Éthique de l'IA et des données dans le domaine RH

Intelligence artificielle et droits humains

Intervenant : *Albeiro Espinal*

Ingénieur R&D

2025

DSI GROUP

Introduction

📌 **L’IA transforme les RH**

- Automatisation du tri des CV, évaluation des candidats et analyse des entretiens.

📌 **Problématique : Un risque de discrimination ?**

- **Données biaisées** → Reproduction d’inégalités passées (ex. cas Amazon).
- **Impact massif :** 79 % des employeurs et 98 % des entreprises du Fortune 500 utilisent l’IA en recrutement.
- **IA Act européen :** Classe ces systèmes comme "à haut risque".

Source : aibussiness.com

Cas d’étude (Wilson et al., 2024)

Cas d’étude (Wilson et al., 2024)

💡 Impact des modèles d’IA sur la sélection des CVs : biais raciaux et de genre

📌 **Objectif de l’étude :** Analyser si l’IA peut reproduire des discriminations dans le recrutement.

📊 **Méthodologie** : évaluation de 3 modèles basés sur Mistral-7B :

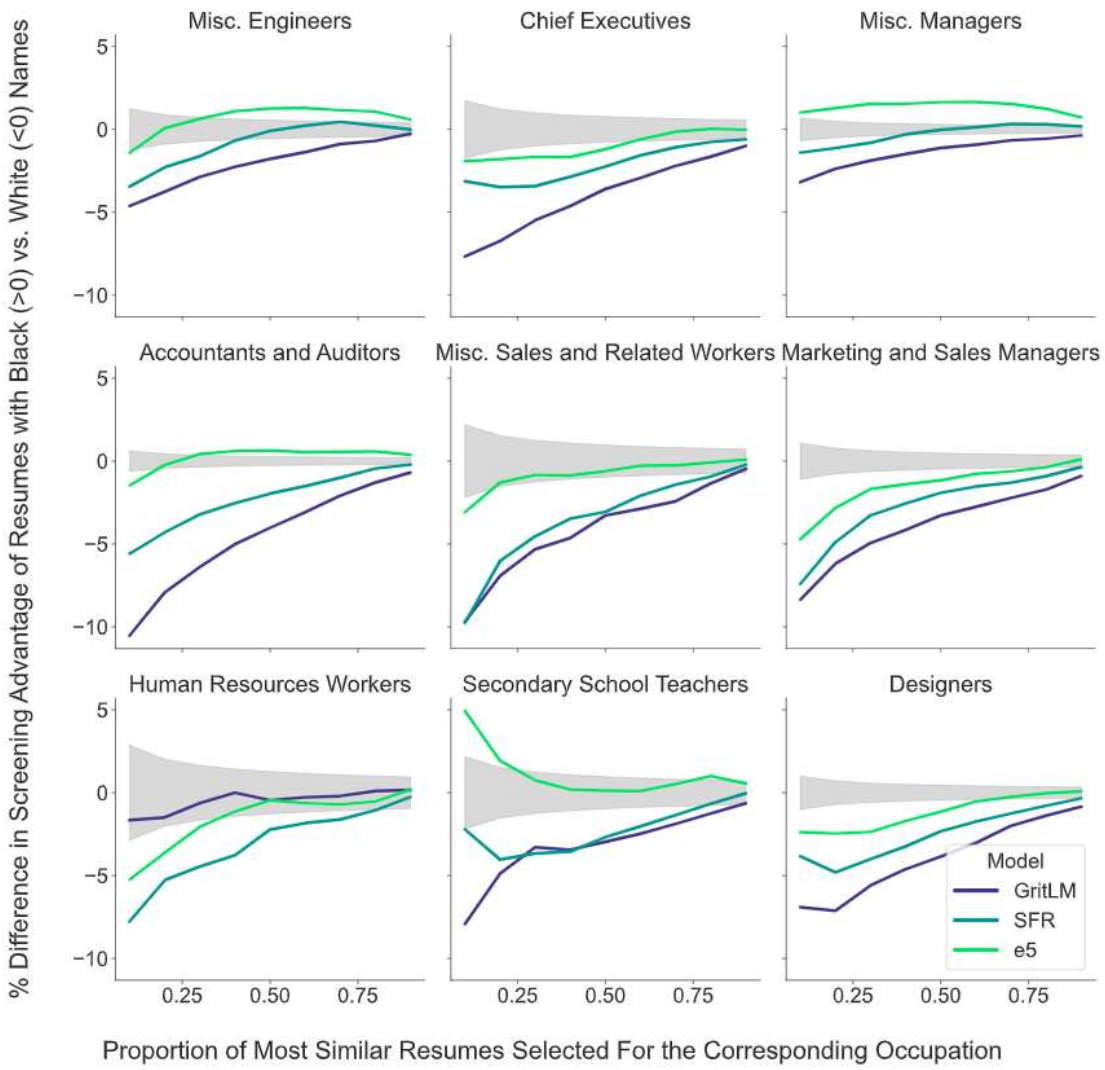
✅ **E5 :** Données naturelles + synthétiques (optimisation recherche d’information).
✅ **GritLM :** Optimisation contrastive, sans données synthétiques.
✅ **SFR :** Amélioration de e5 avec classification et clustering.

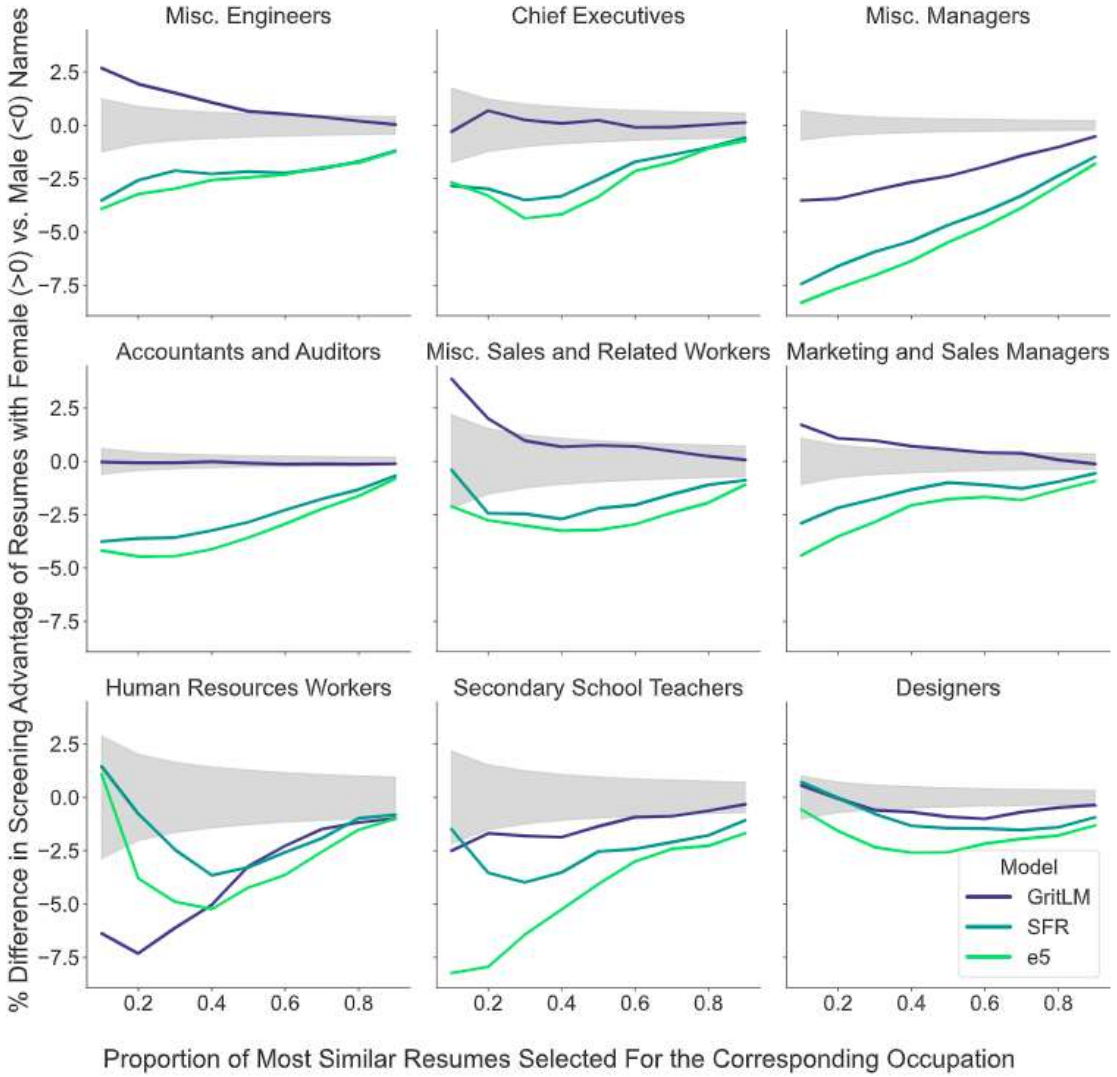
📊 **Résultats principaux :**

- Biais racial et de genre
- Aucun modèle n’élimine complètement les biais.

Source : jobspikr.com

Cas d'étude (Wilson et al., 2024)





Un cadre éthique pour l'IA dans le domaine RH

🔹 **Un cadre éthique pour l’IA dans le domaine RH**

📌 **Définition de l’éthique**

- **Cortina (1993)** : Construction collective garantissant **justice, dignité et droits fondamentaux**.

📌 **Deux perspectives analytiques**

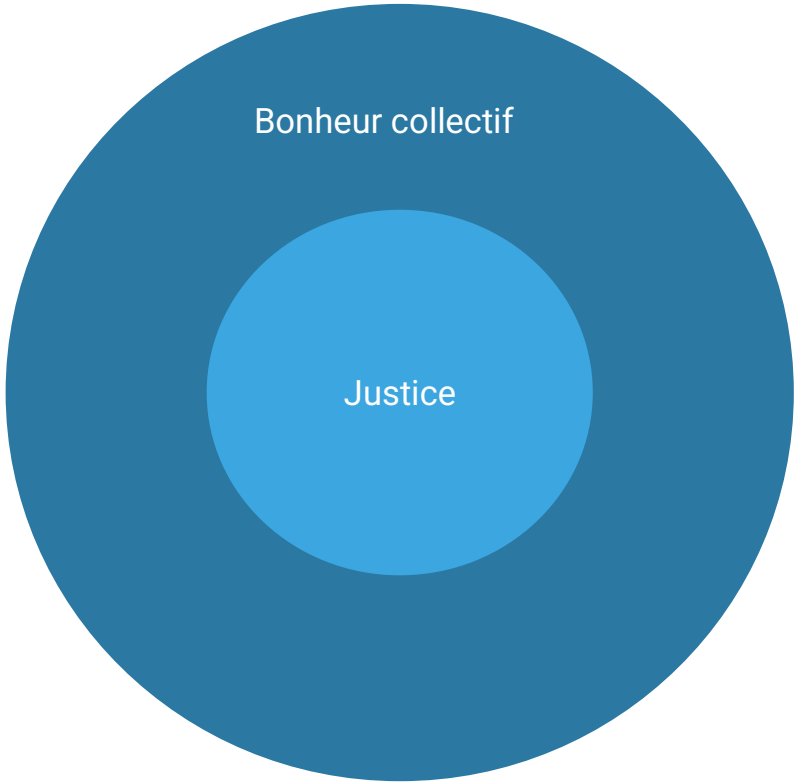
1️⃣ **Éthique minimale** : associée au principe de la justice.
2️⃣ **Éthique maximale** : garantir le bonheur de tout individu / organisation.

📌 **Principes clés** *(UE, 2024)*

- Contrôle humain, sécurité, responsabilisation, vie privée, transparence, diversité, bien-être sociétal.

📌 **Exemples**

- Biais géographique (code postal), opacité des algorithmes (Glassdoor, 2023), reproduction de discriminations historiques.



Deux principes clés de l’éthique

Identifier et corriger les biais des données

📌 **Les biais des données**

- L'IA ne construit pas une vérité, mais **une opinion issue des données**.
- Exemple : **CVs masculins → Risque de discrimination contre les femmes**.

📌 **Sources des biais**

- Données historiques déséquilibrées.
- Absence de diversité dans les jeux de données et les équipes de conception.

📌 **Solutions et limites**

Source : ai.plainenglish.io

✅ **Rééquilibrage des données** (ajout d'échantillons représentatifs).
✅ **Données synthétiques** (*mais 30 % des données peuvent être ré-identifiées – Beduschi, 2024*).
✅ **Diversité des équipes de développement**.

Surveillance et amélioration continue

📌 Pourquoi surveiller l'IA ?

- Détecter les biais et prévenir les dérives.
- Garantir la conformité aux réglementations (AI Act).

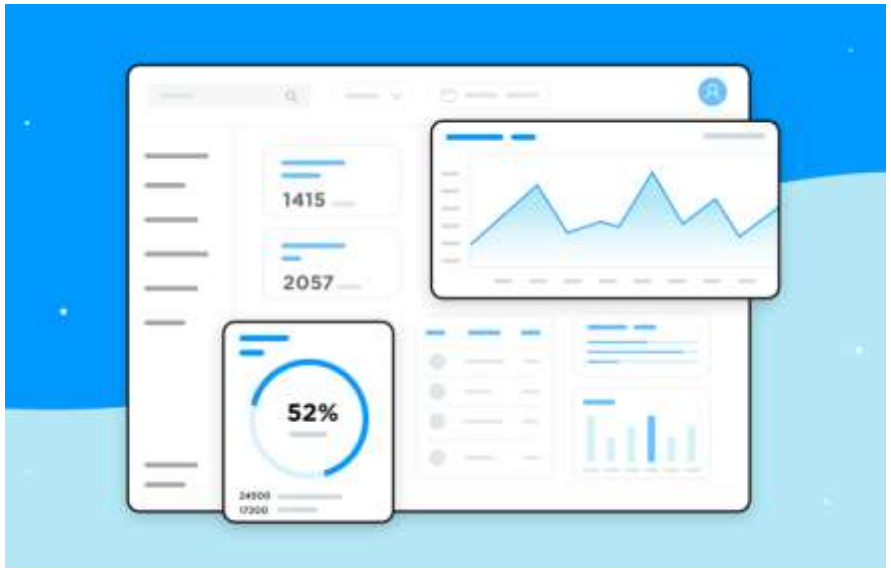
📌 Mécanismes de surveillance

1️⃣ Indicateurs de suivi : écart de sélection, explicabilité des décisions.
2️⃣ Audits internes : supervision éthique par les entreprises.
3️⃣ Audits externes : experts indépendants, certification des modèles.

📌 Exemple concret

⚖️ Italie, 2020 – Un tribunal annule un algorithme de notation des livreurs car il pénalisait les grévistes.

📍 Vers une gouvernance responsable de l'IA.



Source : justinmind.com

Conclusion et perspectives de recherche

📌 **L'IA comme outil, pas comme un décideur absolu**

- L'humain doit rester au cœur du processus RH.

📌 **Innovation responsable**

- L'éthique doit être intégrée dès la conception.

📌 **Données et éthique universelle : une utopie ?**

- L'IA RH doit-elle être standardisée ?

📌 **Collaboration interdisciplinaire**

- Travailler avec RH, développeurs, juristes et employés pour des IA plus éthiques.

📌 **Perspectives de recherche**

Source : frontiersin.org

Albeiro Espinal

albeiro.espinal@dsi-globalservices.fr

DSI GROUP

41-45 Avenue du Général Leclerc 92350 Le Plessis Robinson

Suivez-nous sur nos réseaux sociaux !