

1. Qu'est-ce qu'une IA ?

DÉFINITION DU PARLEMENT EUROPÉEN

« L'IA désigne la possibilité pour une machine de reproduire des comportements liés aux humains, tels que le raisonnement, la planification et la créativité. »

1. Pourquoi le Parlement Européen a-t-il besoin de définir l'intelligence artificielle ?
2. Êtes-vous d'accord avec cette définition de l'IA ? Avez-vous d'autres éléments à y ajouter ?

DÉFINITION PAR CHATGPT

« Une intelligence artificielle (IA) est un système informatique conçu pour accomplir des tâches qui, normalement, nécessitent l'intelligence humaine. L'IA essaie d'imiter certains aspects de la pensée humaine — comme raisonner, planifier, ou prendre des décisions.

Pour répondre aux questions qui lui sont posées, l'IA s'appuie sur une base de connaissances et génère ses réponses en prédisant, mot après mot, les plus probables en fonction du contexte. »

1. Quelles sont les facultés humaines que l'IA accomplit ?
2. L'IA « *imité* » ces facultés : cela signifie-t-il qu'elle ne les possède pas ?
3. Qu'est-ce qui distingue la manière de répondre à une question d'une IA de celle d'un être humain, et que peut-on en conclure ?
4. Socrate avait fait sienne cette injonction écrite au fronton du temple de Delphes : « *Connais-toi toi-même* ». Que signifie cet impératif ? Et, puisque l'IA ChatGPT est capable de se définir, peut-on dire que ChatGPT se connaît lui-même ?

EXERCICE

1. Définir « *artificiel* » ; à quoi peut-on opposer ce concept ? Utilisez des exemples pour distinguer *artificiel* et son contraire.
2. Définir « *intelligence* » (aidez-vous de l'étymologie latine : mot composé du préfixe *inter* (entre) et de *legere* (cueillir, choisir, lire)) ; prendre des exemples pour appuyer votre définition, et cherchez ce que serait le contraire de l'intelligence.
3. À partir de ces réponses, précisez la définition de l'intelligence artificielle donnée précédemment.

2. La pensée est-elle le propre de l'homme ?

2.1. La pensée dans la nature : l'humain et les autres animaux

RENÉ DESCARTES, *Lettre au Marquis de Newcastle* (1646)

« Si on apprend à une pie (1) à dire bonjour à sa maîtresse, lorsqu'elle la voit arriver, ce ne peut être qu'en faisant que la prolation (2) de cette parole devienne le mouvement de quelqu'une de ses passions (3) ; à savoir, ce sera un mouvement de l'espérance qu'elle a de manger, si l'on a toujours accoutumé de lui donner quelque friandise, lorsqu'elle l'a dit ; et ainsi toutes les choses qu'on fait faire aux chiens, chevaux et aux singes ne sont que des mouvements de leur crainte, de leur espérance, ou de leur joie, en sorte qu'ils les peuvent faire sans pensées.



Or, il est, ce me semble, fort remarquable que la parole étant ainsi définie, ne convient qu'à l'homme seul. Car, bien que Montaigne et Charron aient dit qu'il y a plus de différences d'homme à homme, que d'homme à bête, il ne s'est toutefois trouvé aucune bête si parfaite, qu'elle ait usé de quelque signe, pour faire entendre à d'autres animaux quelque chose qui n'eût point de rapport à ses passions ; et il n'y a point d'homme si imparfait, qu'il n'en use ; en sorte que ceux qui sont sourds et muets inventent des signes particuliers, par lesquels ils expriment leurs pensées. Ce qui me semble un très fort argument pour prouver que ce qui fait que les bêtes ne parlent point comme nous, est qu'elles n'ont aucune pensée, et non point que les organes leur manquent. Et on ne peut dire qu'elles parlent entre elles, mais que nous ne les entendons pas ; car, comme les chiens et quelques autres animaux nous expriment leurs passions, ils nous exprimeraient aussi bien leurs pensées, s'ils en avaient. »

(1) *Pie* : oiseau à robe blanche et noire, capable d'imiter la voie humaine

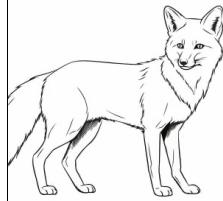
(2) *Prolation* : proférer des paroles sans logique ni sens

(3) *Passion* : comportement involontaire, passif et corporel

1. Pourquoi une pie ou tout autre animal ne pense pas, selon Descartes ?
2. Expliquez la thèse de Montaigne : « *il y a plus de différences d'homme à homme, que d'homme à bête.* » Comment Descartes réfute-t-il cette idée ?
3. Comment Descartes définit-il le langage ? Pourquoi refuse-t-il aux animaux cette capacité ?
4. Que dirait Descartes s'il connaissait l'intelligence artificielle ? Affirmerait-il qu'elle pense ?

MONTAIGNE, *Essais* (1580)

« Existe-t-il une sorte de savoir-faire humain que nous ne retrouvons pas dans les actions des animaux ? Est-il une société réglée avec plus d'ordre, avec une plus grande diversité de charges et d'offices, et maintenue avec plus de constance, que celle des abeilles ? Et pouvons-nous imaginer qu'une telle organisation des fonctions et des actions puisse se faire sans l'usage de la raison et de la sagesse ? (...) »



Voyez par exemple comment font les habitants de Thrace quand ils veulent se risquer sur quelque rivière gelée : ils lâchent un renard devant eux, et quand celui-ci est près du bord, il approche l'oreille de la glace pour savoir si le bruit de l'eau en dessous est proche ou lointain, en déduit que l'épaisseur est plus ou moins grande, et donc avance ou bien recule... Quand on voit cela, ne peut-on penser que lui passent par la tête les mêmes idées que celles que nous aurions nous aussi dans cette situation, et qu'il s'agit là d'un raisonnement et d'une conclusion qui viennent du bon sens naturel, comme : « ce qui fait du bruit est agité ; ce qui est agité n'est pas gelé ; ce qui n'est pas gelé est liquide, et ce qui est liquide ne peut supporter de poids. »

1. Quelle thèse Montaigne défend-il dans le premier paragraphe ?

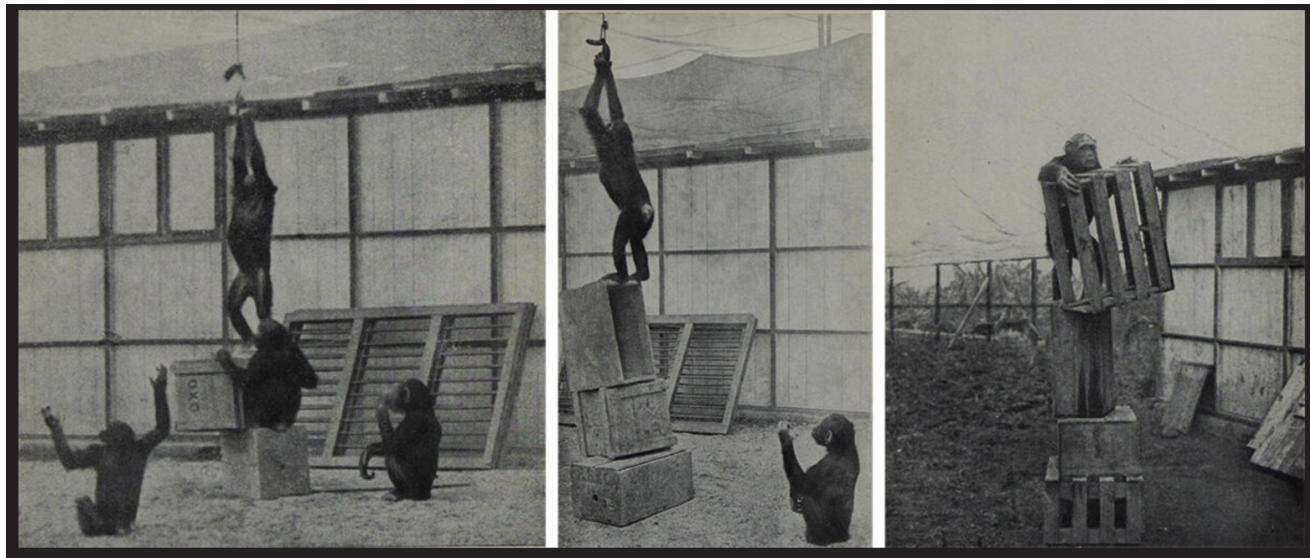
2. En quoi consiste l'intelligence des renards dans l'exemple qu'il donne dans le second paragraphe ?

REVUE « CARNETS DE SCIENCE » (2023)

« Des études commencent à ébranler nos certitudes sur la supériorité humaine. «*On s'est aperçu que l'animal, considéré comme une machine et censé fonctionner par stimulus-réponse, ne fonctionnait pas tout à fait comme ça*», explique Pierre Le Neindre, ingénieur agronome et éthologue (1). *On observe qu'il répond aussi à des situations selon la manière dont il les perçoit, qu'il prend des décisions en fonction de son environnement et de son vécu antérieur.* Un individu conscient de la situation dans laquelle il se trouve, capable de s'adapter aux circonstances et de se concevoir dans le temps : l'éthologie se donne ainsi une définition de l'intelligence animale.

On l'observe d'abord chez le chimpanzé, avec les expériences fondatrices du psychologue allemand Wolfgang Köhler, entre 1913 et 1917. Pour vérifier la capacité du singe à s'adapter à son environnement immédiat, Köhler suspend une banane au plafond d'un local rempli de caisses éparses. Le test vise à découvrir si l'animal va comprendre qu'il lui faut empiler les caisses pour atteindre la banane. S'il fait au début un peu n'importe quoi, soudain le chimpanzé découvre la solution et envisage la forme de l'empilement. Comme s'il avait bel et bien une vision globale de la situation et faisait preuve d'une capacité, proche de celle de l'être humain, à interagir avec son environnement. »

(1) *Éthologie : science du comportement animal*



1. Comment l'éthologie définit-elle l'intelligence ?

2. Décrivez l'expérience de Wolfgang Köhler. Que nous prouve-t-elle ?

3. Les découvertes en éthologie confirment-elles la thèse de Descartes, ou celle de Montaigne ?

4. Selon vous, peut-on comparer l'intelligence de ces singes à celle d'une IA ?

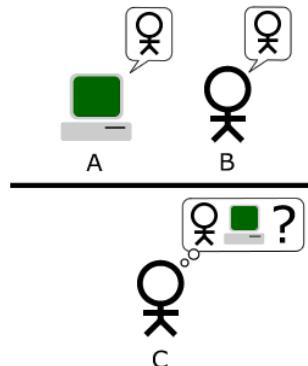
2.2. La pensée dans la machine**EXERCICE – Discuter avec ChatGPT en partant de la question : « Es-tu intelligent, ChatGPT ? »**

Prendre des notes pendant la discussion et écrire une synthèse de quelques lignes.

3. Suffit-il à une machine de parler pour penser ?

LE TEST DE TURING (article de *Philosophie Magazine*)

« À la question épineuse : “Les machines peuvent-elles penser ?”, Turing a proposé une réponse habile et indirecte, en fixant les règles de ce qu'il appelle le “jeu de l'imitation”. En résumé, une personne [C] communique à l'aveugle par messages écrits interposés avec deux partenaires [A et B], dont l'un est un humain [B] et l'autre un ordinateur [C] qui essaie de se faire passer pour un humain. Le test est considéré réussi si le juge n'arrive pas à déterminer qui est l'humain et qui est la machine. En 1950, Turing a prédit : “dans cinquante ans”, il sera possible de programmer un ordinateur “de telle façon qu'un interrogateur moyen n'aura pas plus de 70 % de chances de procéder à la bonne identification après cinq minutes de conversation”. (...)



Lors du *Turing Test 2014*, le programme informatique *Eugene* a réussi à faire croire à dix juges sur trente qu'il était un adolescent ukrainien durant cinq minutes, mais la tâche lui a été simplifiée par le fait qu'il disait avoir 13 ans et ne pas posséder l'anglais comme langue maternelle. Ceci n'élude pourtant pas la question centrale : si un chatbot peut discuter avec vous en se faisant passer pour un humain, cela signifie-t-il que le programme pense ? Le philosophe John R. Searle ne le croit pas, lui qui a opposé au test de Turing le célèbre exemple de la “chambre chinoise” (1980). Supposons qu'un homme soit enfermé dans une pièce et qu'on lui transmette des messages en caractères chinois. Cet homme ne lit pas cette langue, mais il dispose d'un manuel, qui lui permet de trouver, pour chaque question posée, une petite plaquette sur laquelle une réponse appropriée est calligraphiée. Le questionneur, à l'extérieur de la chambre, aura l'illusion de converser avec quelqu'un qui connaît le chinois. Or, il n'en est rien. L'ordinateur ne comprend pas ce qu'il dit. Il exécute des règles syntaxiques, mais sans contenu sémantique – les mots n'ont pas de sens pour lui.

1. Expliquez en quoi consiste le test de Turing.
2. Que prédit Turing ? Cette prédiction a-t-elle été réalisée ?
3. Le fonctionnalisme, ou computationalisme, est une théorie qui conçoit l'esprit comme un système de traitement de l'information et compare la pensée à un calcul (en anglais : “computation”). Expliquez en quoi la thèse de Turing repose sur cette hypothèse fonctionnaliste.
4. Expliquez l'exemple de la chambre chinoise de John Searle et en quoi elle remet en cause le test de Turing.
5. Sur quelle définition de la pensée se repose Searle, et en quoi s'oppose-t-elle à celle de Turing ?

ALAN TURING, *Les ordinateurs et l'intelligence* (1950)

Dans cet article où il imagine son jeu de l'imitation, Alan Turing expose une objection qu'on pourrait lui faire : les machines ne penseront jamais, parce qu'elles ne peuvent pas être conscientes.

« **L'argument de la conscience**. Cet argument est très bien exprimé par le Professeur Jefferson dans son discours d'obtention de la médaille Lister, dont je cite : “*Pas avant qu'une machine n'ait écrit un sonnet ou composé une symphonie à cause d'émotions et pensées ressenties, et pas par le hasard de concordance de symboles, nous ne pourrons être d'accord sur le fait qu'une machine égale un cerveau humain, c'est-à-dire que non seulement cette machine écrit mais qu'en plus, elle sait ce qu'elle a écrit. Aucun mécanisme ne pourrait ressentir de plaisir lorsqu'il réussit (et pas seulement des signaux artificiels, des stratagèmes faciles), ne pourrait être réconforté par des flatteries, ou rendu misérable par ses erreurs, charmé par le sexe, en colère ou déprimé parce qu'il n'arrive pas obtenir ce qu'il veut.*” »

1. Qu'est-ce que la conscience, selon Jefferson ?
2. En quoi cet argument de la conscience remet-il en question la thèse de Turing selon laquelle un ordinateur pourrait un jour penser ?
3. Que pourrait répondre Turing à cette objection ?

Exercice

1. Expliquez les deux citations suivantes :
 - « *La pensée est le dialogue intérieur et silencieux de l'âme avec elle-même.* » (Platon, V^e s. av. J.-C.)
 - « *Penser, c'est dire non ; c'est à elle-même que la pensée dit non, elle se sépare d'elle-même, elle combat contre elle-même. Réfléchir, c'est nier ce que l'on croit.* » (Alain, 20^e s.)
2. A partir de ces citations et de tout ce qui a été vu précédemment, répondez à la question de notre leçon :

« *Une intelligence artificielle pense-t-elle vraiment ?* »