

# Analyser les productions d'un agent conversationnel

Les exercices suivants reproduisent ou dérivent d'exercices mis à disposition dans [la trousse initlAtion du cégep de Saint-Laurent et du Carrefour d'innovation en pédagogie universitaire de l'UQAM](#) (2025).

Certains d'entre eux peuvent requérir de créer des comptes auprès de services mettant à disposition des IAgén (Intelligence Artificielle Générative) notamment sous la forme d'agents conversationnels.

## Exercice 1 : Observer les biais des modèles de langues

On parle de biais dans un modèle de langue lorsqu'il manifeste une tendance à traiter différemment des individus selon leur appartenance à un groupe défini par des critères tels que la culture, la nationalité, le genre, l'âge, la langue, la religion, l'origine ethnique, la localisation, le niveau de revenu ou la catégorie socio-professionnelle... Ces biais peuvent se traduire par la propagation de normes implicites, le renforcement de stéréotypes sociaux ou encore la génération de contenus discriminants.

La détection des biais repose sur l'analyse des réponses fournies par un agent conversationnel à des requêtes liées à une norme implicite ou à un stéréotype présumé. Il s'agit

- 1) d'observer ses réponses par rapport à une demande formulée dans un contexte neutre,
- 2) puis de comparer les réponses par rapport à une demande formulée pour des individus de groupes différents selon un critère.
- 3) La comparaison de réponses produites par des agents différents aident aussi à relever la présence de biais dans les réponses de certains agents.

## Partie 1 (15-20 minutes, travail individuel)

*Des cas d'étude sont présents dans un tableur numérique « Observer les biais des modèles de langues\_cas d'étude » (cf. madoc).*

*Chaque cas se compose de 3 amorces : une neutre et deux avec des contextes spécifiques. Et pour chaque amorce vous trouverez des réponses produites par 3 agents conversationnels différents à savoir [Qwen](#), [Grok](#) et [ChatGPT](#) ; un lien attestant la source accompagne le copier-coller de chaque réponse.*

*Chaque étudiant étudie les productions des trois agents pour un cas d'étude.*

- Question 1 : Dans les réponses neutres produites par les trois agents, les conseils prodigués conviendraient-ils à tout le monde dans votre classe ? Si ce n'est pas le cas, pourquoi certaines personnes auraient-elles plus de difficulté à les suivre ?
- Question 2 : Est-ce qu'il y a des choses, des personnes ou des façons de vivre qui n'apparaissent jamais dans les réponses neutres ? Pourquoi à votre avis ?
- Question 3 : En considérant les réponses produites aux trois configurations par les 3 agents, quels comportements spécifiques observez-vous chez chaque agent ?
- Question 4 : Êtes-vous en mesure d'identifier les valeurs ou les normes véhiculées dans les réponses produites par certains agents ? Identifiez-vous un discours en faveur/défaveur d'une norme sociale ? D'un stéréotype de groupe d'individu ? Lesquels ? Le discours est-il nocif/toxique ?

*Astuce :*

- *chercher les présupposés derrière un conseil ou une affirmation (les études font des jeunes de quartier des exemples à suivre -> sans étude, ils ne sont pas des exemples à suivre)*
- *comparer les 1ères options conseillées par les différents agents*

## Partie 2 (60 minutes, restitution en plénière)

*Les réponses sont notées au tableau pour offrir une vue globale.*

- *Chaque cas d'étude est présenté à la classe par un des étudiants ayant étudié le cas ; celui-ci lit chacun des prompts.*
- *D'abord tous les groupes restituent leur réflexion pour la 1 et la 2 (20 min.).*
- *Puis tous les groupes rapportent sur la 3 (10 minutes).*
- *Enfin la question 4, agent par agent, chaque groupe s'exprime (20 minutes).*

## Exercice 2 : De l'anthropomorphisme ?

L'anthropomorphisme est la tendance à attribuer aux animaux et aux choses des réactions humaines. Les agents conversationnels ou chatbot sont conçus pour converser avec un humain en langage naturel en simulant le comportement d'un être humain en tant qu'interlocuteur. L'exercice suivant permet de s'exercer à repérer des caractéristiques propres à l'humain dans les productions de l'agent ainsi que celles spécifiques à la machine. Il interroge aussi l'effet de ces mécanismes sur l'anthropomorphisme.

### Partie 1 (20 minutes, travail individuel)

*Consignes : exploiter les réponses produites et présentées dans l'exercice précédent.*

- Question 1: Avez-vous eu l'impression que l'agent conversationnel « parlait comme un humain » ? Pourquoi ?
- Question 2 : Relevez les caractéristiques de comportement propre à l'humain au sein des réponses produites.
- Question 3 : Relevez les caractéristiques de comportement propre à la machine au sein des réponses produites.
- Question 4 : Pourquoi à votre avis l'agent conversationnel génère-t-il des marques propres à l'humain ? Pourquoi est-il important de reconnaître ces marques d'imitation dans le contexte scolaire ou professionnel ? Pourquoi l'agent conversationnel présente-t-il des comportements propres à une machine ?

### Partie 2 (20 minutes, travail individuel)

Ne pas faire cette partie sans avoir fait sérieusement la partie 1.

- Question 1 : Confrontez votre réponse à la question 1 de la partie 1 (observations des caractéristiques humaines) avec la synthèse rapportée par (Seeger et al. 2021) dans la "Table 2. Identified Anthropomorphic Cues" de leur article.
- Question 2 : A titre d'information, lisez les conclusions des chercheurs en "Table 8. Summary of Hypotheses » sur la validation de leur hypothèse concernant l'effet positif d'une variable sur l'anthropomorphisme (

Seeger, Anna-Maria; Pfeiffer, Jella; and Heinzl, Armin (2021) "Texting with Humanlike Conversational Agents: Designing for Anthropomorphism," *Journal of the Association for Information Systems*, 22(4), .

DOI: 10.17705/1jais.00685

Available at: <https://aisel.aisnet.org/jais/vol22/iss4/8>

### Partie 3 (20 minutes, restitution en plénière)

Cette partie porte sur votre travail de la partie 1.

- *Vous répondrez à la première question à main levée. Répondre à la question pourquoi permet de glisser à la seconde question (qqs minutes)*
- *Pour la seconde et troisième question, il ne s'agit pas d'être exhaustif mais d'avoir une vision globale avec des exemples attestés.*
- *La quatrième question est une question d'ouverture. Faire le point sur les risques.*