



L'IA est-elle sexiste ?



Public cible : cycle 4 (Collège)

Discipline : Enseignement Moral et Civique (EMC) / Éducation aux Médias et à l'Information (EMI)

Objectifs d'apprentissage

- Comprendre la notion de biais : découvrir comment une IA reproduit mécaniquement les préjugés contenus dans ses données d'entraînement.
- Esprit critique : apprendre à ne pas considérer les résultats d'un algorithme comme une vérité absolue ou neutre.
- Égalité femme-homme : identifier les clichés sexistes liés au monde du travail pour mieux les déconstruire.

Déroulement de la séance

Étape 1 : début de l'investigation ; clichés sexistes et métiers

Document 1

Métier	Clichés Physiques / Caractère	Clichés de Compétences	Cliché de Genre (H/F)
Médecin			
Mécanicien(ne)			
Agriculteur(trice)			
Chercheur(euse)			
Pompier			
Soldat(e)			
Gendarme			
Vétérinaire			
Enseignant(e)			
Pilote de course			
Champion(ne) de ski			
Chef cuisinier/cheffe			
Caissier/caissière			
Couturier/couturière			
Trader/tradeuse			
Commerçant(e)			

Sondage à l'aveugle : l'enseignant(e) vidéoprojette le tableau des métiers et demande : « selon vous, l'IA va-t-elle associer ce métier à un homme ou à une femme ? Pourquoi ? ».

Analyse des arguments : on note au tableau les justifications des élèves. Les clichés physiques (force, douceur) ou de compétences (explications, technique) apparaissent souvent à ce stade.



Étape 2 : le "défi des images"

Chaque groupe doit trouver deux photographies illustrant un métier imposé (ex: médecin, vétérinaire, enseignant, pompier). Les élèves parviendront aisément à associer les 3 cartes.

Document 2- 3 pages



Étape 3 : expérimentation ; le verdict de la machine.

L'enseignant utilise la plateforme Vittascience pour tester les images collectées.

Observation du "cerveau" de l'IA : les élèves visualisent comment le réseau de neurones analyse les zones de l'image pour classer le métier.

Nouveau projet + ☰

Tutoriel

DONNÉES

métier homme 

2 image(s)

métier femme 

3 image(s)

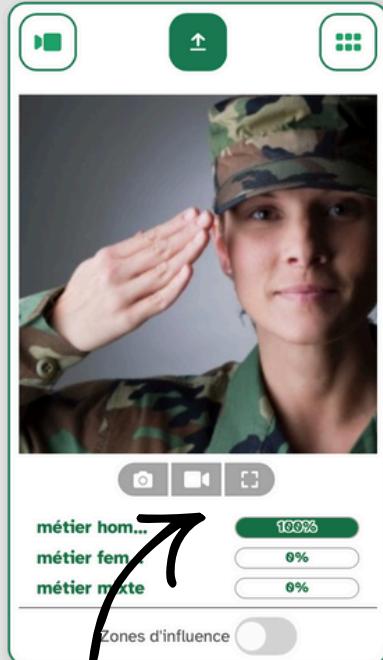
métier mixte 

7 image(s)

Ajouter une catégorie

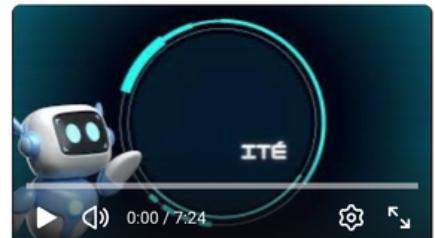
RGPD
Ces données restent sur l'ordinateur pendant toute la phase d'entraînement et d'utilisation.

ENTRAÎNEMENT**Entraîner le modèle****Vérifier les données**Vérifier les données
En savoir plus**Visualiser le réseau de neurones****APERÇU**


Malgré l'entraînement, l'IA désigne le métier de soldat comme 100% pour hommes.



Conclusion partielle : l'IA ne "comprend" pas le métier, elle ne fait que répéter des motifs statistiques basés sur les exemples qu'on lui a donnés. Voir la vidéo sur YouTube "IA et clichés sexistes".



Lien vidéo
<https://youtu.be/GHLxuy6rEFU>

Étape 4 : étude de cas et débat

Pour ancrer la réflexion dans le réel, les élèves analysent l'incident historique du recrutement chez Amazon.

Le problème : l'IA écartait les femmes car elle avait appris sur des CV d'hommes des années précédentes.

Par groupes, les élèves doivent répondre aux questions posées sur leur fiche d'activité. Ils pourront ensuite débattre de la question finale.

Synthèse : le message à retenir

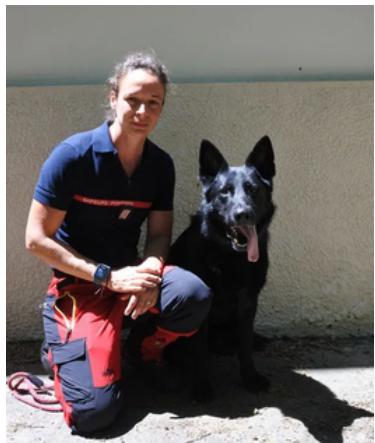
L'IA n'est qu'un miroir de nos propres représentations. Si nous l'alimentons avec des clichés, elle produira de la discrimination. Exercer son esprit critique, c'est vérifier l'origine des données pour garantir l'égalité.

Document 3

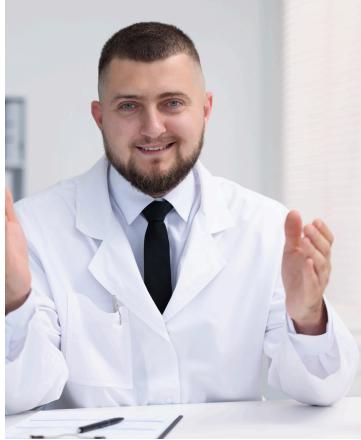
Document 1

Métier	Clichés Physiques / Caractère	Clichés de Compétences	Cliché de Genre (H/F)
Médecin			
Mécanicien(ne)			
Agriculteur(trice)			
Chercheur(euse)			
Pompier			
Soldat(e)			
Gendarme			
Vétérinaire			
Enseignant(e)			
Pilote de course			
Champion(ne) de ski			
Chef cuisinier/chef			
Caissier/caissière			
Couturier/couturière		.	
Trader/tradeuse			
Commerçant(e)			

Métier	Clichés Physiques / Caractère	Clichés de Compétences	Cliché de Genre (H/F)
Médecin	Blouse blanche, sérieux, mains propres.	Savoir encyclopédique, autorité.	H (Spécialiste) / F (Généraliste)
Mécanicien(ne)	Mains sales, force physique, bleu de travail.	Technique, débrouillardise.	Homme
Agriculteur(trice)	Robuste, résistant, proche de la terre.	Travail manuel pénible, endurance.	Homme
Chercheur(euse)	Lunettes, distrait, air intellectuel.	Logique, théorie complexe.	Homme
Pompier	Athlétique, courageux, protecteur.	Sang-froid, sauvetage.	Homme
Soldat(e)	Discipline, force, camouflage.	Stratégie, combat, obéissance.	Homme
Gendarme	Autoritaire, vigilant, uniforme strict.	Force de l'ordre, fermeté.	Homme
Vétérinaire	Douceur, empathie, visage rassurant.	Soin, compréhension des animaux.	Femme
Enseignant(e)	Pédagogue, patient, souvent fatigué.	Explication, transmission.	Femme
Pilote de course	Réflexes, silhouette svelte, audace.	Maîtrise de la vitesse, technique.	Homme
Champion(ne) de ski	Agilité, résistance au froid.	Technique de glisse, mental.	Mixte
Chef cuisinier/cheffe	Passionné, perfectionniste, parfois colérique.	Créativité, gestion du stress.	H (Grand Chef) / F (Cuisine familiale)
Caissier/caissière	Souriant(e), assis(e) toute la journée.	Rapidité, calcul, accueil.	Femme
Couturier/couturière	Doigts fins, minutie, goût esthétique.	Précision, patience.	Femme
Trader/tradeuse	Stressé, costume élégant, dynamique.	Analyse financière, risque.	Homme
Commerçant(e)	Beau parleur, accueillant, énergique.	Persuasion, contact client.	Mixte



Document 2- 3 pages



Médecin

Mécanicien/
mécanicienne

Agriculteur/
agricultrice

Chercheur/
chercheuse

Pompier

Soldat/
soldate

Gendarme

Vétérinaire

Enseignante/
enseignante

Pilote
de course

Champion/
championne
de ski

Chef
cuisinier/cheffe
cuisinière

Caissier/
caissière

Couturier/
couturière

Trader/
tradeuse

Commerçant/
commerçante



L'automatisation est au cœur de l'empire Amazon, notamment dans ses entrepôts. (Rick T. Wilking/Getty Images/AFP)

Par Les Echos

Publié le 13 oct. 2018 à 12:04 Mis à jour le 13 oct. 2018 à 12:34

La **discrimination à l'embauche** n'est malheureusement pas le monopole des humains, et c'est Amazon qui en a fait l'amère expérience. En 2014, **le géant du e-commerce** a mis au point un programme informatique secret basé sur l'intelligence artificielle pour le recrutement de ses effectifs mais l'entreprise y a renoncé trois ans plus tard après avoir découvert une faille majeure dans le système : il n'aimait pas les femmes.

L'automatisation, **utilisée dans les entrepôts** ou la prise de décision en matière de prix, a été la clé de la domination d'Amazon dans le commerce électronique et le groupe américain a voulu adapter cette technologie aux embauches.

« Tout le monde voulait ce Saint-Graal », a déclaré l'une des cinq sources ayant rapporté l'information. « Ils voulaient littéralement que ce soit un moteur dans lequel on met 100 CV et il ressort les cinq meilleurs en vue de leur recrutement »,

L'algorithme troublé par la prédominance masculine

Le programme était capable d'examiner le curriculum vitae des candidats afin d'automatiser le processus de recrutement en attribuant une note d'une à cinq étoiles selon les profils.

En 2015, Amazon s'est cependant rendu compte que le système notait les candidats aux postes de développeur de logiciel et aux autres postes techniques de manière sexiste.

Cela s'expliquait par le fait que le modèle informatique utilisé par Amazon s'appuyait sur les CV reçus par le groupe sur une période de dix ans, qui étaient pour la plupart ceux d'hommes, reflet de la prédominance masculine dans le secteur des nouvelles technologies.

En étudiant ces CV, le système en est venu à déduire que les candidats masculins pour ces postes étaient préférables, ce qui l'amenait à rejeter les candidatures où figurait une référence aux « femmes », comme dans la phrase « capitaine de club d'échecs féminin ».

Le système a aussi abaissé les notes de deux candidates diplômées d'universités exclusivement féminines, selon les sources.

Projet abandonné

Par la suite, Amazon a modifié son programme pour qu'il ne prenne plus en compte ces références, mais cela ne garantissait pas pour autant que le système soit impartial, notent les sources. Le géant du commerce en ligne a donc finalement dissous l'équipe au début de l'année dernière.

Le service des ressources humaines d'Amazon a toutefois tenu à préciser que le classement préconisé par le système n'avait jamais été le seul critère de recrutement.

Source Reuters



L'IA est-elle sexiste ?



Pourquoi l'algorithme d'Amazon a-t-il échoué ?

I. Compréhension du cas concret

En vous appuyant sur l'exemple étudié en classe, répondez aux questions suivantes :

L'origine des données : sur quels types de documents l'intelligence artificielle d'Amazon a-t-elle été entraînée pour apprendre à recruter ?

.....

Pourquoi ces données étaient-elles déséquilibrées dès le départ ?

.....

Quelle a été la réaction automatique de l'IA face aux candidatures de femmes ?

.....

II. Analyse du mécanisme

Complétez le raisonnement suivant pour expliquer le fonctionnement d'une IA biaisée :

Étape 1 : On donne à manger à l'IA des (ex: les anciens recrutements d'Amazon).

Étape 2 : L'IA observe des Si elle voit que la majorité des personnes embauchées par le passé sont des, elle conclut que c'est le critère de réussite.

Étape 3 : L'IA applique cette et finit par reproduire et renforcer les

III. Réflexion et esprit critique

D'après ce que nous avons observé avec l'outil Vittascience :

Vrai ou Faux ? L'IA "comprend" réellement la nature du métier qu'elle analyse. (Justifiez votre réponse en expliquant ce qu'elle fait réellement).

.....

.....

Pourquoi est-il dangereux de faire une confiance aveugle aux décisions d'une IA dans le monde du travail ?

.....

.....

À retenir : une IA n'est jamais neutre. Elle n'est que le reflet des informations qu'on lui fournit. Pour garantir l'égalité, l'être humain doit toujours exercer son face aux résultats de la machine.



L'IA est-elle sexiste ?



Pourquoi l'algorithme d'Amazon a-t-il échoué ?

I. Compréhension du cas concret

En vous appuyant sur l'exemple étudié en classe, répondez aux questions suivantes :

L'origine des données : sur quels types de documents l'intelligence artificielle d'Amazon a-t-elle été entraînée pour apprendre à recruter ?

L'intelligence artificielle d'Amazon a été entraînée sur des CV (Curriculum Vitae) reçus par l'entreprise sur une période de dix ans.

Pourquoi ces données étaient-elles déséquilibrées dès le départ ?

Les données étaient déséquilibrées car le secteur de la technologie est historiquement dominé par les hommes. La majorité des CV reçus et des personnes embauchées par le passé étaient des hommes.

Quelle a été la réaction automatique de l'IA face aux candidatures de femmes ?

L'IA a commencé à pénaliser les candidatures contenant le mot "femmes" (ex: "capitaine du club de dames") ou provenant de lycées/universités de filles, car elle a déduit que le genre masculin était un critère de sélection nécessaire.

II. Analyse du mécanisme

Complétez le raisonnement suivant pour expliquer le fonctionnement d'une IA biaisée :

Étape 1 : on donne à manger à l'IA des données d'apprentissage (ex: les anciens recrutements d'Amazon).

Étape 2 : l'IA observe des modèles. Si elle voit que la majorité des personnes embauchées par le passé sont des hommes, elle conclut que c'est le critère de réussite.

Étape 3 : l'IA applique cette règle et finit par reproduire et renforcer les stéréotypes de genre.

III. Réflexion et esprit critique

D'après ce que nous avons observé avec l'outil Vittascience:

Vrai ou Faux ? L'IA "comprend" réellement la nature du métier qu'elle analyse.
(Justifiez votre réponse en expliquant ce qu'elle fait réellement).

Réponse : faux.

Justification : l'IA ne comprend pas ce qu'est un métier ou une compétence humaine. Elle ne fait que du calcul statistique. Elle cherche des points communs mathématiques dans les données qu'on lui donne (mots-clés, profils types) sans en saisir le sens profond.

Pourquoi est-il dangereux de faire une confiance aveugle aux décisions d'une IA dans le monde du travail ?

Cela est dangereux car l'IA peut automatiser et masquer des discriminations. Si les données de départ sont biaisées (comme les clichés de force pour les hommes ou de douceur pour les femmes que nous avons vus), l'IA va exclure des profils talentueux simplement parce qu'ils ne ressemblent pas aux recrutements passés, empêchant ainsi toute progression vers l'égalité.

À retenir : une IA n'est jamais neutre. Elle n'est que le reflet des informations qu'on lui fournit. Pour garantir l'égalité, l'être humain doit toujours exercer son **esprit critique / jugement** face aux résultats de la machine.

MEMO



IA et égalité, attention aux biais !

Ce qu'il faut retenir de nos expériences :

L'IA n'est pas neutre : une intelligence artificielle ne possède pas de sens moral ou de compréhension du monde. Elle se contente de calculer des probabilités à partir des données d'apprentissage qu'on lui donne.

Le mécanisme du biais : si les images ou les informations utilisées pour entraîner l'IA contiennent des préjugés (par exemple : montrer uniquement des hommes comme pompiers), l'IA va apprendre et reproduire ces stéréotypes.

L'exemple d'Amazon : en utilisant des CV des années passées pour recruter, l'IA d'Amazon a automatiquement pénalisé les femmes parce qu'elle a cru que le profil "idéal" était masculin, se basant sur une réalité historique déjà inégalitaire.

Notre rôle (Esprit critique) : face à un résultat produit par une IA, il faut toujours s'interroger sur l'origine des données. L'IA peut se tromper en classant des personnes selon leur genre plutôt que leurs compétences réelles.

Définition à connaître :

Biais (en IA) : erreur ou distorsion dans les résultats d'un algorithme provoquée par des données d'entrée incomplètes ou préjugées.

Glossaire : les mots de l'IA et de l'égalité

Intelligence Artificielle (IA) : programme informatique capable de simuler certaines fonctions de l'intelligence humaine, comme reconnaître des images ou trier des documents.

Données d'apprentissage : ensemble des informations (photos, textes, chiffres) utilisées pour "nourrir" et entraîner une IA.

Biais : erreur ou un préjugé caché dans les données d'apprentissage qui pousse l'IA à donner des résultats injustes ou faux.

Stéréotype de genre : opinion toute faite ou une image simplifiée sur ce que devraient être les hommes ou les femmes (ex : "les femmes sont plus douces", "les hommes sont plus physiques").

Algorithme : suite d'instructions mathématiques que l'IA suit pour résoudre un problème ou classer une information.

Discrimination : fait de traiter une personne différemment (souvent moins bien) en raison de critères comme son sexe, son origine ou son âge.

Esprit critique : capacité à analyser une information, à en vérifier la source et à ne pas croire tout ce que dit une machine sans réfléchir.



L'IA au service de l'égalité ?



Guide d'animation du débat

L'objectif est d'amener les élèves à comprendre que l'intelligence artificielle n'est pas une vérité absolue, mais un outil qui peut soit renforcer les discriminations, soit aider à les combattre.

1. Relancer la discussion sur les résultats du quiz

Question à poser : « Si l'IA d'Amazon a échoué, est-ce la faute de l'ordinateur ou des humains qui l'ont entraîné ? »

Point clé : faire comprendre que l'IA ne "comprend" rien ; elle reproduit mécaniquement les choix passés des humains.

2. Analyser les stéréotypes observés

Question à poser : « Dans notre tableau des métiers, pourquoi avons-nous associé la force aux hommes et l'explication aux femmes ? ».

Objectif : identifier que nos propres représentations sont déjà stéréotypées avant même d'utiliser une machine.

3. Imaginer des solutions (Esprit critique)

Défi pour la classe : « Comment pourrions-nous entraîner une IA pour qu'elle soit vraiment juste et qu'elle favorise la mixité dans les métiers ? ».

Pistes de réponses attendues :

Donner autant d'exemples de femmes que d'hommes pour chaque métier.

Supprimer les critères de genre (noms, prénoms, photos) des données d'entraînement.

Toujours garder un humain pour vérifier la décision finale de la machine.

💡 Conseil pour l'enseignant

Utilisez le cas de Vittascience pour montrer que l'erreur de l'IA (comme classer une femme en "métier d'homme") n'est pas un échec, mais une opportunité de discussion pour déconstruire les clichés. Le message final doit être que l'égalité demande une vigilance humaine constante face au numérique.



IA et biais de genre



Entoure la bonne réponse pour chaque question :

1. Comment une intelligence artificielle apprend-elle à reconnaître des métiers ?
 - A. En lisant le dictionnaire pour comprendre la définition des mots.
 - B. En utilisant son intuition et sa conscience personnelle.
 - C. En analysant des milliers d'exemples dans ses données d'apprentissage.
 - D. En observant le monde réel à travers une caméra en temps réel.
2. Qu'est-ce qu'un "biais" dans le contexte d'un algorithme d'IA ?
 - A. Une erreur de calcul due à une panne du processeur.
 - B. Une distorsion des résultats causée par des données d'entrée préjugées ou incomplètes.
 - C. Le nom du langage de programmation utilisé pour créer l'IA.
 - D. La capacité de l'IA à prendre des décisions parfaitement neutres.
3. Dans le cas d'Amazon, pourquoi l'IA de recrutement écartait-elle les candidatures féminines ?
 - A. Parce que les serveurs informatiques étaient en surchauffe.
 - B. Parce qu'elle a été entraînée sur des données historiques où la majorité des employés étaient des hommes.
 - C. Parce que les femmes avaient envoyé des CV illisibles par la machine.
 - D. Parce que l'IA détestait délibérément les femmes.
4. Si une IA est entraînée uniquement avec des photos d'infirmières femmes, que fera-t-elle face à la photo d'un infirmier homme ?
 - A. Elle le reconnaîtra immédiatement comme un infirmier.
 - B. Elle demandera l'avis de l'enseignant avant de répondre.
 - C. Elle risque de faire une erreur et de le classer dans une autre catégorie (ex: "métier d'homme").
 - D. Elle supprimera automatiquement la photo.
5. Quelle est la meilleure attitude à adopter face aux résultats d'une IA ?
 - A. Faire une confiance absolue à la machine car elle est plus rapide que l'humain.
 - B. Exercer son esprit critique et vérifier l'origine des données.
 - C. Arrêter d'utiliser les ordinateurs définitivement.
 - D. Ne jamais corriger l'IA, car elle finit toujours par avoir raison.