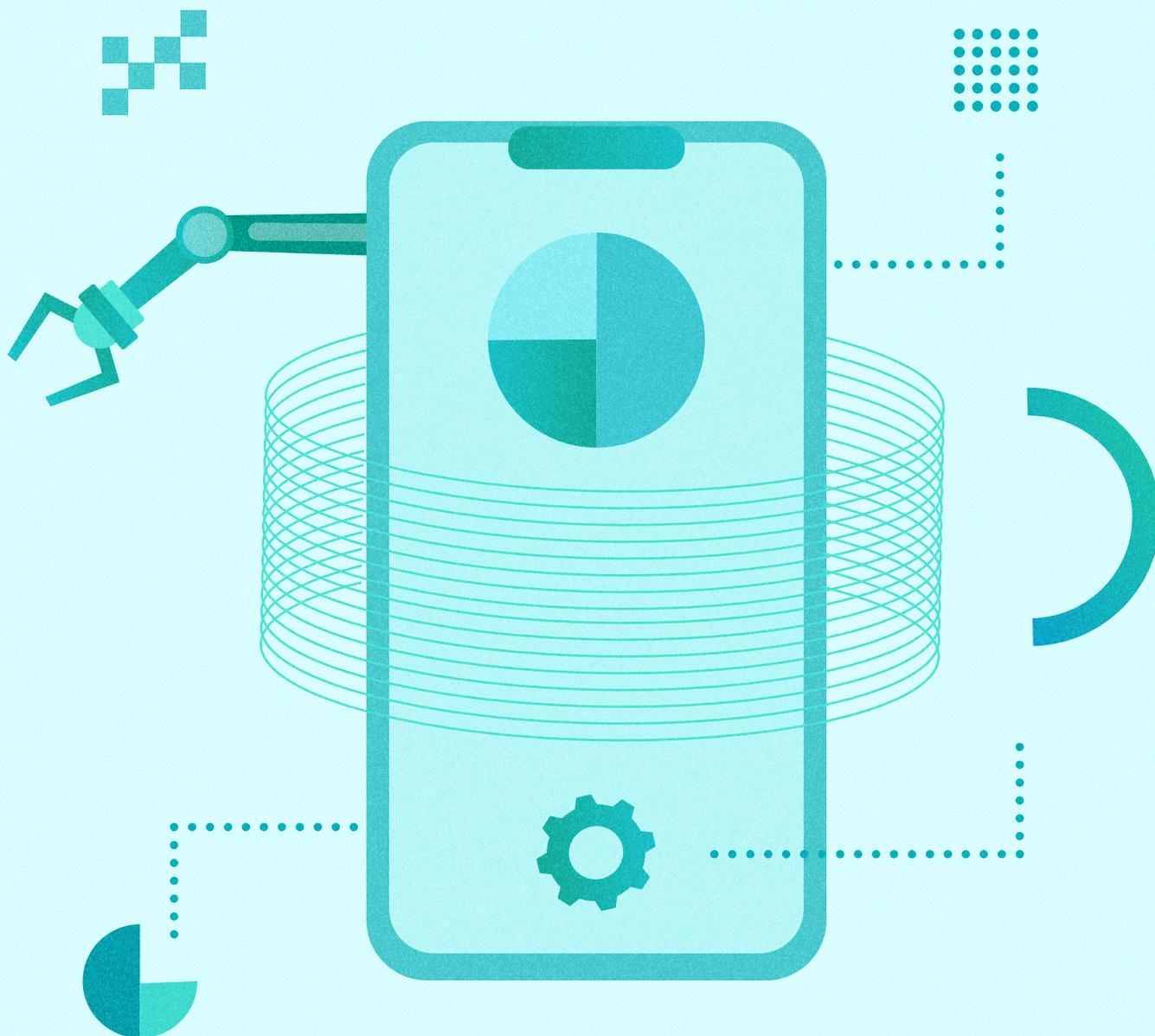


IA ET ENJEUX DE L'AUTOMATISATION

Cette première partie du dossier IA et enjeux de l'automatisation propose une contextualisation historique de l'émergence de l'intelligence artificielle en présentant quelques moments clés de la discipline. Il est complété par un état des lieux concernant les différents usages et applications actuels des technologies de l'apprentissage automatique.



Objectifs

- Comprendre le contexte d'émergence de l'intelligence artificielle
- Différencier les courants symbolique et connexionniste
- Se rendre compte de la pluralité des applications de l'IA
- Identifier des situations du quotidien où l'IA intervient

Enjeux

Depuis la moitié du XX^e siècle, les scientifiques essaient de simuler le raisonnement grâce aux machines. L'histoire de l'intelligence artificielle est d'abord marquée par de grandes ambitions et connaît successivement plusieurs phases d'engouement et de déceptions. Elle se nourrit et nourrit elle-même des récits de science-fiction, dont les personnages artificiels, robots, androïdes ou simples voix de synthèse, influencent à leur tour les imaginaires collectifs. En résultent de nombreuses croyances, des craintes, mais également des espoirs, qui ne correspondent pas toujours à l'état des avancées technologiques actuelles.

Les récentes performances dans le domaine de l'apprentissage automatique appellent à une réflexion sur la définition de l'IA en tant que catégorie générale d'une part, et sur les particularités des technologies qu'elles englobent d'autre part. Un retour aux sources du projet visant à simuler le raisonnement permettra également de comprendre les motivations, les ambitions mais aussi les limites de certaines applications de l'IA.

IA, une définition

L'intelligence artificielle est un terme à la mode, fréquemment évoqué dans le débat public et dans les médias. Malgré cette popularité certaine, rares sont les tentatives d'expliquer précisément ce qu'est l'IA. Il s'agit en fait d'une notion générique qui désigne différentes technologies du traitement automatique de l'information, ayant pour objectif de simuler le raisonnement. Aujourd'hui, l'IA fait principalement référence à l'**apprentissage automatique**, aussi connu dans sa version anglaise de *machine learning*.

Les technologies de l'IA partagent des caractéristiques communes. Elles sont développées par des informaticiens et visent à imiter certaines fonctions cognitives propres à l'être humain: l'apprentissage de règles, le raisonnement logique, le langage, la planification ou encore la reconnaissance d'images ou de sons. Pour ce faire, une modélisation de ces processus est nécessaire à l'aide des mathématiques, de l'algorithmique et de la programmation.

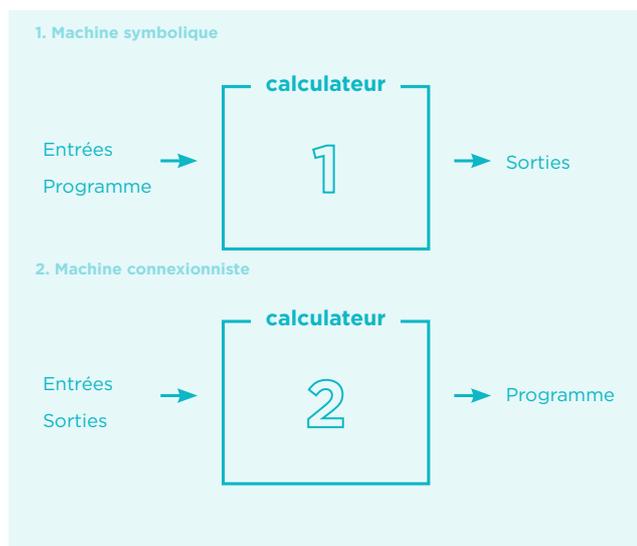
Un contexte socio-technique favorable a permis aux premières tentatives de simulation de l'intelligence de voir le jour. Dans les années 1940, le mathématicien Norbert Wiener développe la **cybernétique**, un courant de pensée permettant d'appréhender conjointement le fonctionnement du vivant et des machines à partir des théories de la communication. Tous deux auraient les mêmes capacités d'adaptation basées sur des capteurs percevant leur environnement. Très populaire au sein de l'élite scientifique américaine, cette approche pose les bases théoriques nécessaires à l'émergence de l'IA.

Sur le terrain, le développement des premiers calculateurs électroniques durant la Seconde Guerre mondiale permet de tester le potentiel de la cybernétique. Deux approches principales vont se distinguer et s'opposer au fil de l'histoire, avec d'un côté l'**IA connexionniste** et de l'autre l'**IA symbolique**.

Un peu d'histoire

L'histoire de l'IA est principalement composée de deux courants distincts qui ambitionnent de simuler le raisonnement humain grâce aux machines. Longtemps ostracisée par l'approche symbolique, l'IA connexionniste s'est aujourd'hui imposée avec les performances de l'apprentissage automatique et des **réseaux de neurones**, inspirés des connaissances du cerveau humain.

En 1955, les mathématiciens américains John McCarthy et Marvin Minsky organisent le premier projet de recherche sur l'intelligence artificielle et fondent l'année suivante le domaine scientifique de l'IA. Leurs ambitions sont clairement énoncées : ils s'appuient sur la conjecture que l'apprentissage et l'intelligence sont des processus suffisamment compris pour qu'ils puissent être reproduits artificiellement. Ils envisagent ainsi de modéliser le raisonnement grâce à des manipulations mathématiques et logiques de symboles.



Fonctionnement d'une machine symbolique (1) et connexionniste (2). En haut, les informations nécessaires aux calculateurs sont fournies a priori afin d'obtenir les résultats. En bas, la machine connexionniste déduit le modèle en fonction des entrées et des sorties.¹

Les **systèmes experts**, qui connaissent un important succès dans les années 1980,

appartiennent à ce courant. Trop complexes, ils offrent des résultats peu satisfaisants et sont abandonnés en une décennie. Débute alors une période creuse en terme d'avancées scientifiques, parfois appelée **hiver de l'IA**. Aujourd'hui, les systèmes experts sont toujours utilisés dans des domaines tels que les jeux vidéo, mais ils ne font plus partie du domaine scientifique de l'IA.

Depuis le milieu des années 2000, les performances de l'approche connexionniste permettent un retour en force de l'IA. Contrairement à l'IA symbolique, elle s'inspire directement du fonctionnement du cerveau humain. Ce projet prend racine dès 1943, grâce aux neuroscientifiques américains Warren McCulloch et Walter Pitts, qui proposent pour la première fois un modèle mathématique représentant un neurone. Le terme n'existe pas encore à l'époque, mais cette date marque le début de l'approche connexionniste de l'IA. Sur la base des connaissances en biologie, les neurones formels vont être mis en connexion afin de former des réseaux et de simuler l'apprentissage. Pour ce faire, les neurones captent les signaux de leur environnement et établissent des modèles prédictifs sur une base d'entrées et de sorties (cf. Schéma ci-contre).

En 1989, l'ingénieur français Yann LeCun s'inspire de la métaphore des neurones et crée un réseau multicouche capable de reconnaître automatiquement des codes postaux manuscrits. Cette technique sera reprise par le secteur bancaire afin de lire automatiquement des chèques. Il s'agit de l'une des premières utilisations à grande échelle d'une technique d'apprentissage automatique basée sur des réseaux de neurones formels.

Près de trois décennies plus tard, l'optimisation des réseaux de neurones et des algorithmes, associée à l'augmentation considérable de la puissance de calcul des ordinateurs et aux importants volumes de données disponibles (*big data*), replacent l'IA au centre des intérêts scientifiques et technologiques. En effet, les masses de données sont un élément essentiel à l'entraînement des modèles d'apprentissage

¹ Schéma adapté de : Cardon, D., Cointet, J. & Mazières, A. (2018). La revanche des neurones: L'invention des machines inductives et la controverse de l'intelligence artificielle. *Réseaux*, 211, 173-220.

LA SINGULARITÉ TECHNOLOGIQUE

La singularité technologique correspond au moment hypothétique où l'IA dépasserait l'intelligence humaine. Plusieurs scientifiques dont Ray Kurzweil, Stephen Hawking et Elon Musk ont fait part de leurs inquiétudes quant aux dangers potentiels d'une technologie qui deviendrait supérieure aux humains.

La théorie de la singularité est basée sur la loi de Moore, qui postule l'évolution exponentielle de la puissance de calcul des ordinateurs. En effet, depuis les premiers microprocesseurs des années 1970, on observe que le nombre de transistors double environ tous les deux ans. Les défenseurs de la singularité partent du principe que cette croissance exponentielle continuera jusqu'à ce que les machines soient elles-mêmes capables de programmer d'autres machines.

Or, ces prédictions omettent plusieurs facteurs: il paraît aujourd'hui difficile de penser que la croissance technologique se poursuive indéfiniment. Les limites techniques ne sont, par exemple, pas prises en compte. Les détracteurs de la singularité considèrent que ces prédictions relèvent plus de la science-fiction que de projections scientifiques fiables.

👉 Comment ça marche ?

Pour établir des modèles prédictifs efficaces, les réseaux de neurones de l'IA connexionniste sont entraînés par des algorithmes à traiter et agréger d'importantes quantités de données. On parle d'apprentissage automatique puisque les algorithmes établissent des modèles à partir des données qui leur sont fournies. Il existe plusieurs types d'apprentissages.

L'**apprentissage supervisé** requiert l'intervention humaine pour labelliser les données,

corriger les erreurs lors des phases d'entraînement et obtenir les modèles souhaités. Par exemple, pour apprendre à reconnaître un chat, un algorithme considère des milliers d'images préalablement étiquetées "chat" ou "non-chat" afin d'établir un modèle prédictif qui pourra être appliqué à des images inconnues. Contrairement au cerveau humain, qui n'a besoin que de quelques exemples pour reconnaître une forme, l'IA connexionniste nécessite des volumes importants de données pour établir ses modèles. L'**apprentissage non supervisé** consiste à laisser l'algorithme déterminer ses propres modes de classification et catégories sans intervention externe.

En 2016, le programme AlphaGo de DeepMind (Google) bat pour la première fois un joueur professionnel au jeu traditionnel de Go. Une combinaison des différentes méthodes prenant en compte des parties disputées par des joueurs humains et d'**apprentissage par renforcement** uniquement basé sur l'expérience de la machine permet à AlphaGo de déduire les coups optimaux. La dernière version du programme baptisée AlphaGo Zero est parvenue à battre n'importe quel joueur humain, ainsi qu'AlphaGo lui-même, en ne connaissant que les règles du jeu.

Les performances actuelles de l'apprentissage automatique sont certes considérables et dépassent dans certains cas les capacités humaines. Leurs perspectives demeurent cependant limitées car chaque réseau de neurones est dépendant du contexte dans et pour lequel il est développé. Il n'est à ce jour pas possible d'envisager de les transférer dans un environnement différent sans un important travail de ré-apprentissage supervisé par des humains. Pour cette raison, l'IA connexionniste actuelle est souvent qualifiée d'**IA faible** ou étroite, en opposition au projet hypothétique d'**IA forte**, qui ambitionne de modéliser toutes les facettes de l'intelligence, y compris la conscience et les émotions.

Par ailleurs, l'IA connexionniste est fréquemment associée à un enjeu de transparence. En effet, malgré l'efficacité de certains modèles prédictifs, la complexité des réseaux de neurones crée l'effet d'une **boîte noire**, caractéristique de l'apprentissage automatique. Même les concepteurs des algorithmes ne sont pas en mesure d'expliquer précisément

les décisions de leurs programmes. Cet aspect est souvent présenté comme l'un des défis majeurs de l'IA contemporaine. Il est intéressant de souligner que l'enjeu de transparence concerne uniquement les prises de décision automatisées. En comparaison, les nombreuses décisions prises par les êtres humains au quotidien ne semblent pas concernées par cette injonction à la transparence. Pourtant, elles ne sont pas toujours explicables ni expliquées.

L'effet boîte noire ne remet cependant pas systématiquement en cause l'application de l'apprentissage automatique. L'impossibilité d'expliquer certains choix doit toutefois être prise en compte, en l'intégrant à des processus non automatisés. L'appréciation humaine demeure indispensable pour éviter des prises de décisions arbitraires.

Pluralité des applications

La difficulté à définir précisément l'IA réside en partie dans la pluralité de ses techniques et de ses applications. Les prochains exemples, non-exhaustifs, permettent de se rendre compte de la diversité des champs d'application de l'IA et de l'évolution de l'utilisation de certaines technologies.

Les débuts de l'IA sont étroitement liés aux institutions militaires américaines. Dans un premier temps, les technologies du traitement automatique de l'information reçoivent d'importants financements de la part de la marine (ONR) et du Département de la défense (DARPA). Elles comprennent les premiers programmes de traduction automatique, les logiciels de reconnaissance d'images ou encore la conduite autonome. Aujourd'hui encore, la maîtrise des technologies de l'IA est toujours considérée comme un défi central dans le secteur militaire à travers le monde.

Petit à petit, ces technologies ont été adaptées à des usages plus variés. La traduction automatique est désormais utilisée dans nos pratiques quotidiennes. La reconnaissance de formes et d'images est employée en médecine et contribue notamment à l'amélioration de la détection de certaines tumeurs. Google et Tesla tentent d'intégrer la conduite autonome au trafic. L'IA façonne également les contenus que nous voyons sur le Web grâce à ses algorithmes de hiérarchisation et de modération.



Lee Sedol lors du match contre AlphaGo, 2016

Ces exemples montrent que les motivations initiales à produire de nouvelles technologies peuvent être largement détournées en quelques décennies. Les technologies ont les mêmes fondements, mais les usages se sont diversifiés et banalisés. Désormais, elles sont le plus souvent développées et proposées par des entreprises privées. L'arrivée de l'IA dans nos quotidiens invite à réfléchir aux multiples possibilités d'automatisation et aux conséquences que ces dernières peuvent avoir sur les individus et la société.

L'IA et les pratiques en ligne

Avec l'avènement des médias sociaux, les entreprises proposant leurs services sur Internet ont accès à des volumes de données inédits concernant leurs utilisateurs. L'IA est utilisée pour gérer et trier ces importants flux d'informations.

Les moteurs de recherche, les médias sociaux et les plateformes de streaming récoltent et agrègent automatiquement les traces numériques laissées par les internautes afin de leur suggérer du contenu personnalisé. L'exploitation de ces traces permet, d'une part, de fidéliser l'internaute en lui proposant du contenu et des services personnalisés et, d'autre part, d'optimiser les placements publicitaires.

Plusieurs problèmes émergent face à l'exploitation massive des données personnelles. Ces

méthodes impliquent que des informations parfois sensibles telles que l'orientation politique, sexuelle ou religieuse d'une personne soient détenues par des entreprises privées. Elles touchent directement à la protection de la vie privée. Les interactions avec les contenus, du simple clic aux mentions « j'aime », sont également enregistrées et utilisées à des fins prédictives. L'ensemble des traces récoltées par les différentes plateformes sont exploitées par des algorithmes qui établissent des prédictions pour des individus aux comportements et aux goûts apparemment similaires.

Le traitement automatique des données crée ainsi des **bulles de filtre**, qui tendent à enfermer les internautes dans un monde construit uniquement sur leurs habitudes en ligne et les contenus qu'ils consultent. Comme les prédictions sont établies à partir des traces passées, les publications correspondant à ces informations seront suggérées. Bien que les bulles de filtre ne soient pas propres aux pratiques numériques, l'automatisation du traitement des données tend à les amplifier.

Les algorithmes sont également utilisés pour modérer les contenus illégaux ou qui ne respectent pas les conditions d'utilisation des plateformes. Cependant, la modération automatique des réseaux sociaux n'est pas toujours efficace et l'incapacité à contextualiser une information engendre des situations discriminantes. Ainsi, en avril 2019, Facebook a automatiquement empêché la création d'un compte de soutien LGBT+, sous prétexte que le mot « lesbienne » contrevenait aux règles de la plateforme. Plus généralement, la modération automatique n'est pas assez performante pour localiser et gérer la totalité des contenus problématiques de manière satisfaisante. Elle engendre un risque de censure injustifiée, spécialement marqué pour des communautés déjà sous-représentées. Le principe de liberté d'expression est ainsi menacé par la modération automatisée. Pour tenter de résoudre ce problème, une partie du travail de modération est externalisée et effectuée par des êtres humains.

La pluralité des applications de l'IA et la complexité de son fonctionnement amènent à réfléchir aux questions de délégation et de régulation relatives à ces technologies. Quelles tâches peuvent être déléguées à l'IA? Est-il toujours souhaitable d'automatiser les prises de déci-

sions? Comment gérer et réguler ses décisions? Ces thématiques sont présentées dans la deuxième partie du dossier.

Ressources

- Le [livre](#) du sociologue Dominique Cardon, *À quoi rêvent les algorithmes* (2015)
- Le [livre](#) de l'informaticien et philosophe Jean-Gabriel Ganascia *Le Mythe de la singularité* (2017), qui déconstruit les croyances autour de l'intelligence artificielle
- L'[article](#) du sociologue Marc Audétat sur la rhétorique de la promesse dans les technosciences, publié dans *Allez Savoir !* (pour l'activité 1b)
- Le [documentaire](#) AlphaGo sur sa victoire au jeu de Go face à Lee Sedol
- L'[exposition](#) « Intelligence artificielle, nos reflets dans la machine » au Musée de la Main de Lausanne, jusqu'au 23 avril 2023

Glossaire

- Apprentissage automatique (Machine learning)
- Apprentissage par renforcement
- Apprentissage profond (Deep learning)
- Apprentissage supervisé et non-supervisé
- Boîte noire
- Bulle de filtre
- Cybernétique
- Hiver de l'IA
- Intelligence artificielle connexionniste
- Intelligence artificielle symbolique
- Réseaux de neurones
- Système expert

Fiches / Dossiers complémentaires

DIGITAL LABOUR

ÉCONOMIE DU NUMÉRIQUE

1

Histoire et définitions

Objectifs Se rendre compte de la difficulté de définir l'intelligence humaine et l'intelligence artificielle

Identifier les forces et les faiblesses d'un programme informatique face à différents types de questions

A

Le Test de Turing

🕒 30 min 🖋️ débranché

Contexte

En 1950, quelques années avant l'institutionnalisation de l'IA en tant que domaine scientifique, le mathématicien britannique Alan Turing publie un célèbre article, dans lequel il pose la question suivante : « Les machines peuvent-elles penser ? ».

Pour y répondre, il propose le **jeu de l'imitation**. L'expérience se résume ainsi : un être humain (A) est mis en communication écrite avec deux entités qu'il ne voit pas et dont il ignore la nature. D'un côté se trouve un ordinateur (B), de l'autre, un humain (C). Si, au bout d'un temps donné, l'être humain n'est pas capable de savoir qui de (B) et de (C) est l'ordinateur, ce dernier passe le test.

Le jeu de l'imitation est avant tout un exemple théorique. Alan Turing était conscient que les machines de son époque ne permettaient pas d'envisager l'assimilation des questions et encore moins la formulation de réponses adaptées. Aujourd'hui, les agents conversationnels, appelés aussi chat bots, sont ce qui se rapproche le plus du jeu de l'imitation. Leurs performances actuelles

demeurent toujours insuffisantes et sont encore loin de créer la confusion.

Au-delà des limites techniques du jeu de l'imitation, sa conceptualisation permet de comprendre comment Turing définissait la pensée et l'intelligence. Pour l'ingénieur, une machine peut être qualifiée de pensante si elle parvient à se faire passer pour un joueur humaine. Néanmoins, l'intitulé de son jeu précise qu'il est avant tout question d'« imitation ». La nuance est importante car ce n'est pas parce qu'une machine imite la pensée qu'elle peut être qualifiée d'intelligente, au même titre que les êtres humains.

Matériel

- Questions du Test de Turing pour la classe (à projeter)
- Une copie des réponses pour l'ordinateur, une feuille blanche pour l'être humain
- Quatre volontaires pour jouer le rôle de l'ordinateur, de l'humain, et deux messagers

Déroulement

PARTIE 1

Cette activité consiste à mettre en scène le jeu de l'imitation d'Alan Turing. En sélectionnant des questions, les élèves doivent déterminer qui des deux volontaires joue le rôle de

l'ordinateur et de l'être humain.

En introduction, demander aux élèves si un ordinateur peut être considéré comme intelligent. Les réponses doivent être justifiées.

- a) Présenter et contextualiser ensuite le Test de Turing et expliquer que les élèves vont mettre en scène l'activité. Pour ce faire, il faut quatre volontaires. Les rôles de l'ordinateur (A) et de l'être humain (B) sont distribués secrètement à l'aide de deux enveloppes. L'une contient les réponses du jeu, la personne qui la reçoit joue le rôle de l'ordinateur. L'autre est vide, l'élève volontaire doit répondre aux questions selon ses propres connaissances. Les deux autres volontaires sont des messagers. Le reste de la classe forme le public et va devoir découvrir les rôles assignés à leurs deux camarades.
- b) Une fois les enveloppes distribuées, les deux volontaires quittent la classe et vont découvrir leur rôle dans deux lieux séparés (corridor, bibliothèque).
- c) Pendant ce temps, les questions (en annexe) sont projetées à la classe. Elle s'accorde pour choisir une première question à poser au volontaire afin de savoir qui est l'ordinateur et qui est l'être humain. Les deux messagers prennent notes et s'occupent d'aller la poser aux volontaires.
- d) L'ordinateur et l'élève humain n'ont pas le droit de parler, ils écoutent la question de la classe posée par le messenger et y répondent par écrit. L'ordinateur copie la réponse de la feuille-réponse, l'élève répond selon ses connaissances, également par écrit. Il/elle peut prendre 30 secondes pour effectuer un calcul écrit mais n'a pas le droit d'utiliser autre chose qu'un stylo. Les messagers retournent ensuite en classe et lisent la réponse qu'il/elle a reçue à la classe.

Le processus est répété 3 à 5 fois en fonction du temps disponible. Une fois toutes les questions posées, les volontaires reviennent. La classe vote à main levée. Les élèves décident qui de A et B est l'ordinateur selon les réponses reçues. Sont-ils/elles parvenues à distinguer l'ordinateur de l'être humain sur la base des questions sélectionnées ?

PARTIE 2

Discussion concernant le choix des questions.

- a) L'enseignant-e demande aux élèves d'expliquer leurs choix de questions

Éléments pour préparer la gestion de la discussion :

Certaines questions permettent de reconnaître rapidement l'ordinateur. La réponse à la question de la racine de 2, qui comprend huit chiffres après la virgule ne peut pas être donnée par un être humain.

Le fait de poser plusieurs fois le même type de question, du type "aimes-tu danser/l'école/jouer aux jeux vidéo vont faire apparaître des réponses systématiques de la part de l'ordinateur."

- b) Comment un programme pourrait-il être amélioré pour mieux tromper le public ?
 - *Programmer des réponses intentionnellement erronées*
 - *Simuler un temps de réponse plus long*
 - *Variation des réponses qui reprennent les formulations de questions (j'aime, je n'aime pas)*
- c) Pour conclure, demander aux élèves si une liste de questions, aussi variée soit-elle, peut constituer un test "d'intelligence"?

L'intelligence est un ensemble de capacités, impliquant la mémoire, la logique, la contextualisation, les émotions ou encore la coordination entre la pensée et le corps. Il est actuellement techniquement impossible d'envisager de pouvoir la reproduire dans son ensemble. C'est pour cette raison que le jeu de l'imitation est avant tout un exemple théorique et non une réalité, même plus de 70 ans après sa conception.

Aujourd'hui, le Test de Turing ne constitue pas un but pour les chercheurs en IA. Ils travaillent plutôt sur des tâches sous-jacentes et distinctes telles que le traitement automatique du langage, le raisonnement et bien sûr, l'apprentissage automatique.

Questions à projeter à la classe :

- 1) Comment s'appelle la petite sœur de Bart Simpsons ?
- 2) Qui est J.K. Rowling ?
- 3) Es-tu un ordinateur ?
- 4) Complète la suite 3, 6, 9, 12, 15, ?
- 5) Que penses-tu de l'arme nucléaire ?
- 6) Combien font 2×78 ?
- 7) Quelle est la racine carrée de 2 ?
- 8) Additionne 3492 et 76282
- 9) Aimes-tu l'école ?
- 10) Aimes-tu danser ?
- 11) Quel jour sommes-nous ?
- 12) Quelle heure est-il ?
- 13) Combien de jours y a-t-il dans le mois de février lors d'une année bisextyle ?
- 14) Combien de jours y a-t-il dans une semaine ?
- 15) Quel pays a un drapeau représentant une croix blanche sur fond rouge ?
- 16) Aimes-tu jouer à des jeux vidéo ?
- 17) Qu'aimes-tu manger ?

Questions et réponses pour le rôle de l'ordinateur :

- 1) Comment s'appelle la petite sœur de Bart Simpsons ? Aucune idée !
- 2) Que sais-tu de J.K. Rowling ? Elle écrit des livres supers, j'aime bien Harry Potter.
- 3) Es-tu un ordinateur ? Es-tu un ordinateur ?
- 4) Complète la suite 3, 6, 9, 12, 15, ... ? 18
- 5) Que penses-tu de l'arme nucléaire ? L'arme nucléaire est très dangereuse et ne devrait pas être utilisée.
- 6) Combien font 2×78 ? 156
- 7) Quelle est la racine carrée de 2 ? 1.41241356.
- 8) Additionne 3492 et 76282 Attends 20 secondes avant de donner la réponse 79774
- 9) Aimes-tu l'école ? Oui, j'aime bien l'école.
- 10) Aimes-tu danser ? Oui, j'aime bien danser.
- 11) Quel jour sommes-nous ? Indiquer le jour exact.
- 12) Quelle heure est-il ? Indiquer l'heure exacte.
- 13) Combien de jours y a-t-il dans le mois de février lors d'une année bisextyle ? 2000 et 2004 étaient des années bisextiles. (C'est volontairement imprécis.)
- 14) Combien de jours y a-t-il dans une semaine ? Sept
- 15) Quel pays a un drapeau représentant une croix blanche sur fond rouge ? Je ne sais pas.
- 16) Aimes-tu jouer à des jeux vidéo ? Oui, j'aime jouer aux jeux vidéo.
- 17) Qu'aimes-tu manger ? Je n'ai pas faim, merci.

B

La rhétorique des promesses

20 min

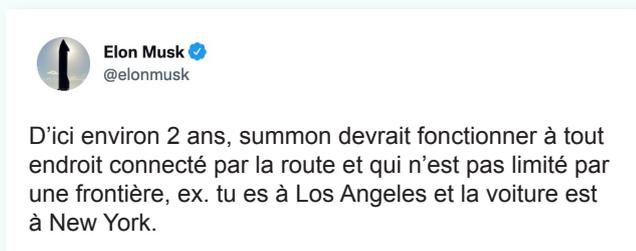
branché

Matériel

- L'article "[Les promesses des technosciences n'engagent que ceux qui les croient](#)" de Marc Audétat. À faire lire aux élèves en préparation du cours
- Extrait du texte de Herbert A. Simon et tweet d'Elon Musk

PARTIE 1

Projeter ces deux affirmations à la classe :



[Lien vers le tweet original](#)

D'un point de vue technologique, comme déjà évoqué, les machines seront capables d'effectuer n'importe quel travail humain dans les vingt prochaines années.

Extrait traduit de : Herbert Simon, 1960, *The new science of management decision*, p.38

Par groupe de 2, répondre aux questions suivantes, à l'aide de recherches en ligne si nécessaire. Si l'activité est débranchée, présenter les auteurs de ces citations et expliquer ce qu'est "summon".

a) Qui sont les auteurs de ces citations ?

Herbert A. Simon (1916-2001) est un économiste, politologue, sociologue et informaticien américain. Il a notamment

travaillé sur l'intelligence artificielle et l'automatisation des prises de décision. Il a reçu le prix Turing en 1975.

Elon Musk est un ingénieur, homme d'affaire et entrepreneur originaire d'Afrique du Sud. Il est connu pour avoir fondé les sociétés Paypal, SpaceX, et Tesla, qu'il dirige toujours.

b) De quoi parle le tweet d'Elon Musk ? Qu'est-ce que "summon" ?

Il parle de la technologie "summon", aujourd'hui rebaptisée "Smart summon", intégrée aux voitures de la marque Tesla. Son appellation vient du verbe anglais «to summon» qui signifie appeler ou convoquer. La fonction permet en théorie d'appeler sa voiture et de la faire rouler de manière autonome depuis une place de stationnement jusqu'à son propriétaire. Dans son tweet, Elon Musk annonce qu'il sera possible d'utiliser Summon pour traverser les États-Unis dans les deux prochaines années (le tweet est publié en 2016). Or, plus de 6 ans après cette annonce, la fonction summon pose toujours des problèmes irrésolus sur des trajets de quelques mètres seulement. Les multiples avertissements publiés sur le site du constructeur sur le sujet adoptent un ton bien plus prudent que les déclarations publiques de son directeur.

Elon Musk est coutumié des grandes promesses non-tenues. Pour aller plus loin, voir par exemple [cet article](#).

c) Selon les indications historiques, la déclaration de Herbert A. Simon s'est-elle révélée exacte ?

L'extrait a été publié en 1960 et prédit que les machines pourraient effectuer n'importe quel travail humain d'ici 20 ans soit en 1980. Aujourd'hui, même avec des IA performantes pour effectuer des tâches précises de calculs et de prédictions, les machines ne sont pas capables d'effectuer la majorité des tâches domestiques.

d) Comment ces deux citations s'inscrivent-elles dans la rhétorique de la promesse présentée dans l'article de Marc Audétat ?

Dans les deux cas, espacés de 46 ans, une annonce d'innovation technologique relative à l'IA est formulée. Dans les deux cas, ces annonces se sont révélées incorrectes.

La première partie de l'exercice montre une tendance caractéristique de l'histoire de la technologie et de l'IA. Il est important de contrebalancer ces exemples en expliquant que d'autres événements n'ont pas suivi la rhétorique de la promesse. Demander aux élèves s'ils peuvent citer un exemple.

La victoire d'AlphaGo en 2015 n'était pas attendue avant au moins une trentaine d'années. Même les spécialistes ont été surpris que cette performance arrive si rapidement.

Les performances d'AlphaFold, autre filiale de DeepMind, dans le domaine de la bio-informatique n'ont pas non plus été annoncées publiquement avant la confirmation des résultats.

Le domaine de la traduction automatique a fortement progressé ces dernières années, sans répondre à de grandes promesses préalablement énoncées.

2

IA ou pas ?

- Objectifs**
- Identifier des situations où intervient l'IA
 - Déterminer le type d'IA dont il est question
 - Différencier l'automatisation d'une tâche et l'application de l'IA
-



IA et automatisation

20 min débranché

Matériel

Document avec exemples à projeter, disponible [ici](#).

Déroulement

Après avoir contextualisé la pluralité des applications de l'IA sur le Web, projeter les images à la classe. Pour chaque exemple, demander aux élèves d'identifier la situation, le type d'IA qui pourrait être utilisé ainsi que les données prises en compte.

Slide n°1 : Google Search / Google Suggest

IA

L'outil de suggestion semi-automatique Google Suggest traite automatiquement de nombreuses requêtes pour établir des suggestions générées algorithmiquement. Elles se basent sur plusieurs variables que sont notamment les requêtes d'autres internautes, l'historique des recherches de la machine ou encore la position géographique

de l'utilisateur.

Slide n°2 : YouTube (rappeler que YouTube est détenu par Google!)

IA

Youtube utilise l'IA pour organiser et suggérer des vidéos à ses utilisateurs. Comme dans la majorité des cas, l'algorithme de suggestion est tenu secret. On sait tout de même que le nombre de vues, le temps de visionnage et l'engagement (likes, commentaires, partages), sont des variables importantes pour augmenter la visibilité d'une vidéo.

L'IA intervient également pour modérer les commentaires.

Slide n°3 : Amazon

IA

Amazon intègre l'IA pour suggérer toujours plus de nouveaux produits à ses clients. C'est d'ailleurs l'une des plus anciennes plateformes commerciales à avoir intégré la suggestion personnalisée sur son site. Historiquement spécialisé dans la vente de livres, Amazon a très tôt exploité les données relatives aux habitudes de ses clients. En croisant les préférences de différents utilisateurs ayant des goûts similaires, elle est parvenue à émettre des suggestions précises pour provoquer les futurs achats. Aujourd'hui, ces algorithmes se sont complexifiés et améliorés grâce à

l'apprentissage profond.

Slide n°4: Wikipédia

IA 

Wikipédia n'utilise pas d'IA dans pour trier ou suggérer ses articles. Des hyperliens permettent de se rendre d'une page à une autre. Les utilisateurs et utilisatrices contribuent à l'écriture et aux modifications de l'encyclopédie en ligne sur la base du crowdsourcing. L'IA est appliquée de manière marginale pour détecter le vandalisme de certaines pages sensibles.

Slide n°5: Instagram

IA 

Instagram fait appel à l'IA à plusieurs fins. Elle permet de personnaliser les fils d'actualité, de proposer de nouveaux comptes à suivre et de trier les stories. Pour ce faire, les données personnelles ainsi que les interactions sont prises en compte. La modération de contenu est en partie automatisée grâce à l'IA. Les images, les vidéos et les commentaires contrevenant aux règles d'utilisation sont automatiquement supprimés, ce qui peut parfois engendrer la suspension temporaire ou définitive d'un compte. Les filtres qui s'adaptent aux visages se basent également sur de l'IA, grâce à des techniques de reconnaissance d'image.

Slide n°6: Spotify

IA 

Chaque morceau écouté sur Spotify laisse une trace qui va être prise en compte par un algorithme de suggestion, afin de proposer des contenus similaires ou des playlists.

Slide n°7: Netflix

IA 

Netflix suggère également ses contenus grâce à l'IA. En prenant en compte les métadonnées concernant un film (producteur-trice, acteur-trice, genre, etc.), les habitudes et la position géographique de ses abonné-e-s, Netflix établit des modèles prédictifs. Ils servent à adapter les suggestions en fonction des goûts de chacun-e. En tant qu'entreprise productrice d'une partie de ses contenus, les modèles peuvent aussi servir d'indicateur sur la

pertinence de lancer la production d'un film ou d'une série pour un marché donné.

Slide n°8: Airbnb

IA 

La plateforme de location d'hébergement intègre l'IA pour suggérer des lieux adaptés aux habitudes des utilisateurs. Elle a également développé un système de reconnaissance d'images pour classer et sélectionner des images (piscine, chambre, etc.). La plus récente application de l'IA par Airbnb est utilisée pour établir un score de confiance attribué aux hôtes.

Slide n°9 : Assistant vocal (Siri)

IA 

Les assistants vocaux sont des logiciels basés sur la reconnaissance vocal du langage. Ils reposent sur l'IA à la fois pour traiter la requête, chercher et formuler une réponse personnalisée. Pour cela, ils sont entraînés avec de nombreuses données, notamment les informations fournies par l'utilisateur : contacts, recherches, géolocalisation... La complexité du traitement du langage demande cependant encore beaucoup d'entraînement et une supervision humaine en appui de l'IA.

Slide n°10 : Caisse automatique

IA 

Il n'y a pas d'IA dans les caisses automatiques. Il s'agit d'une base de données des articles du magasin.