

Atelier 3 - IA et biais

Ce document fait office de compte-rendu pour l'atelier 3 portant la thématique IA et biais, se déroulant lors du Workshop Sciences et Société du 12 décembre 2024.

Durant ce temps, nous avons réfléchi en groupe avec la dizaine de participants de l'atelier, afin d'élaborer plusieurs scénarios d'animation autour de la thématique des biais dans l'IA. Vous trouverez ci-dessous les propositions que nous avons faites.

Médiation 1 : « Prompt'it »

Public et cadre

Grand public et familles

En contexte Événementiel (festivals, stand en médiathèque, ...)

Points d'accroches et contraintes

Le PhotomIAton :

1. Photo de la personne
2. Une première IA génère un prompt, une description de la photo
3. A partir du prompt, l'IA génère une nouvelle image

La personne reçoit sa photo avatar IA.

Objectif principal

Sensibiliser aux biais de l'IA

Message clé

L'IA est une production humaine qui reproduit les stéréotypes d'une tranche de la société.

Impact

Avoir un regard critique sur les risques de l'IA

Connaissances

Ce qu'est une IA et son fonctionnement, quelques exemples de biais humains et reproduits par IA

Déroulé

Après l'accroche du PhotomIAton, on lance une partie en trois temps. N'importe qui peut venir ou partir de la partie grâce à des règles simples.

1. On montre un portrait généré par IA, puis 6 prompts potentiels qui auraient pu servir à générer ce portrait. Les joueurs et joueuses doivent essayer de deviner quel prompt est le bon.

2. A partir d'une autre photo générée par IA, chaque participant-e formule un prompt qui pourrait avoir généré l'image. On compare ensuite avec le véritable prompt, et la personne la plus proche du vrai prompt gagne
3. Le troisième jeu se fait en connecté. Les participant-e-s devront essayer de générer des images sans employer certains mots. La troisième étape permettrait de mettre en évidence certains biais de l'IA.
Exemple : générer l'image d'une chef d'orchestre femme, sans utiliser le mot « femme ».

Médiation 2 : « Qui est-ce ? Vs IA »

Public et cadre

Grand public et familles

En contexte Évènementiel (festivals, stand en médiathèque, ...)

Points d'accroches et contraintes

Débranché

Facile à expliquer

Adapté à du flux

Se joue à deux → Comment y jouer en groupe ?

Objectif principal

Sensibiliser aux biais de l'IA

Message clé

L'IA est une production humaine qui reproduit les stéréotypes d'une tranche de la société.

Impact

Avoir un regard critique sur les risques de l'IA

Connaissances

Types de biais, Impact des biais sur les personnes

Déroulé

2 joueurs ou groupes de joueurs : 1 dans le rôle de l'IA et un dans le rôle de l'humain.

On reprend les règles du Qui est-ce ? en ajoutant quelques variantes.

L'humain va essayer de deviner quelque chose en posant des questions à l'autre joueur.

L'IA répond en tenant compte de sa carte biais qui lui donne quelques critères et infos scientifiques pour analyser les images.

Il pourrait y avoir plusieurs postes de jeu avec un biais différent à chaque fois (genre, ethnie, salaire, ...)

Médiation 3 : « Le labyrinthe des idées »

(Cette médiation a été établie avec l'IA Copilot en lui donnant les contraintes de réflexion que nous nous étions fixés au sein du groupe de participants)

Public et cadre

Grand public et familles

En contexte Évènementiel (festivals, stand en médiathèque, ...)

Objectif principal

Sensibiliser aux biais de l'IA

Message clé

L'IA est une production humaine qui reproduit les stéréotypes d'une tranche de la société.

Impact

Avoir un regard critique sur les risques de l'IA

Déroulé

Les joueurs sont en exploration dans un labyrinthe virtuel conçu par une IA. Chaque salle du labyrinthe représente un défi ou dilemme lié à un biais spécifique. Les joueurs doivent collaborer pour surmonter ces défis en reconnaissant et en corrigeant les biais, afin d'avancer dans le labyrinthe et d'atteindre la sortie.

Déroulement :

1. Intro : Brève description sur l'IA.
2. Rôles : Chaque joueur reçoit un rôle (Data scientist, psychologue, utilisateur final, ...) + des compétences liées à son rôle pour résoudre les problèmes.
3. Exploration du labyrinthe : choix des portes, chemin, défis
4. Défis : Salle des biais cognitifs, salle des biais algorithmiques,
5. Collaboration pour résoudre ces biais
6. Progression en fonction de la résolution.