



● Les défis de l'Intelligence Artificielle entre éthique et performance

Daniela Oliveira, PhD

Colloque Performance et valorisation des
données

2023-03-23

Plan

Introduction

Où l'IA peut nous aider

Défis de l'utilisation de l'IA

Cas concrets

Des leçons apprises

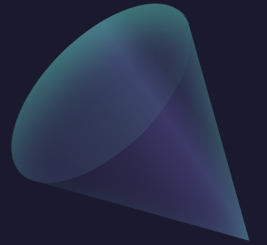
Et alors?



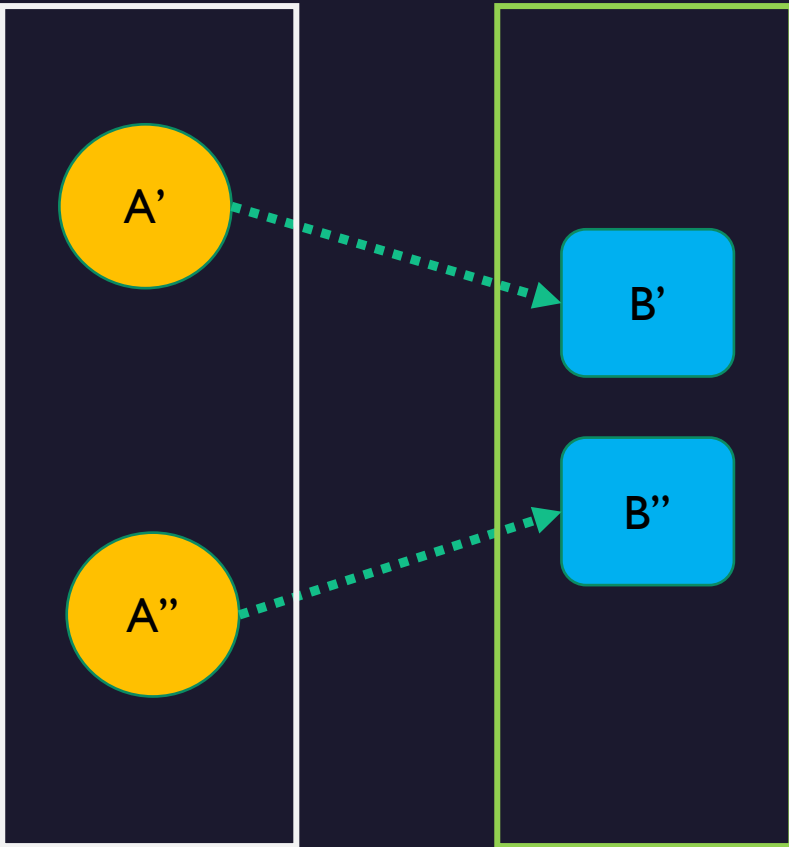


Introduction

Programmation usuelle

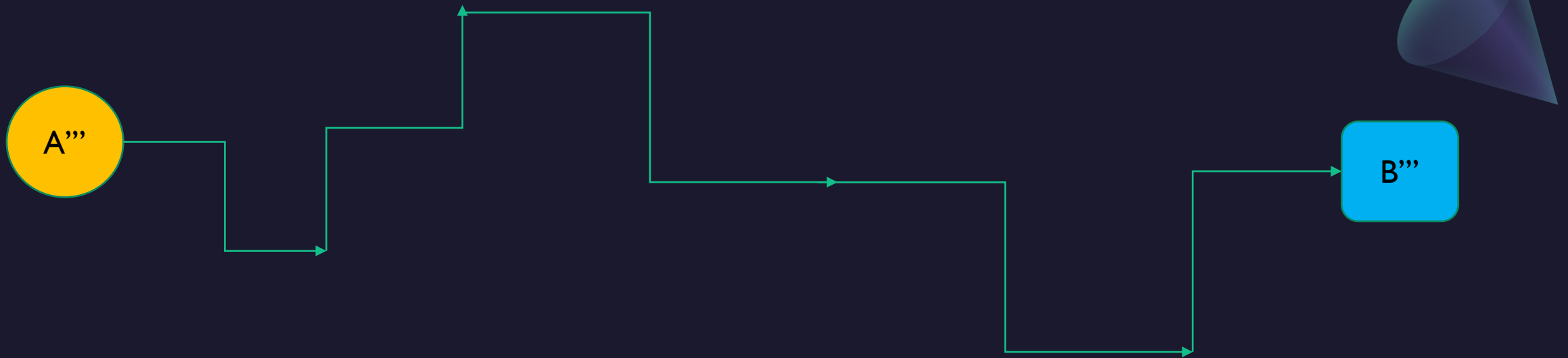


Apprentissage automatique



IA = programmation de la recherche de fonctions
(apprentissage automatique)

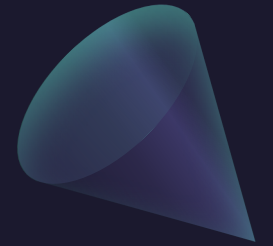
IA Symbolique



ou d'arbres de décision complexes (IA symbolique)

L'IA

- Apprentissage automatique
 - Les fonctions entre caractéristiques et résultats
- IA symbolique
 - Arbres de decision automatisées par l'humain
- Répétition de patrons



Pourquoi l'IA

Augmente le potentiel
d'action

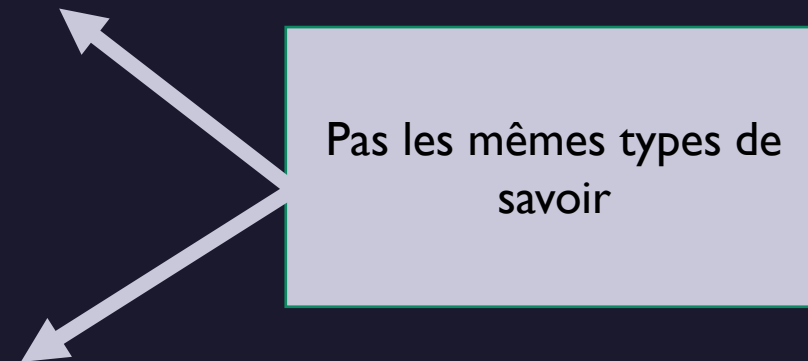
- Capacité de calcul
 - Dépasse largement celle de l'humain
- Automatisation de processus cognitifs
 - Parfois tacites
 - Peut mener un professionnel novice à agir de la même façon qu'un professionnel chevronné

Où l'IA peut nous aider

Là où des patrons du passé peuvent être un chemin d'action dans le présent

Si on y pense...

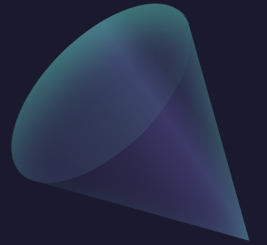
- Identification de tendances de l'action
 - Des stratégies de traitement réussies mais sans étude clinique formelle
 - Des caractéristiques de demandes réussies
- Des analyses complexes
 - Analyse de différents aspects d'un dossier
 - Optimisation d'horaire en tenant compte de la disponibilité des ressources humaines, matérielles et budgétaires



Mais...

Défis de l'utilisation de l'IA

- Apprentissage automatique
 - Très dépendante de la fréquence et distribution de données
- IA symbolique
 - Dépendante de l'existence d'expertise
 - Savoir doit évoluer



Enjeux - Apprentissage automatique

- Exemple

- Le programme X existe depuis 2010 pour des adultes (18 ans et plus)
- Des données sont disponibles depuis 2018
- Depuis 2019 des personnes âgées de plus de 16 ans sont admises

- Enjeu

- Les données existantes ne représentent peut-être pas tout le parcours du programme
- Si le jeu de données est composé de tous les cas depuis 2018 et traité de la même façon, des tendances de la clientèle adulte se imposeront à la clientèle de 16-18 ans



BIAIS

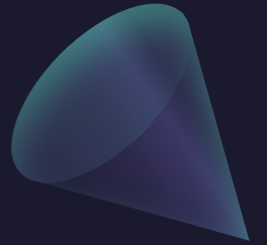
Enjeux – IA symbolique

- Exemple

- L'analyse de l'admission au programme X a été faite par des experts dans le domaine

- Enjeu

- Les novices ont eu du contact avec la solution mais pas avec les experts
- Les connaissances du domaine ont évolué
- Il n'y a plus d'experts dans l'organisation

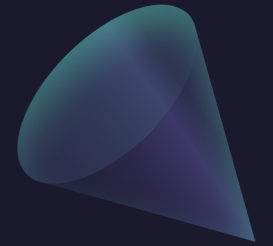


Cas concrets

Et réflexions

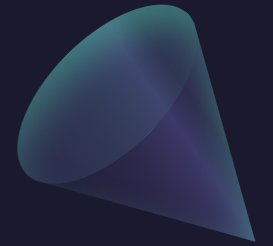
Cas concret 1 – Apprentissage automatique

- Un programme d'accompagnement accru de patients a été mis sur pied
- Des contraintes de ressources humaines et financières ont mené à l'inclusion seulement de patients **sévèrement malades**
- Design de la solution:
 - Caractéristique: “sévèrement malade”
 - Résultat: admis dans le programme



Cas concret 1 – Apprentissage automatique (suite)

- Jeu de données utilisé:
 - Pas de couleur de peau
 - “Sévèrement malade” = A dépensé plus que X dans l’année précédente
- Utilisation: sur plus de 200 millions de cas

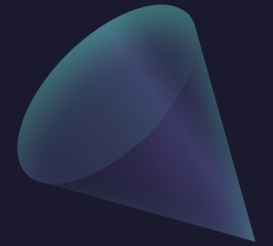


Cas concret 1 – Apprentissage automatique (suite)

- Évaluation A: Comparaison avec des données sur l'identification en tant que Noirs
- Résultat A: Relativement plus de patients non-Noirs étaient admis

Il y a-t-il un problème?

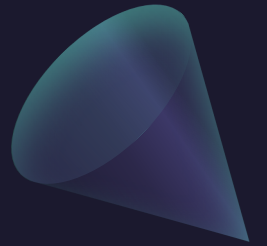
- Évaluation B: Comparaison avec des données sur la comorbidité (présence de plus qu'une condition nuisible à la santé)
- Résultat B: 29% plus de patients Noirs devraient être admis



Cas concret 1 – Apprentissage automatique (suite)

Quel est le problème?

- “Sévèrement malade” = A dépensé plus que X dans l’année précédente
 - Les patients Noirs dépensent moins pour le même niveau de comorbidité



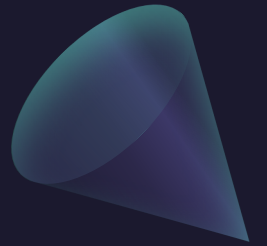
Si on y pense encore...

- Identification de tendances de l'action
 - Des stratégies de traitement réussies mais sans étude clinique formelle
 - Des caractéristiques de demandes réussies
- Des analyses complexes
 - Analyse de différents aspects d'un dossier
 - Optimisation d'horaire en tenant compte de la disponibilité des ressources humaines, matérielles et budgétaires

« Réussies »?

Cas concret 2 – IA symbolique

- Système de classification d'ostéo-arthrite: Kellgren-Lawrence (1957) – IA symbolique
- Jeu de données utilisé:
 - Mineurs dans Lancashire, en Angleterre
 - Pas de mention de couleur de peau ou sexe
- Utilisée pour prédire le niveau de douleur à partir de rayons-X



Cas concret 2 – IA symbolique (suite)

- Évaluation A: Comparaison avec de la prédiction de la douleur par des professionnels de la santé versus description de la douleur par des patients
- Résultat A: Des patients avec moins de revenu, moins d'éducation ou Noirs ont plus de douleur que prévu par les professionnels de la santé

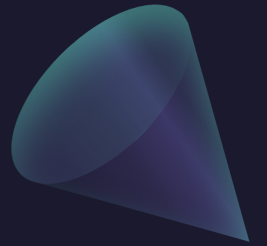
Il y a-t-il un problème?

- Évaluation B: Identification de tendance (Apprentissage automatique) avec rayons-X plus description de douleur par le patient
- Résultat B: L'application est capable de prédire avec précision le niveau de douleur du patient

Cas concret 2 – IA symbolique (suite)

Quel est le problème?

- Échantillon utilisé pour la classification n'est pas significatif de la population où cette classification est appliquée

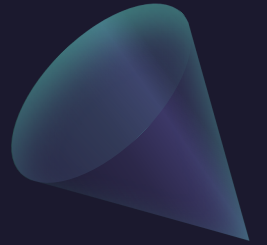


Si on y pense encore une fois...

- Identification de tendances de l'action
 - Des stratégies de traitement réussies mais sans étude clinique formelle
 - Des caractéristiques de demandes réussies
- Des analyses complexes
 - Analyse de différents aspects d'un dossier
 - Optimisation d'horaire en tenant compte de la disponibilité des ressources humaines, matérielles et budgétaires

« Réussies »?

« Complexes »?



Défis de l'utilisation de l'IA

- Apprentissage automatique
 - Très dépendante de la fréquence et distribution de données
- IA symbolique
 - Dépendante de l'existence d'expertise
 - Savoir doit évoluer

« Évolution du savoir »?

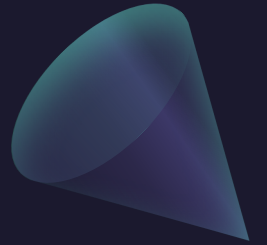


Des leçons appprises

Et suite

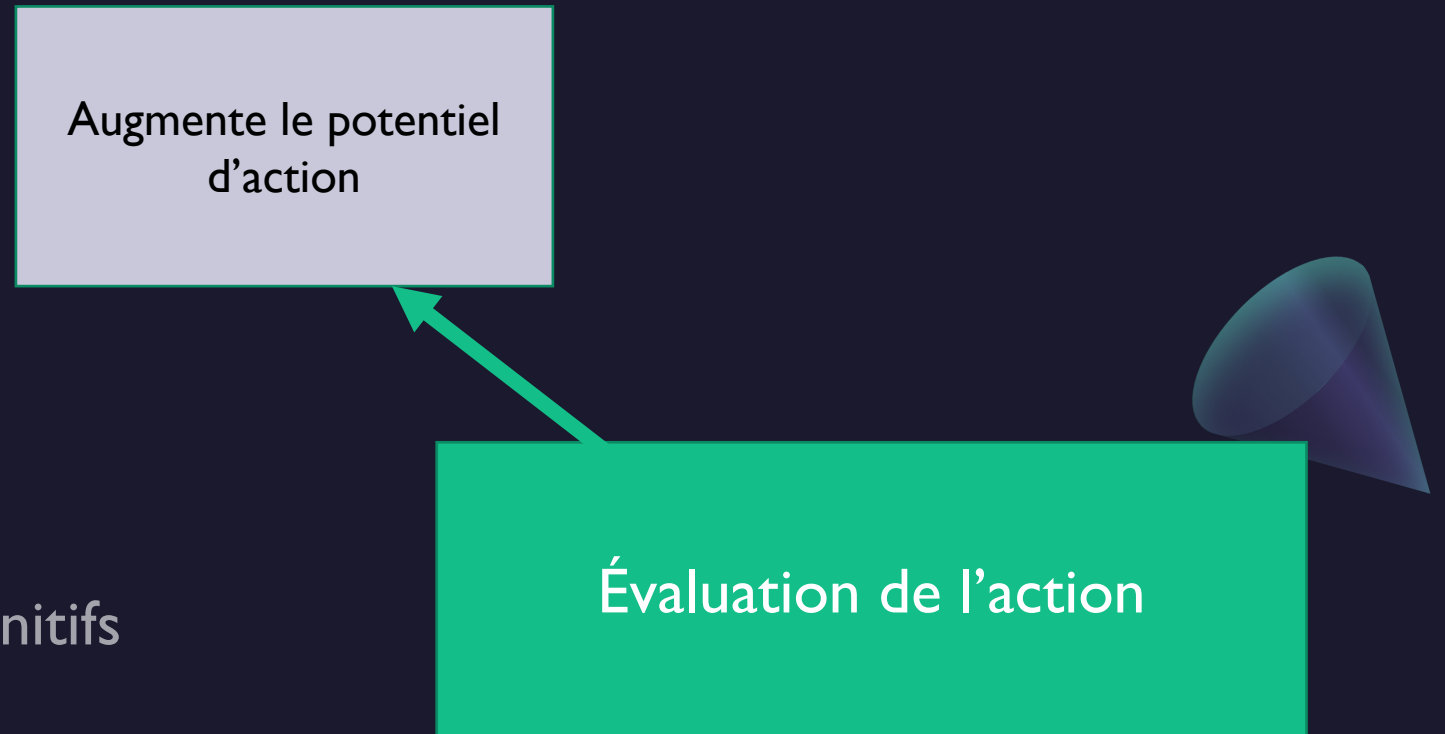
Évaluation de l'IA

- L'adéquation au contexte dépend d'une réflexion sur ce contexte
 - Quels sont les savoirs impliqués?
 - Quel est le contexte d'utilisation?
 - Qu'est-ce que c'est qu'une "réussite"?
- Le type de technologie n'est pas un gage de réussite
 - L'adéquation de la technologie au contexte d'utilisation l'est



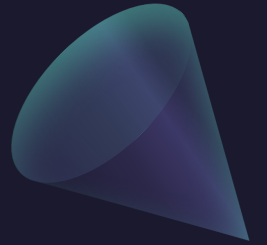
Pourquoi l'IA

- Capacité de calcul
 - Dépasse largement celle de l'humain
- automatisation de processus cognitifs
 - Parfois tacites
 - Peut mener un professionnel novice à agir de la même façon qu'un professionnel chevronné



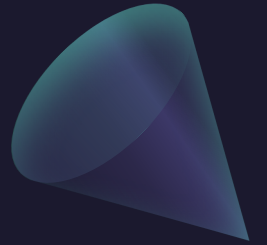
Évaluation de l'IA = évaluation de l'action

- L'automatisation de processus cognitifs ne doit pas suivre la même gestion que l'automatisation de tâches
 - Les risques sont différents
 - Adéquation au contexte est beaucoup plus importante pour l'IA
- = Évaluation de la performance



L'IA et l'évaluation de la performance

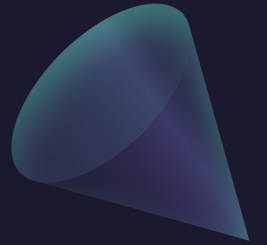
- Planification de l'évaluation de la performance avant adoption
 - Quel est le contexte d'utilisation?
 - Qu'est-ce que c'est qu'une "réussite"?
 - Quelles sont les contraintes?
- Indicateurs à partir du contexte d'utilisation



Mais...

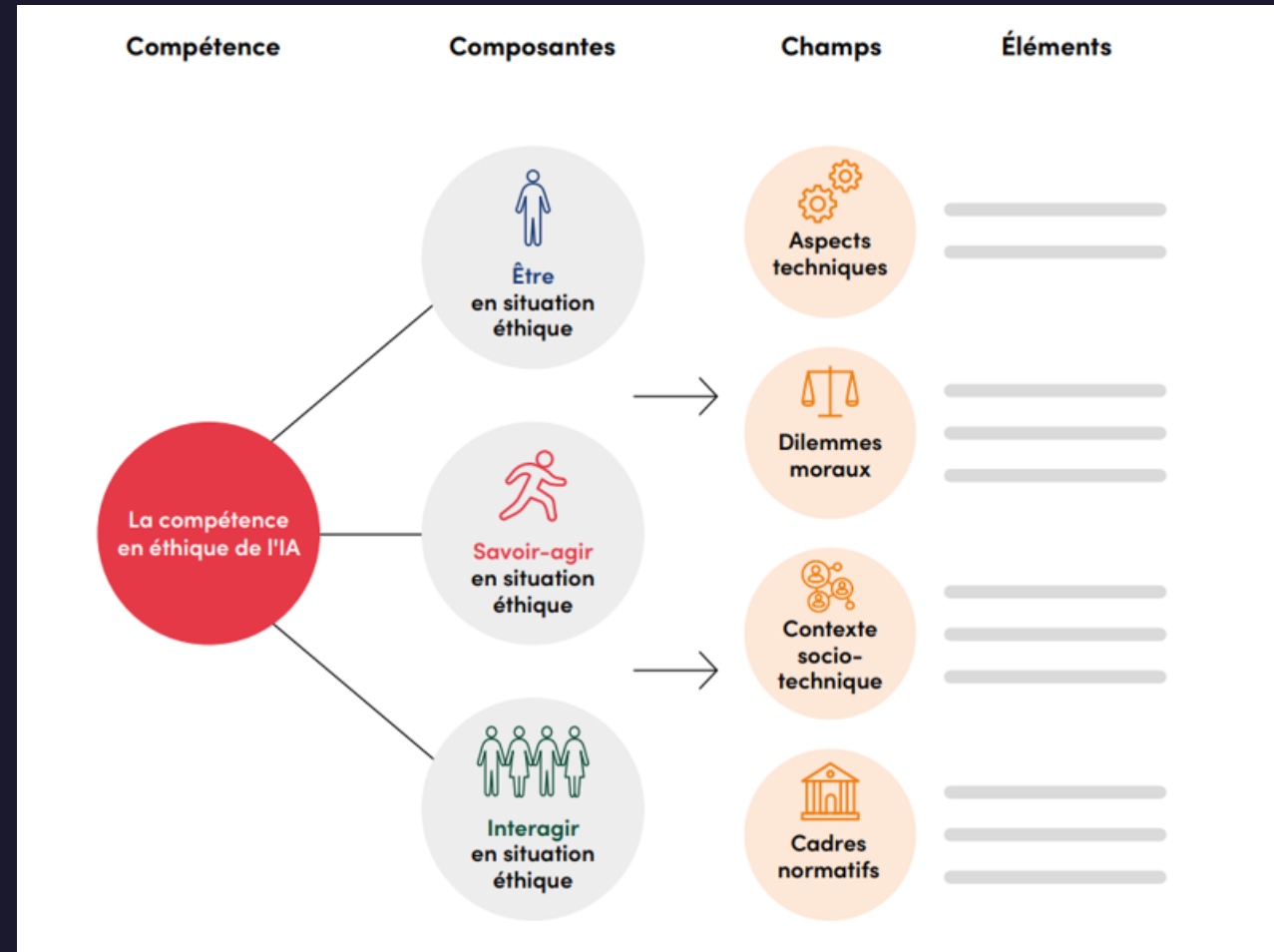
Qui?

- Qui est capable de répondre à toutes ces questions?
 - Les développeurs en IA?
 - Les gestionnaires?
 - Les employés administratifs?
 - Les patients?
 - L'utilisateur? Le client? Le donneur d'ouvrage?



Toutes ces réponses – l'IA responsable appliquée

- Processus collaboratif et continu

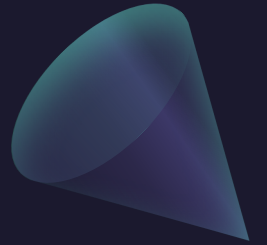




Sur une note d'espoir...

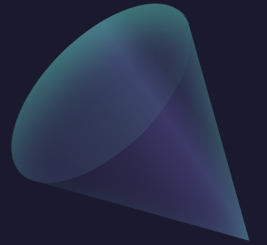
Formation (entraînement) de données responsable

- Dans le premier cas concret, l'algorithme a été formé à nouveau
 - Collaboration avec l'équipe de chercheurs
 - Biais réduit de 84%



Directives pour une IA responsable appliquée

- *Blueprint for an AI Bill of Rights*
 - Départ en octobre 2022





En bref

L'IA potentialise l'action humaine. C'est important que cette potentialisation soit accompagnée et que l'impact de cette action, évalué

Pour aller plus loin

Référentiel de compétence en éthique de l'IA

Déclaration de Montréal pour un développement responsable de l'IA (declarationmontreal-iaresponsable.com)

En anglais:

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447-453. <https://doi.org/10.1126/science.aax2342>

Pierson, E., Cutler, D. M., Leskovec, J., Mullainathan, S., & Obermeyer, Z. (2021). An algorithmic approach to reducing unexplained pain disparities in underserved populations. *Nature Medicine*, 27(1), Article 1. <https://doi.org/10.1038/s41591-020-01192-7>

March 20, 2023 - Conversation@Mila - Y. Bengio & K. Crawford on AI's future/Le futur de l'IA – YouTube

