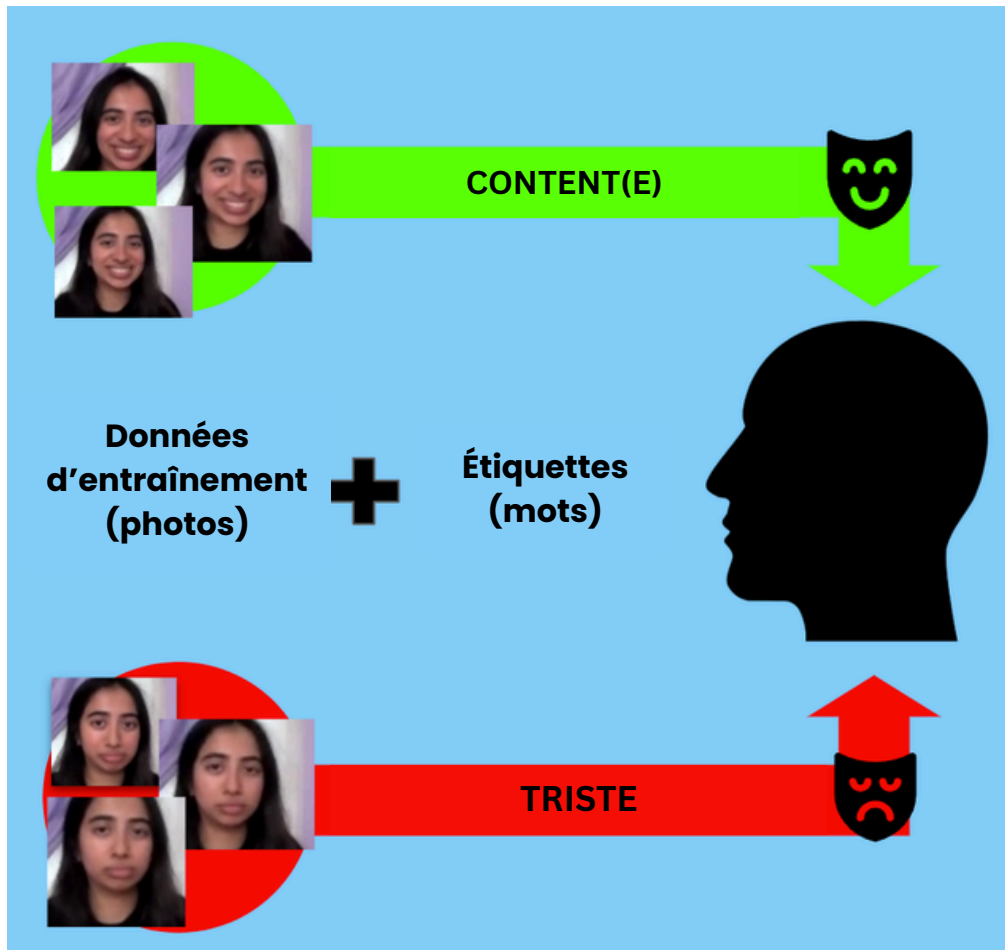
L'entraînement des algorithmes d'IA



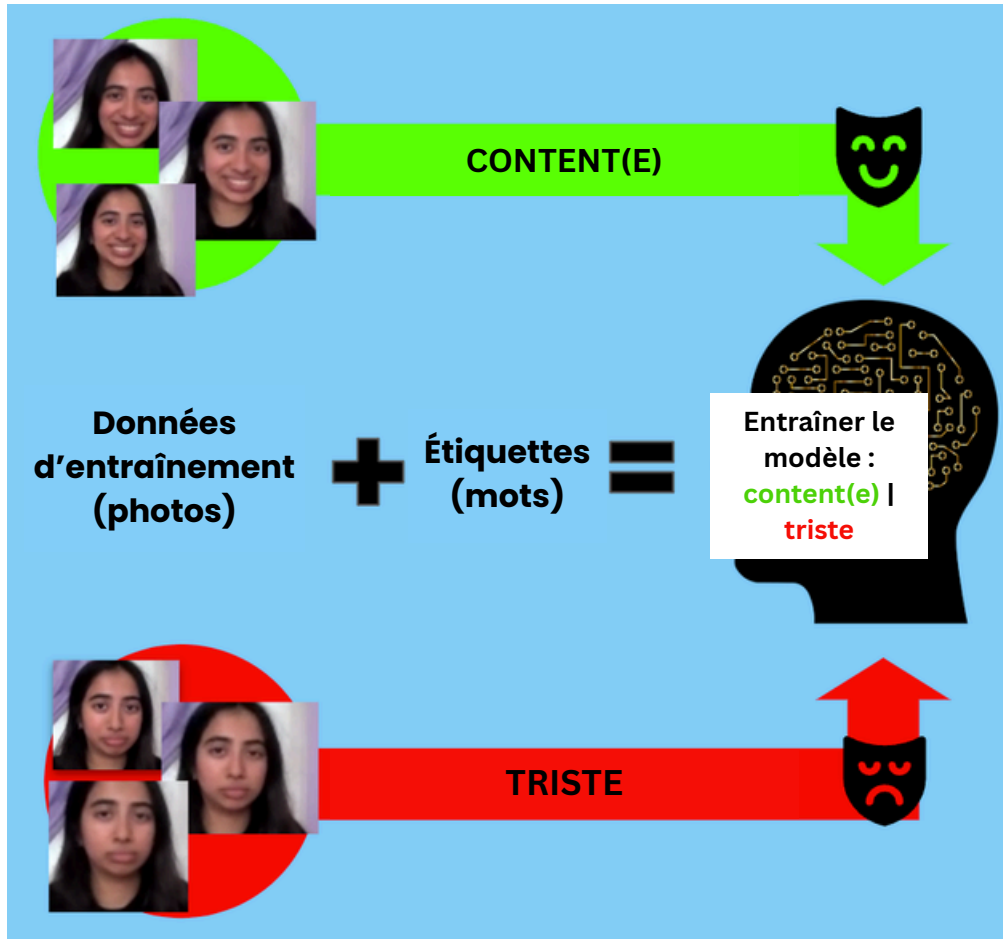
Pour mieux comprendre comment l'IA peut prédire ce que vous aimerez ou n'aimerez pas, examinons le processus par lequel une IA pourrait apprendre quels types de choses vous rendent heureux ou triste.



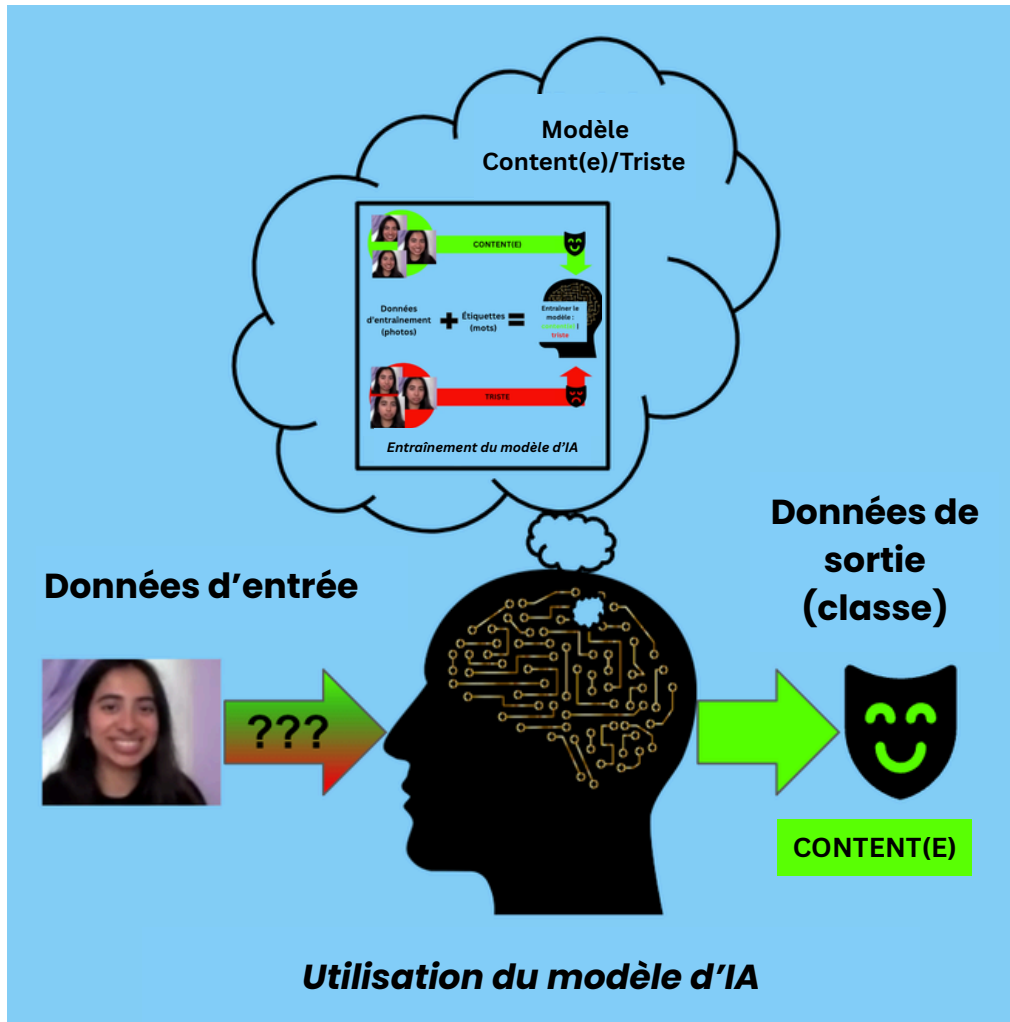
Avant que l'IA puisse commencer à prédire vos réponses, elle doit étiqueter les données qu'elle collecte. Dans ce cas, l'IA est entraînée sur des captures d'écran d'un visage pour distinguer les visages heureux des visages tristes. Elle aura besoin d'un nombre raisonnable, disons un minimum de 30, pour voir 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 étiquetées avec des variables, comme les mots "heureux" et "triste." L'IA se fiche des étiquettes (parce qu'elle n'est pas une personne capable de s'en soucier), elle a juste besoin de suffisamment de données pour pouvoir trier les objets dans les schémas des ensembles de données que nous lui fournissons. Avec suffisamment d'exemples de visages souriants étiquetés "heureux" et de visages abattus étiquetés "tristes," l'IA peut créer un modèle de travail pour tester ses performances avec des ensembles de données beaucoup plus grands. Bien qu'elle puisse créer un modèle préliminaire avec seulement 30 exemples, elle bénéficiera d'exemples beaucoup plus nombreux et diversifiés pour renforcer et affiner le modèle.



Avec ce modèle, l'IA peut ensuite "produire" son estimation de l'émotion que le visage exprime et tester si son estimation est correcte. Au fur et à mesure que l'IA ingère plus de données et teste ses propres prédictions, elle développe une précision croissante parce que ses choix sont basés sur un nombre de plus en plus grand d'exemples humains réels. L'algorithme qui permet à l'IA de construire son propre ensemble de règles s'appelle un "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, ce sont les humains qui établissaient les règles que les machines suivaient.

Un algorithme informatique traditionnel aurait pu être un programme qui fournissait des images qu'un humain avait déjà étiquetées, puis demandait à l'ordinateur de produire des réponses basées sur ces données pré-étiquetées.

Dans l'apprentissage automatique, nous pouvons étiqueter de grands ensembles d'images. Mais ensuite, nous laissons l'algorithme inférer les similarités, puis se tester lui-même pour savoir s'il étiquette correctement les nouvelles données.

Plus l'ordinateur est entraîné sur des données et en a appris, plus la machine elle-même crée les règles à partir du modèle qu'elle a créé.

Programmation traditionnelle



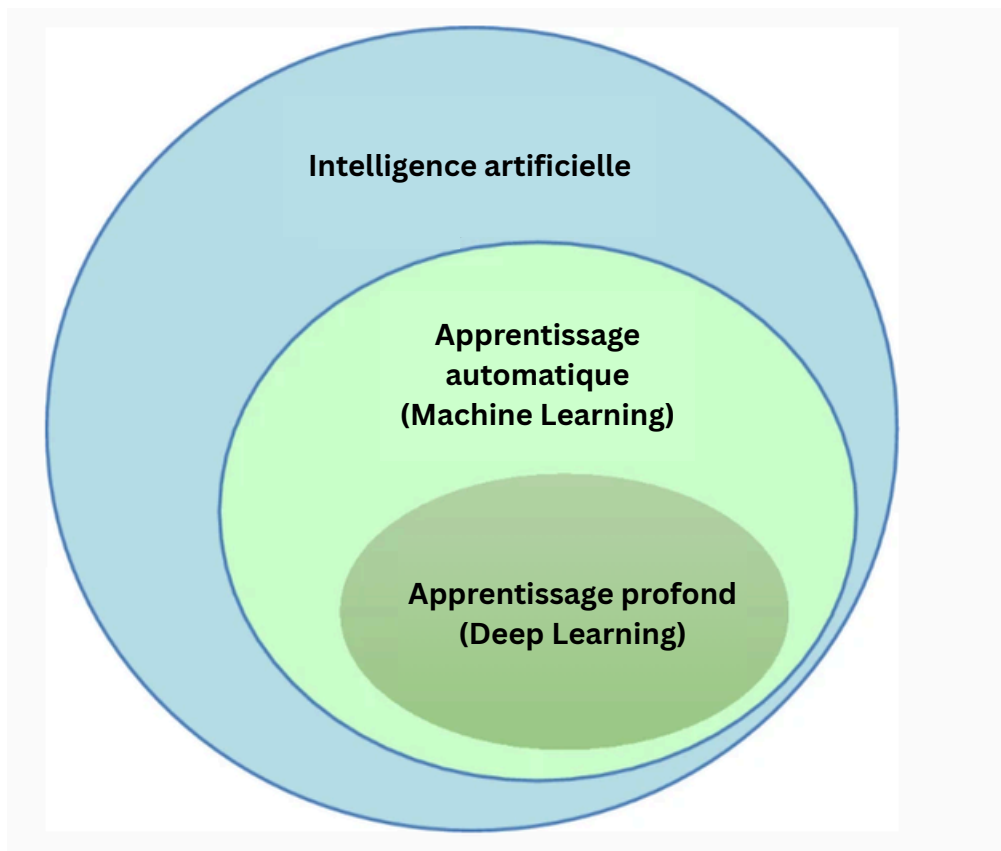
Apprentissage automatique (Machine Learning)



C'est important! Il sera toujours utile de garder à l'esprit que le seul fait qu'un algorithme informatique puisse créer de nombreuses règles ne signifie pas qu'il devrait créer toutes les règles. Cela ne signifie surtout pas que toutes les règles sont toujours correctes!

Le type d'apprentissage que fait l'IA est encore ce que nous appelons l'apprentissage "étroit." Contrairement aux humains, l'IA n'est pas équipée pour mettre son apprentissage en contexte de la façon dont les humains le font.

Plus les algorithmes d'IA sont axés sur les données, plus ils entrent dans l'état appelé "apprentissage profond." Ici, les couches de réseaux neuronaux sont nombreuses et complexes. Elles sont moins transparentes et plus difficiles à expliquer. Si une IA fait une rencontre tragique avec un kangourou, c'est probablement parce qu'il n'y a pas eu suffisamment d'implication humaine dans l'entraînement et la conception de l'algorithme pour toutes les différentes situations qu'un humain pourrait rencontrer.



Revenons à notre modèle "heureux/triste." Il y a de nombreuses raisons pour lesquelles l'IA pourrait encore se tromper ou faire des corrélations non pertinentes. Peut-être, par exemple, a-t-elle été entraînée sur des milliers de visages, mais ces visages n'étaient que caucasiens. Ou peut-être que ses utilisateurs provenaient d'une culture plus susceptible de réagir avec du malheur à quelque chose qui produisait du bonheur dans une culture dominante.

Les humains peuvent et doivent continuer à influencer les règles sur lesquelles fonctionne l'apprentissage automatique. Le secret d'une utilisation optimale de l'apprentissage automatique est 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 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.

Une façon de clarifier ce type de décisions est d'utiliser des "fiches de modèle." Ce sont des documents que les créateurs d'IA peuvent utiliser pour apporter de la transparence sur la façon dont un modèle d'IA est destiné à être utilisé, ses limites et avertissements, les données sur lesquelles il a été entraîné et les étiquettes et caractéristiques utilisées pour entraîner les données.

Voici, par exemple, une fiche de modèle créée pour ChatGPT, l'agent conversationnel impressionnamment compétent. Cette fiche de modèle a été introduite quelques mois après que la première vague d'utilisateurs a réalisé que ChatGPT n'était pas toujours aussi compétent qu'il le semblait au premier abord :

ChatGPT



Exemples

« Explique l'informatique quantique en termes simples » →

« As-tu des idées créatives pour l'anniversaire d'un enfant de 10 ans ? » →

« Comment faire une requête HTTP en JavaScript ? » →



Capacités

Se souvient de ce que l'utilisateur a dit plus tôt dans la conversation.

Permet à l'utilisateur d'apporter des corrections ou des précisions par la suite.

Entraîné à refuser les demandes inappropriées.



Limites

Peut occasionnellement générer des informations incorrectes.

Peut occasionnellement produire des instructions nuisibles ou un contenu biaisé.

Connaissances limitées du monde et des événements postérieurs à 2021.

Obstacles invisibles

Les fiches de modèle sont l'une des façons dont un créateur d'IA peut aider un utilisateur à déterminer si cette IA convient au problème qu'il veut résoudre ou à l'artefact qu'il veut créer.

Mais tous les systèmes pilotés par l'IA qu'un citoyen vivant à l'ère numérique rencontrera ne sont pas accompagnés d'une fiche de modèle. L'un des défis les plus répandus de la vie à l'ère de l'IA est d'être conscient des façons dont l'IA fonctionne de manière invisible ou obscure pour vous. Lorsqu'une plateforme de diffusion en continu vous fait des recommandations, ou lorsque 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.

Chaque année, l'IA devient meilleure pour prédire ce que les humains feront ou diront. Si vous avez essayé un outil d'IA conversationnelle comme ChatGPT, vous avez vu à quel point l'IA peut efficacement assembler les phrases qu'elle croit constituer de bonnes réponses à vos questions.

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

L'une de nos tendances humaines les plus universelles est l'anthropomorphisme, 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 réellement, il est donc facile pour nous de supposer que l'IA dans nos vies a des pouvoirs humains, voire magiques.

Mais selon toute vraisemblance, il y aura toujours des choses que les humains font mieux que l'IA, surtout en ce qui concerne la compréhension du contexte et ce qu'on appelle l'apprentissage "large." Nous sommes encore meilleurs pour établir le type de connexions visuelles et spatiales qui nous aident à juger à quelle distance se trouve un kangourou de notre voiture, même si nous ne sommes pas australiens!

À mesure que les agents conversationnels deviennent plus courants dans le monde, il deviendra de plus en plus important de savoir quand vous interagissez avec l'un d'eux.

Nous ne savons pas si l'IA remplacera les humains. Cela ne semble pas probable, quelle que soit la qualité avec laquelle une machine peut être programmée pour imiter les humains. Il est certainement juste de dire que l'IA pousse les humains à devenir plus intelligents et encore plus compétents en aidant les humains à s'améliorer dans les tâches d'apprentissage répétitives pour lesquelles nous ne sommes pas aussi bien adaptés.

Mais cela nous met également sous pression pour démontrer notre capacité supérieure à faire le type de travail que l'IA ne peut pas faire, comme penser de façon critique aux situations, prendre des décisions sur ce que nous valorisons, comprendre pourquoi nous valorisons ces choses et nous préoccuper du bien-être des autres.

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

L'une des plus grandes opportunités présentées 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. Au fur et à mesure que nous devenons meilleurs pour connaître et juger ces différences, nous deviendrons meilleurs pour prendre des décisions sur la façon d'utiliser l'IA de manière éthique et efficace.

Apprendre sur l'IA peut enseigner aux humains plusieurs choses sur leur propre intelligence :

1. **Limites:** L'une des leçons les plus importantes que nous pouvons tirer de l'IA est qu'il existe des limites inhérentes à notre propre intelligence. Les modèles d'IA peuvent effectuer des calculs complexes à une vitesse fulgurante et peuvent se souvenir de vastes quantités de données avec facilité, ce qui peut mettre en évidence les limites de nos propres capacités cognitives. Cela peut nous aider à identifier où nous devons nous améliorer et où nous devons compter sur la technologie pour nous aider.
2. **Résolution de problèmes:** Les algorithmes d'IA sont conçus pour résoudre des problèmes spécifiques, et étudier comment ils le font peut nous enseigner 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 défis, que ce soit en affaires, en sciences ou dans la vie quotidienne.
3. **Biais:** Les algorithmes d'IA ne sont pas plus impartiaux que les données qui leur sont fournies. Si les données contiennent des biais, ceux-ci se refléteront dans les résultats. En étudiant l'IA, nous pouvons apprendre les risques liés aux biais et comment les atténuer dans notre propre prise de décision.
4. **Créativité:** Bien que les algorithmes d'IA deviennent de plus en plus sophistiqués, ils manquent encore de la créativité et de l'imagination que possèdent les humains. Cela peut servir de rappel des qualités uniques qui nous rendent humains, comme notre créativité, notre intuition et notre intelligence émotionnelle.

Dans l'ensemble, l'étude de l'IA peut nous aider à comprendre à la fois les forces et les limites de notre propre intelligence, et peut nous inspirer à trouver de nouvelles façons innovantes de résoudre des problèmes et de créer de la valeur.

E COMME ENTHOUSIASME ET LA FIN DE CETTE AMORCE (MAIS SEULEMENT LE DÉBUT DE VOTRE PARCOURS)

Bien que les progrès de l'IA semblent s'accélérer, la route vers une utilisation à son plein potentiel sera probablement jalonnée de débats sur ses utilisations et ses dangers, et d'une prise de conscience croissante des compétences que nous devons maîtriser pour mieux la comprendre. Nous aurons besoin de savoir plus que simplement comment interagir avec l'IA, et devons développer des compétences de base dans des littératies émergentes comme la littératie des algorithmes et des données, la modélisation et l'apprentissage automatique.

L'enthousiasme à propos de l'IA peut souvent se transformer en anxiété, voire en désillusion et en désespoir. Une partie de l'évolution de l'IA sera sous notre contrôle, une autre partie ne le sera pas. C'est pourquoi il est important d'ancrer notre enthousiasme dans la joie d'apprendre.

Dans les années à venir, il y aura tant à apprendre sur l'IA, ce qu'elle peut et ne peut pas faire. Il y aura également tant à apprendre sur l'espèce humaine, ce qu'elle ne peut pas faire, mais aussi des choses que nous pouvons maintenant faire et qui étaient autrefois au-delà de notre capacité à imaginer.

Alors que nous nous tournons vers l'avenir, pour maintenir notre enthousiasme, il pourrait être utile de regarder vers le passé et de penser au principe des sept générations employé 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 précédés, et tout ce que nous aspirons pour les générations qui nous suivront, nous pouvons au moins savoir que nous avons fait ce que nous pouvions pour assurer que l'IA deviendra et restera une partie vitale d'une planète florissante dans laquelle toutes les espèces ont de la valeur.

RESSOURCES

[The Algorithm and Données Literacy Project](#)

["L'IA et l'apprentissage créatif: préoccupations, opportunités et choix"](#) - Mitch Resnick

[L'apprentissage automatique en une journée](#) - Apps for Good

[Le guide d'IA d'Actua](#)

[Danser avec l'IA](#)

[Un guide de référence rapide sur l'intelligence artificielle pour les enseignant·e·s](#)