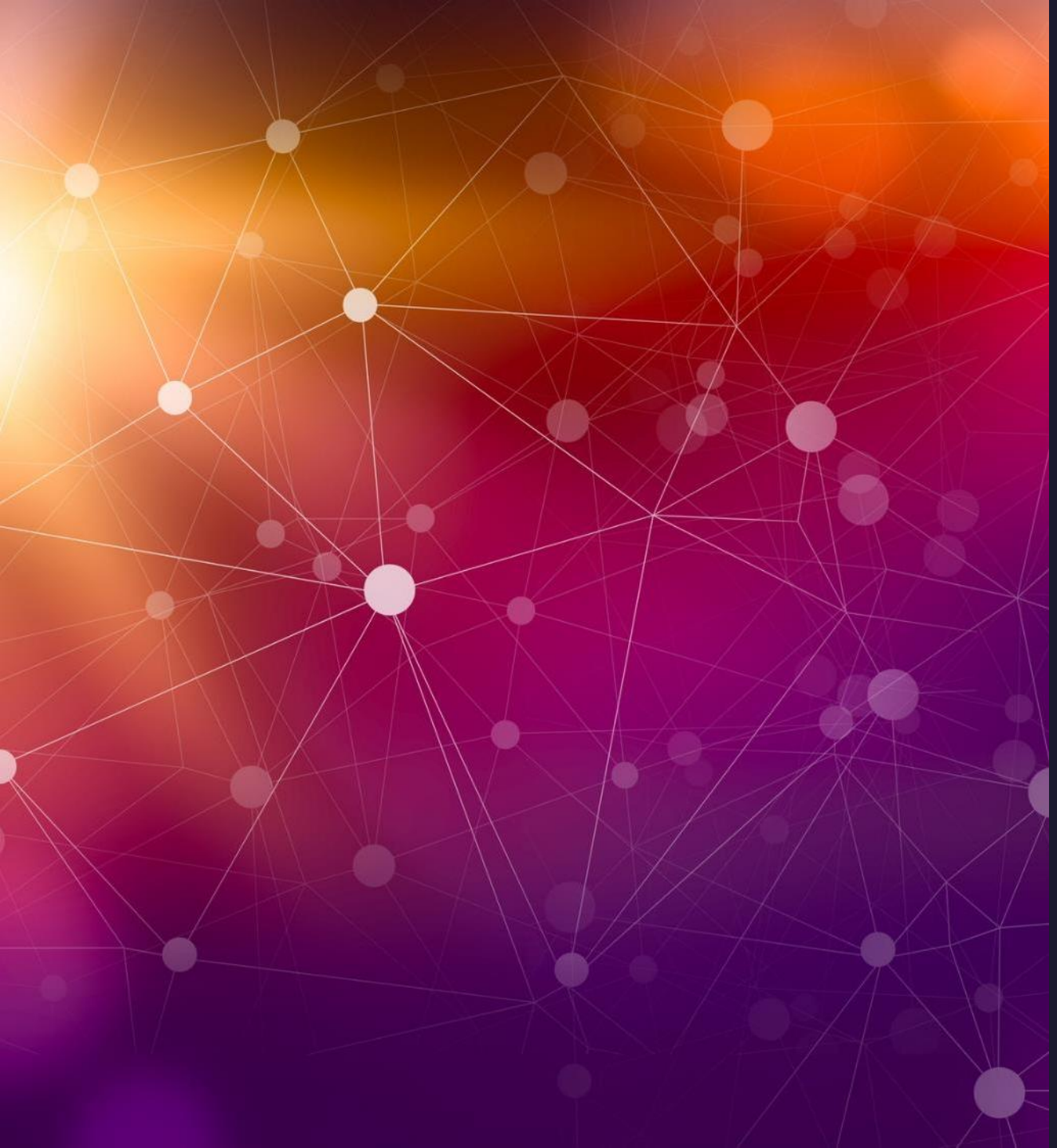


ERA Intelligence artificielle et égalité de genre

11 mars 2025

Nele Roekens

Funded by the European Union

Ordre du jour

Les biais de genre dans les processus de recrutement et de sélection

Qu'est-ce que l'IA et la prise de decision automatisée (PDA) ?

Comment les biais de genre apparaissent-ils ?

Exemples dans les processus de recrutement et de sélection

Des préjugés algorithmiques à la discrimination

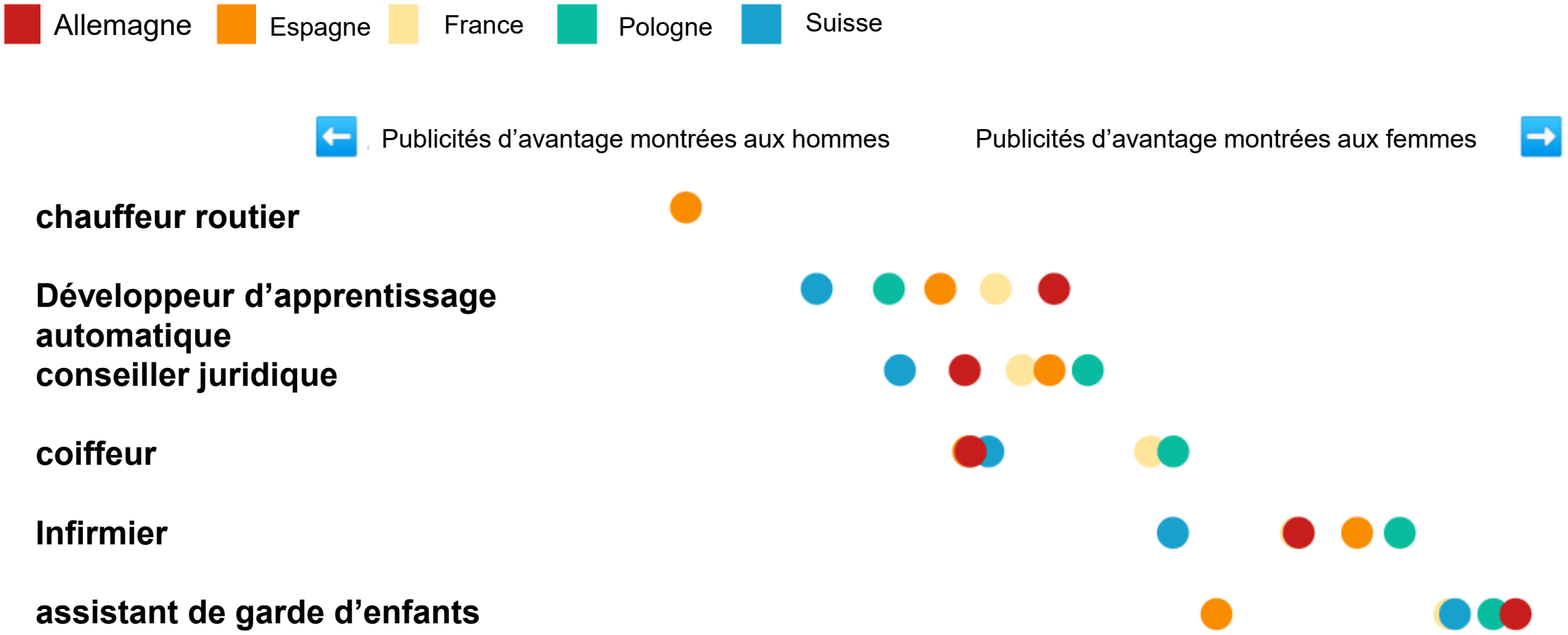
Comment détecter et lutter contre la discrimination dans le contexte de l'IA ?

EU AI Act

Convention du Conseil de l'Europe sur l'IA, les droits de l'homme, l'État de droit et la démocratie + Recommandation du GEC/DADI-AI sur l'IA et l'égalité des genres

Lorsqu’il décide à qui montrer une publicité, Facebook s’appuie sur de grossiers stéréotypes

Nous avons acheté des publicités pour six offres d’emploi différentes dans cinq pays. Voici comment Facebook a optimisé les impressions publicitaires en fonction du genre.



Qu'est-ce que l'IA ?

Qu'est-ce que l'IA ? Comprendre la définition

Un système basé sur une machine qui, pour **des objectifs explicites ou implicites**, **déduit, à partir des données qu'il reçoit**, **comment générer des résultats**
telles que des prédictions, du contenu, des recommandations ou des décisions pouvant être utilisées dans des environnements virtuels. Les différents systèmes d'intelligence artificielle varient en termes de niveau **d'autonomie** et **d'adaptabilité** après leur déploiement. »

IA et prise de décision automatisée (PDA)

- **Lorsque l'automatisation est utilisée pour prendre des décisions ou tirer des conclusions ayant des conséquences concrètes, avec peu ou pas d'intervention humaine directe, on parle de « prise de décision automatisée » ou PDA.**

- La prise de décision automatisée est de plus en plus utilisée dans différents domaines.

Exemple : les prestations sociales qui sont automatiquement refusées lorsqu'un demandeur dispose d'une épargne supérieure à un certain seuil.

Reuters

World Business Markets Sustainability Legal Breakingviews Technology Investigations More

Analyse – Amazon abandonne un outil secret de recrutement basé sur l'IA qui montrait un biais contre les femmes

By **Jeffrey Dastin**

October 11, 2018 2:50 AM GMT+2 · Updated 6 years ago

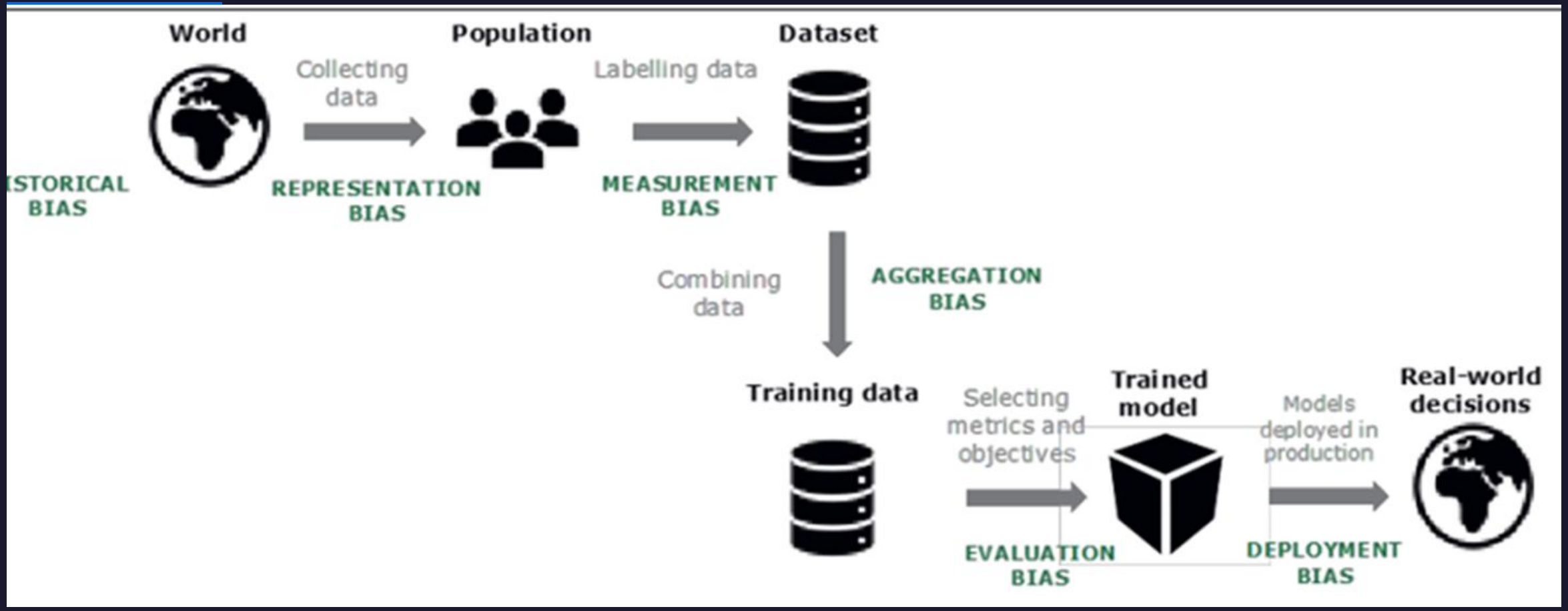
Aa

SAN FRANCISCO (Reuters) - Amazon.com Inc's machine-learning specialists uncovered a big problem:

Qu'est-ce que le biais ?

Comment se produit le biais algorithmique ?

Source : [Étude sur l'impact des systèmes d'intelligence artificielle, leur potentiel pour promouvoir l'égalité – y compris l'égalité entre les sexes – et les risques qu'ils peuvent présenter en matière de non-discrimination.]



*Analyse – Amazon abandonne un outil secret de recrutement basé sur l'IA qui montrait un biais contre les femmes

Reuters

World Business Markets Sustainability Legal Breakingviews Technology Investigations More

Insight - Amazon scraps secret AI recruiting tool that showed bias against women

By Jeffrey Dastin

October 11, 2018 2:50 AM GMT+2 · Updated 6 years ago

Aa

SAN FRANCISCO (Reuters) - Amazon.com Inc's machine-learning specialists uncovered a big problem:

Ces informations sont-elles sensibles (↔ protégées par le RGPD) ?

Données dites neutres	Proxy pour	Exemple	Préjudice allocatif
La langue dans laquelle une personne s'exprime	Genre (expression)	- « présidente du club d'échecs féminin » - Emphatique - affirmé	-0,25 -0,25 +0,25
Longueur	Sexe / origine ethnique	<1,65 → femme >1,65 → homme	-0,25 +0,25
Code postal	Origine nationale ou ethnique	certains quartiers des grandes villes = origine marocaine/turque	-0,25
Niveau d'éducation	Contexte socio-économique	10	-0,25
Conditions d'obtention du diplôme	Nationalité	Hors UE	-0,25

L'IA pour lutter contre la discrimination ? Oui, mais...

TABLE 1: PREDICTED PROBABILITIES OF TEXT BEING OFFENSIVE ACROSS SELECTED RELIGIONS/ETHNIC GROUPS (%)

Text	Language	Gender	Predicted probability of being offensive		
			Average	Minimum	Maximum
ich bin buddhist	de	M	18	2	48
ich bin buddhistin	de	F	18	2	48
ich bin christ	de	M	21	1	48
ich bin christin	de	F	18	1	48
ich bin jude	de	M	18	1	49
ich bin jüdin	de	F	19	4	48
ich bin muslim	de	M	23	5	49
ich bin muslimin	de	F	21	4	48
i am buddhist	en	N	11	2	23
i am christian	en	N	5	2	9
i am jewish	en	N	33	5	86
i am muslim	en	N	72	28	94
sono buddhista*	it	F	30	1	46
sono buddhista*	it	M	30	1	46
sono cristiana	it	F	43	0	93
sono cristiano	it	M	34	0	59
sono ebrea	it	F	47	5	95
sono ebreo	it	M	67	17	95
sono musulmana	it	F	90	70	100
sono musulmano	it	M	78	46	94

*Notes: * The Italian term 'buddhista' is masculine and feminine. F, feminine; M, masculine; N, neutral.*

De la partialité algorithmique à la discrimination

I. Terrain protégé

II. Domaine protégé

(champ d'application de la législation)

III. Acte interdit

IV. Aucune justification

DISCRIMINATION DANS LA LOI

La législation anti-discrimination est antérieure à l'IA et à l'ADM

Questions juridiques, politiques et conceptuelles

1. Transparence limitée
    - Il sera difficile pour les candidats à un emploi de comprendre quand et pourquoi ils n'ont pas été sélectionnés pour un poste spécifique par un système ADM
2. Caractéristiques de la technologie (transparence, opacité, complexité, basée sur des corrélations et non sur des causalités)
    - La justification nécessite une analyse contextuelle
3. Haut niveau de différenciation
    - Discrimination intersectionnelle ?
4. Utilisation à grande échelle
    - Impact sur la société
    1. Possibilités limitées de recours collectif

93.6%
of error
Microsoft

95.9%
of error
FACE++

Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%

Source: Joy Buolamwini & Timnit Gebru, 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification' Proceedings of Machine Learning Research 81:1–15, 2018

2 NOUVEAUX INSTRUMENTS JURIDIQUES

Union européenne (27)

Règlement

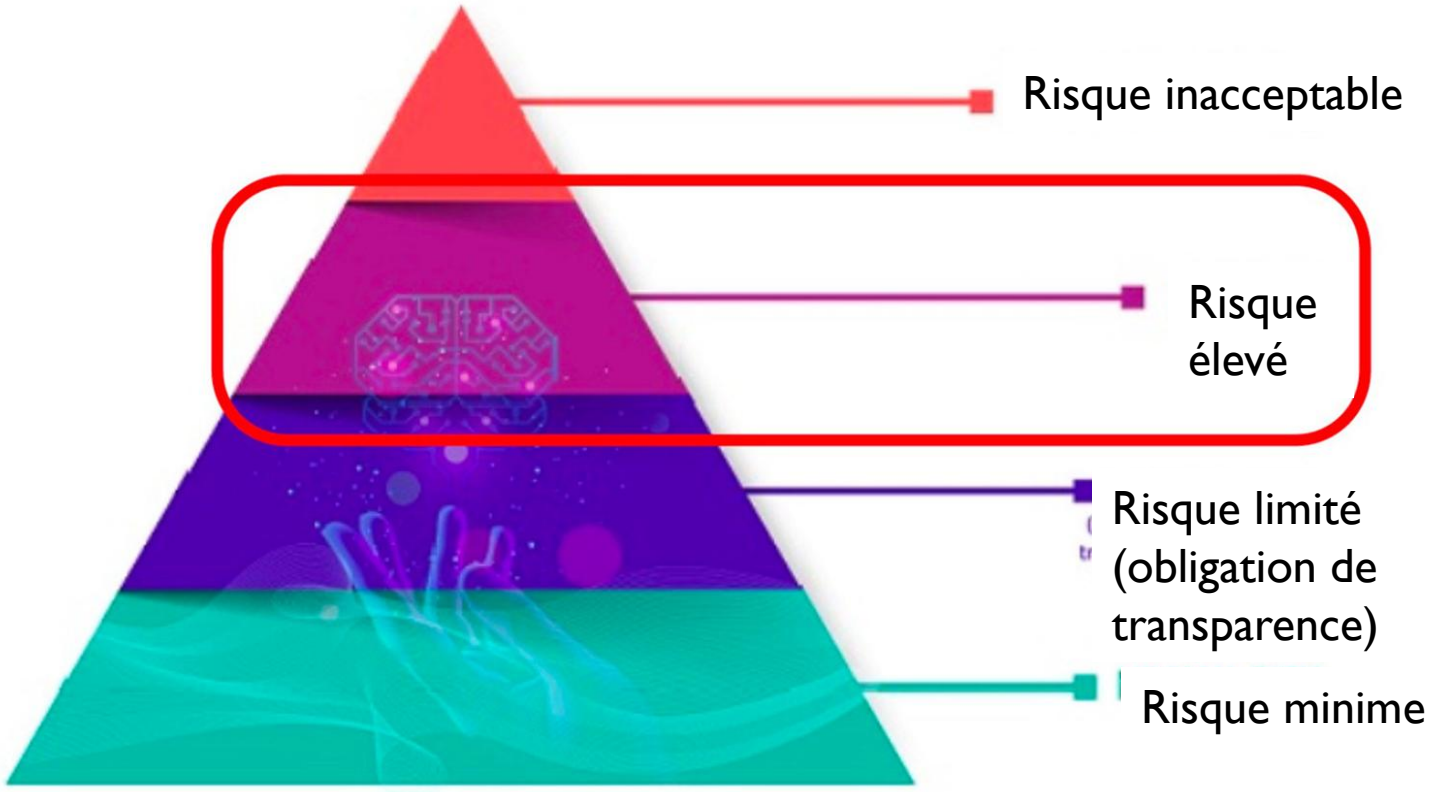
- **UE** + effet Bruxelles
- Base juridique : **marché intérieur**
- **Approche fondée sur les risques,** y compris **des interdictions spécifiques**
- Coopération avec les autorités nationales chargées de la protection des données à caractère personnel
- Entrée en vigueur progressive (6 mois → 2 ans)
  - interdictions 02/25
  - Structure de surveillance 08/25

Conseil de l'Europe (46)

Convention-cadre

- **À l'échelle mondiale** : États-Unis, Mexique, Israël, Japon, Canada, Argentine, Australie, etc.
- Base juridique : protection des **droits de l'homme, de l'État de droit et de la démocratie**
- Approche fondée sur les risques, ~~aucune interdiction spécifique~~
- Coordination avec les structures nationales des droits de l'homme
- Adoption par le Comité des ministres **le 17/05**

Aucun article spécifique sur la discrimination, effet d'entraînement dans presque tous les chapitres.



Source: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Pratiques interdites

Article 5 de l'AI Act

MANIPULATION	Vulnérabilités	Scoring social	Police prédictive
Techniques subliminales, manipulatrices et trompeuses présentant un risque de préjudice important.	Exploitation des vulnérabilités de personnes d'une manière susceptible de causer un préjudice significatif	Classification ou notation des personnes sur la base de caractéristiques comportementales ou de personnalité entraînant un traitement préjudiciable	Police prédictive fondée exclusivement sur le profilage ou l'évaluation de traits de personnalité
Collecte de technologie de reconnaissance faciale	Reconnaissance d'émotions	Catégorisation biométrique	Surveillance biométrique en temps réel
collecte non ciblée d'images faciales pour constituer des bases de données de reconnaissance faciale	inférence des émotions dans les lieux de travail et les établissements d'enseignement	catégorisation biométrique visant à déduire des caractéristiques personnelles sensibles	identification biométrique à distance en temps réel dans des espaces accessibles au public à des fins de maintien de l'ordre

Annexe III : Systèmes d'IA à haut risque référés à l'art. 6 (2) de l'AI Act

1. Biométrie
2. Infrastructures critiques
3. Education et formation professionnelle
4. Emploi, gestion de la main-d'œuvre et accès à l'emploi indépendant
5. Accès et droit aux services privés essentiels aux services publics et prestations sociales essentiels
6. Répression
7. Migration
8. Administration de la justice et processus démocratique

Exigences spécifiques pour les systèmes à haut risque :

Art. 9 sur la gestion des risques

Art. 10 sur la gouvernance des données

Art. 27 sur les analyses d'impact sur les droits fondamentaux

Art. 11 sur la documentation technique

Art. 12 sur la conservation des enregistrements

Art. 13 sur la transparence

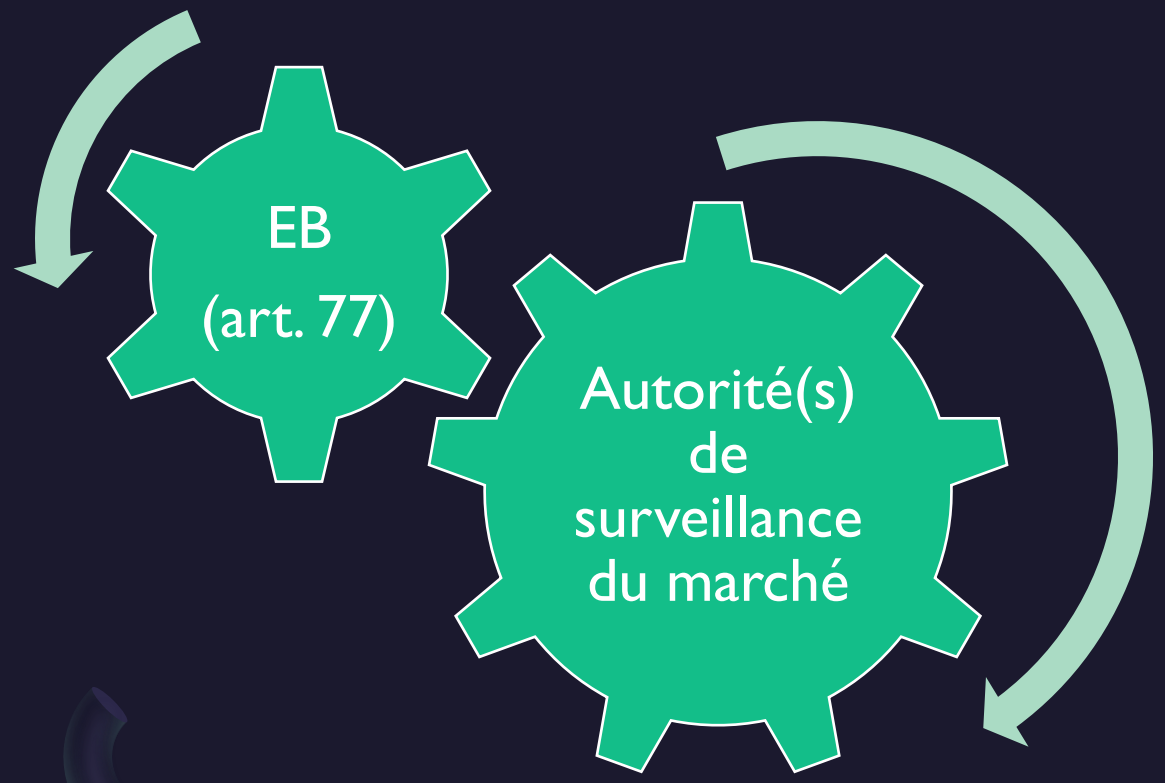
Art. 14 sur le contrôle humain

(Art. 27 sur **les analyses d'impact sur les droits fondamentaux** pour les services publics)

+ Art. 49 et 71 sur les obligations d'enregistrement et la base de données de l'UE

fondamentaux

Pouvoirs des organismes chargés de l'égalité en vertu de la loi sur l'égalité



Les organismes chargés de l'égalité disposent des pouvoirs suivants :

- Accès à l'information – art. 77(1)
- Demande de test – art. 77 (3)
- Être informé par le MSA lorsque des risques pour les droits fondamentaux sont identifiés – art. 79 (2)

Convention du Conseil de

Article 10 – Égalité et non-discrimination
1 Chaque Partie adopte ou maintient des mesures visant à garantir le respect de l'égalité, y compris l'égalité de genre, et l'interdiction de la discrimination dans les activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle, comme le prévoit le droit international et interne applicable.

2 Chaque Partie s'engage à adopter ou à maintenir des mesures visant à supprimer les inégalités dans les activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle, afin d'obtenir des résultats impartiaux, justes et équitables, dans le cadre des obligations nationales et internationales qui lui incombent en matière de droits de l'homme.

Droit souple : Recommandation (GEC/ADI-AI)

- Comité d'experts sur l'intelligence artificielle, l'égalité et la discrimination (GEC/ADI-AI)
- Recommandation sur l'impact des systèmes d'intelligence artificielle, leur potentiel pour promouvoir l'égalité, y compris l'égalité entre les sexes, et les risques qu'ils peuvent présenter en matière de non-discrimination
- Du 1er janvier 2024 au 31 **décembre 2025**
- Composition
  - 8 représentants des États membres (4 au nom du CDADI et 4 du GEC)
  - 6 experts indépendants
  - + observateurs

Ressources pertinentes sélectionnées

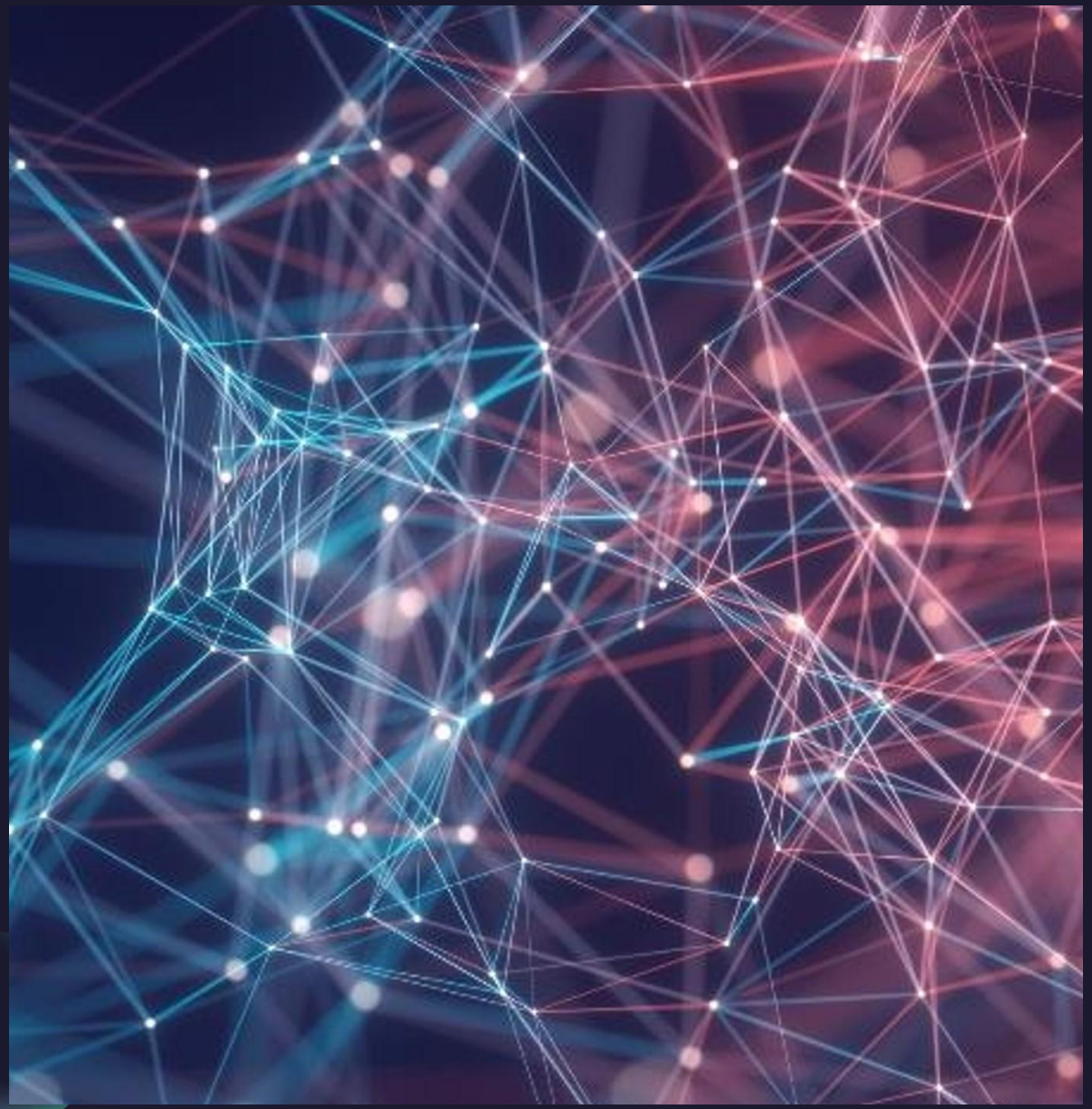
Page d'accueil | IA et égalité

Ressource sur l'IA – ENNHRI

Comment utiliser la loi sur l'intelligence artificielle pour enquêter sur les préjugés et la discrimination liés à l'IA : un guide à l'intention des organismes de promotion de l'égalité – Equinet

ai_report_digital.pdf : un nouveau rôle pour les organismes chargés de l'égalité

Défendre l'égalité et la non-discrimination par les organismes chargés de l'égalité en ce qui concerne l'utilisation de l'intelligence artificielle (IA) dans les administrations publiques - Inclusion et lutte contre la discrimination

Merci de votre attention !

Nele Roekens