

**FAVORISER UNE INTELLIGENCE
ARTIFICIELLE (IA) INCLUSIVE ET
RESPONSABLE POUR LA
RÉUSSITE DE TOUTES LES
POPULATIONS ÉTUDIANTES EN
ENSEIGNEMENT SUPÉRIEUR**

**IA ET EDI EN ENSEIGNEMENT
SUPÉRIEUR:
UN MARIAGE HEUREUX?**

Andréane Sabourin Laflamme

Professeure de philosophie, Cégep André-Laurendeau

Doctorante, Faculté de droit, Université de Sherbrooke



L'IA GÉNÉRATIVE : UNE ÉTAPE CHARNIÈRE DANS L'HISTOIRE DE L'IA

- Type d'IA qui, à partir d'une requête textuelle (*Prompt*), génère un nouveau contenu (texte, image, vidéo, audio, code, etc.)
 - Images : Dall-e, Midjourney, Stable Diffusion, LensaAI, etc.
 - Texte: Chat GPT, Bing, Bard (Google), etc.
 - Audio: Resemble AI, Soundful music generator, etc.
- La course à l'IA générative : quand la rapidité l'emporte sur la prudence.
 - Dissensions internes chez Microsoft et Google.
 - Dissolution de l'équipe éthique chez Microsoft avant le lancement du robot conversationnel de Bing.



L'IA EN ENSEIGNEMENT SUPÉRIEUR

Intégration croissante des SIA en éducation

Déferlante de produits et services éducatifs faisant appel à l'IA

L'IA laisse entrevoir des possibilités inédites en éducation:

- Approches pédagogiques personnalisées et adaptées
- Correction et retroaction personnalisée et automatisée
- Production de contenu pédagogique
- Prédiction de réussite



IA ET ENSEIGNEMENT SUPÉRIEUR

En classe

Systemes d'enseignement adaptatif ou personnalisé

Systemes de tutorat intelligent

Les systemes d'evaluation automatique

À l'extérieur de la salle de classe

Prédiction de réussite ou d'échec, de persévérance scolaire ou d'abandon

Évaluation automatisée des professeur.e.s

Robots conversationnels (ChatGPT)

ÉQUITÉ, DIVERSITÉ ET INCLUSION (EDI) EN ENSEIGNEMENT SUPÉRIEUR

L'EDI prend appui sur le droit à l'égalité

- art. 10, Charte des droits et libertés de la personne
- art. 15, Charte canadienne des droits et libertés
- droit international

L'EDI va plus loin que le droit formel et s'intéresse :

- aux les conditions qui peuvent favoriser ou nuire à l'accès égal aux ressources et aux opportunités (équité)
- à la valorisation des différences et des perspectives (diversité)
- à la création d'environnement où tous les individus se sentent valorisés, respectés et inclus (inclusion)

IA ET EDI: UN MARIAGE HEUREUX ?



- Toute percée technologique concernant l'utilisation de l'IA dans l'éducation est une occasion d'améliorer l'accès à l'éducation des groupes les plus vulnérables.
- L'IA peut contribuer à relever certains défis, dont celui de répondre adéquatement aux besoins diversifiés des personnes.
- Les données sur l'apprentissage permettent un suivi personnalisé, ce qui permet d'identifier les difficultés d'une personne et d'y réagir en temps réel.





LE MYTHE DE LA NEUTRALITÉ TECHNOLOGIQUE EN QUESTION

- Les technologies sont des vecteurs de valeurs
- Manque de diversité dans les équipes de développement en IA
 - Les femmes représentent moins d'un cinquième de la main d'œuvre en IA (Didier Bert, La Presse, 2022)
- Préjugés et biais inconscients se retrouvent dans les dispositifs technologiques
 - Ex: le distributeur de savon
 - Ex: la reconnaissance faciale

BIAIS ALGORITHMIQUES

- Biais algorithmiques : reproduit les préjugés et stéréotypes inclus dans les données d'entraînement ou dans les algorithmes.
- Sources possibles des biais
 - Données d'entraînement
 - Conception
 - Utilisation
- Risque de reproduction, d'amplification et d'automatisation des biais sociohistoriques présents dans nos sociétés

Successful CEO, Dall-E, généré le 14 avril 2023





DISCRIMINATION: DÉFINITION

- Discriminer, c'est traiter une personne différemment en raison de ses caractéristiques personnelles ET l'empêcher d'exercer ses droits.
- Traiter une personne différemment, c'est la distinguer, l'exclure ou la préférer en raison de ses caractéristiques personnelles. L'âge, l'origine ou le sexe sont des exemples de caractéristiques personnelles.
- Il y a 14 caractéristiques personnelles qui sont des motifs interdits de discrimination.

Source: CDPDJ



- La discrimination directe concerne un traitement différent, défavorable ou moins favorable, envers un individu ou un groupe en raison d'une ou plusieurs caractéristiques protégées.
- Différence de traitement fondée sur un des motifs interdit.
- Faite de manière volontaire, ouverte ou avouée.

DISCRIMINATION DIRECTE



- Beaucoup plus subtile que la discrimination directe, mais tout de même interdite par la loi, même s'il n'y a pas d'intention de discriminer.
- La notion de discrimination indirecte réfère non pas à un traitement différent, mais à un impact différent.
- Ex: politique ou une pratique qui est « neutre » à première vue, mais qui désavantage disproportionnellement des individus appartenant à un groupe protégé.



DISCRIMINATION INDIRECTE

DISCRIMINATION ALGORITHMIQUE

Définition : Lorsque des SIA biaisés ont un effet discriminatoire sur certains individus ou groupes d'individus.

Problème: automatisation et amplification et de la discrimination.

Discrimination algorithmique par proxy

- Lorsqu'un SIA établit une corrélation entre un motif non interdit et un motif interdit et que cette corrélation est utilisée à des fins discriminatoires ou a des effets discriminatoires.
- Peut être directe ou indirecte.
- Très difficile à identifier et à corriger. Pour certains, elle est inévitable.



LE TRAVAIL HUMAIN DERRIÈRE CHATGPT

- *More than a Glitch*, Meredith Broussard (2023)

BUSINESS • TECHNOLOGY

Exclusive: OpenAI Used Kenyan Workers Less Than \$2 Per Hour to Make ChatGPT Toxic



This image was generated by OpenAI's image-generation software, Dall-E 2. The prompt was: "A seemingly endless view of African workers at desks in front of computer screens in a printmaking style." TIME does not typically use AI-generated art to illustrate its stories, but chose to in this instance in order to draw attention to the power of OpenAI's technology and shed light on the labor that makes it possible. Image generated by Dall-E 2/OpenAI.

LES ENJEUX ÉTHIQUES LIÉS À L'IA

LE PROBLÈME DE LA BOÎTE NOIRE

Opacité (volontaire ?)

Traçabilité


Responsabilité

Redevabilité



RÉSULTATS ERRONÉS ET DÉSINFORMATION

- ChatGPT est un menteur compulsif (*On Bullshit*, Harry Frankfurt)
- L'objectif est de produire un texte plausible. (prédiction de pertinence)
 - Un pas de plus dans l'ère de la post-vérité?
- Quand les LLM "hallucinent" (génèrent de fausses informations d'apparence plausible).
 - Les LLM: des *perroquets stochastiques* ?
- Neutralité et objectivité factices
- Biais d'automatisation

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?  Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21).



DÉCLARATIONS ET CHARTES EN ÉTHIQUE DE L'IA

- Encadrement législatif limité.
- Multiplication des déclarations et chartes
- Issues de différents milieux: entreprises, gouvernements, ONG, organisations internationales, milieu académique, etc.
- Consensus autour de certains principes. (Jobin et al., 2019)



QUELLE RÉGULATION POUR LES LLM ?

Absence de cadre légal et éthique pour l'utilisation de l'IA en enseignement supérieur

Canada

Enquête du Commissaire à la vie privée sur ChatGPT (avril 2023)
Projet de loi C-27 (Loi sur l'IA)

États-Unis

Plainte à la Federal Trade Commission (CAIDP)
Blueprint for an AI Bill of Rights

Europe

EU AI Act

Italie

Interdiction de ChatGPT (RGPD)

L'IMPORTANCE DE LA COMPÉTENCE NUMÉRIQUE ET DE L'ÉTHIQUE DE L'IA EN ENSEIGNEMENT SUPÉRIEUR



Former à l'éthique de l'IA en enseignement supérieur : Référentiel de compétence



FRÉDÉRIK BRUNEAULT,
Cégep André-Laurendeau /
École des médias, UQAM
ANDRÉANE SABOURIN LAFLAMME,
Cégep André-Laurendeau / UQAM
ANDRÉ MONDOUX,
École des médias, UQAM

Février 2022