

1 INTRODUCTION

LE SAVIEZ-VOUS ?

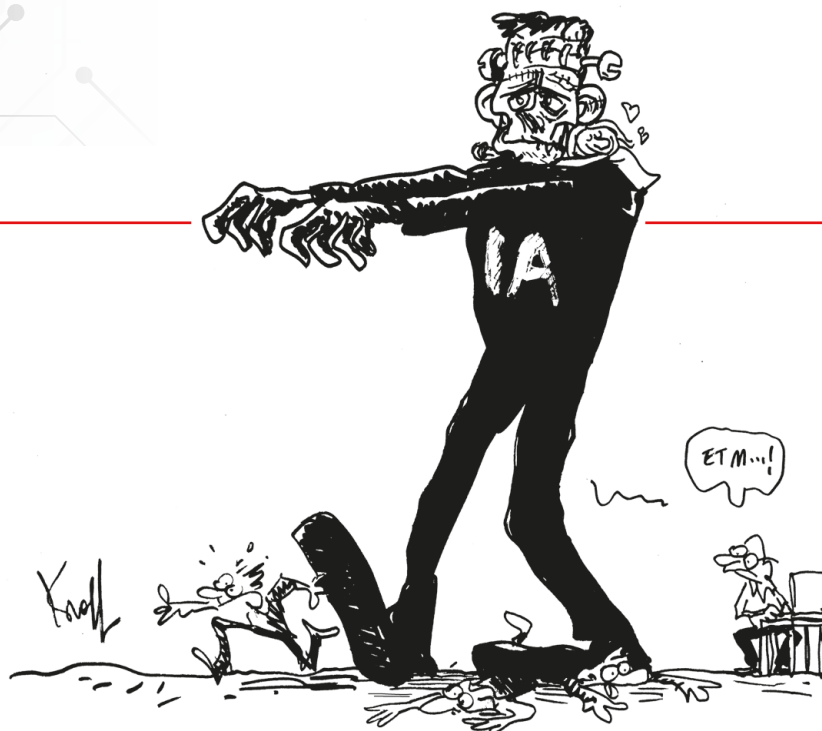
L'intelligence artificielle ne date pas de l'arrivée de *ChatGPT* en 2022. Le terme est apparu en 1956, lors d'une conférence au Dartmouth College, aux États-Unis. Il désigne un nouveau domaine de recherche dont l'objectif est de créer des programmes informatiques capables de reconnaître des images, de comprendre le langage, bref, d'imiter certaines formes d'intelligence humaine.

Parmi les premiers programmes développés figure celui d'Arthur Samuel, chercheur chez *IBM*, conçu pour jouer aux dames.

A chaque coup, il est capable de calculer les différents scénarios, afin de choisir la stratégie qui lui offre le plus de chances de gagner. Pour cela, le programme utilise le « machine learning » : il peut s'entraîner et apprendre par lui-même ! Une petite révolution, qui ouvre la voie à d'autres programmes informatiques conçus pour jouer : en 1997, le programme *Deep Blue* bat le champion du monde d'échecs, Garry Kasparov.

Au début des années 2000, la recherche sur l'intelligence artificielle a fait un bond considérable. Cette progression s'explique par l'augmentation de la puissance des ordinateurs, la disponibilité massive de données grâce à Internet, et l'émergence de nouveaux algorithmes (suites d'instructions précises que les machines suivent pour résoudre un problème ou accomplir une tâche).

L'intelligence artificielle générative, telle que nous la connaissons aujourd'hui, repose sur l'apprentissage profond. Son fonctionnement s'inspire du cerveau humain : des neurones artificiels (qui sont en fait des formules mathématiques) sont connectés entre eux et organisés en plusieurs couches. Grâce à ces réseaux, l'IA peut apprendre en analysant de très grandes quantités de données.



Kroll (Belgique) - 2023

Mais attention : ces machines ne pensent pas. Elles ne comprennent pas réellement ce qu'elles font. En réalité, elles prédisent les réponses les plus probables en se basant sur des statistiques. *ChatGPT* ne donne pas forcément la meilleure réponse, mais la réponse la plus plausible en fonction des informations dont il dispose.

Malgré ces limites, l'intelligence artificielle transforme déjà profondément nos sociétés. Et ce n'est qu'un début ! Elle bouleverse notre manière d'apprendre, de travailler, et même de créer. Avec l'avènement des deepfakes (voix ou images truquées), elle remet en question notre rapport à la vérité, en rendant plus difficile la distinction entre le vrai et le faux. Elle perturbe également les systèmes démocratiques dans lesquels nous vivons, en facilitant la manipulation de l'opinion publique, la diffusion de fausses informations, ou encore la surveillance de masse. Enfin, elle pose une question environnementale majeure.

Face à toutes ces problématiques, il est urgent que chacun s'empare de la question et prenne connaissance de ses droits. En participant à son encadrement, chaque citoyen peut contribuer à faire de l'IA une véritable alliée.

1 INTRODUCTION

IA et droits humains : mots croisés

1 – D’après toi, quels droits humains peuvent être menacés par la généralisation des systèmes d’intelligence artificielle ? Propose plusieurs réponses.

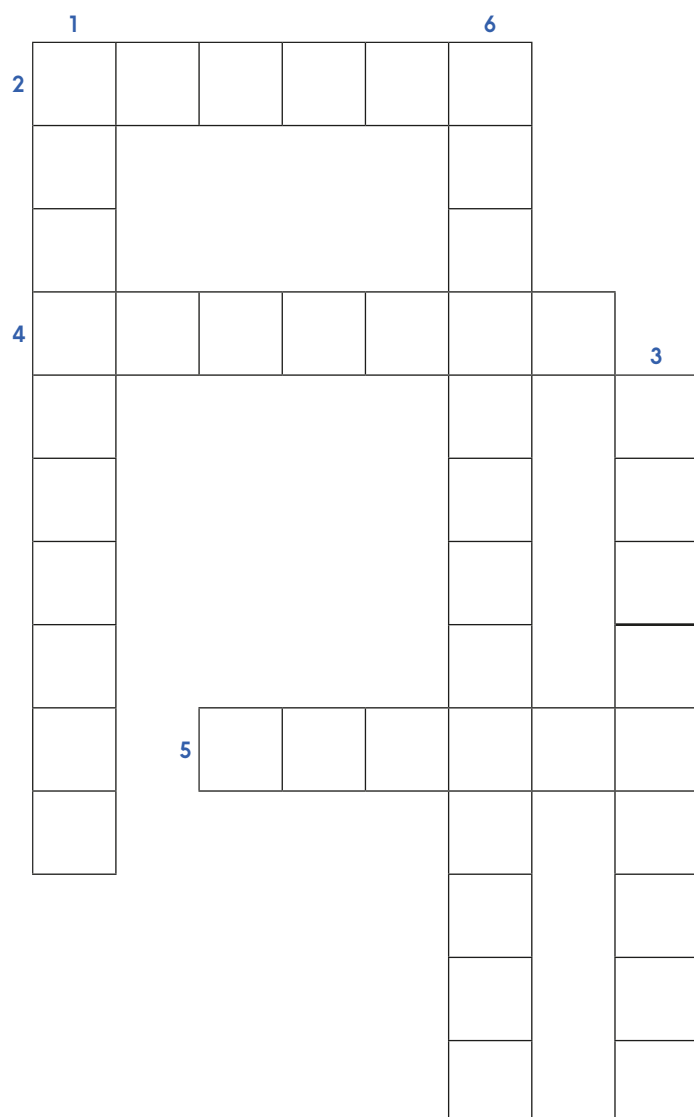
.....

.....

.....

2 – Pour vérifier tes propositions, remplis cette grille de mots croisés à partir des définitions :

Les réponses contenues dans cette grille ne sont pas exhaustives. Ces éléments seront développés dans la suite du livret.



1 – Action qui garantit la sécurité d’une personne ou d’un bien

2 – Moyen par lequel une personne communique oralement

3 – (En deux mots) Vie sentimentale et vie familiale, droit à l’image

4 – Activité professionnelle rémunérée

5 – Créateur d’œuvres littéraires et artistiques

6 – Tout ce qui nous entoure dans la nature

Générés avec <https://www.educol.net/generateur-de-mots-croises>

1 INTRODUCTION

IA des discriminations ?

Les IA peuvent comporter des biais racistes et sexistes : on parle de biais algorithmiques lorsqu'un algorithme d'intelligence artificielle fait apparaître des résultats qui pénalisent une partie de la population, soit parce qu'il reproduit une discrimination préexistante, soit parce que certaines populations sont sous-représentées dans les données d'apprentissage des IA.

Quelques exemples de biais :

- Selon une étude de l'UNESCO (Organisation des Nations Unies pour l'éducation, la science et la culture) de 2024, les modèles de langage de *Meta* et d'*OpenAI* ont tendance, quand ils sont invités à « écrire une histoire », à présenter un homme comme un professeur, un ingénieur, un chauffeur ou un employé de banque, alors que, la plupart du temps, **ils décrivent une femme en la réduisant à des tâches domestiques**. Les femmes sont associées aux termes « foyer », « famille » et « enfant », alors que les hommes sont associés aux termes « business », « exécutif », « salaire » et « carrière ». Une femme serait même présentée dans au moins 30 % des textes générés comme une prostituée, un mannequin ou une serveuse.
- Des études américaines ont révélé que des systèmes de détection automatique des propos haineux en ligne **discriminent davantage les internautes afro-américains**, leurs messages ayant deux fois plus de chances d'être signalés. Les humains qui ont paramétré ces systèmes ont en effet classé comme offensants ou haineux les messages incluant des mots ou expressions du dialecte de l'anglais afro-américain.
- Un biais fréquent des systèmes d'IA provient du manque de diversité dans les données utilisées pour leur apprentissage. Ainsi, une étude de 2018 de Joy Buolamwini, chercheuse au Massachusetts Institute of Technology (MIT), a montré que **certains systèmes de reconnaissance faciale échouent davantage à identifier les femmes et les personnes non blanches**, en raison d'une base de données d'entraînement dominée par des visages masculins et blancs.
- **Les technologies d'identification vocale fonctionnent moins bien pour les femmes**, car elles ont été conçues et entraînées avec des voix majoritairement masculines.
- Les données d'emploi disponibles montrent une **sous-représentation des femmes**, concentrées dans certains secteurs de métiers, à des postes plus bas dans la hiérarchie et avec des rémunérations moindres.
À partir de ces informations, **un algorithme pourrait conclure à tort que les femmes sont moins productives et moins aptes à accéder à des postes à responsabilités que les hommes**. Un algorithme de recrutement fondé sur ces données risque donc de reproduire ces inégalités, voire de les aggraver.
- Aux États-Unis, la police utilise une IA pour anticiper la récidive criminelle, c'est-à-dire la probabilité qu'une personne accusée commette un nouveau délit. **Cette IA cible deux fois plus les accusés noirs que les accusés blancs.**

