

Le contrôle humain et “due process”

Colloque Décision humaine, décision de l'IA

Université d'Artois

Winston Maxwell

9 novembre 2023



Projet ANR-20-CHIA-0023

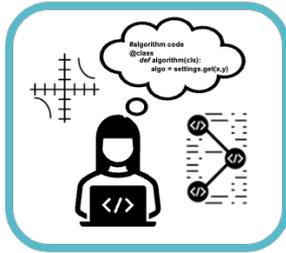
Explainable AI for Anti-Money Laundering



INSTITUT
POLYTECHNIQUE
DE PARIS

Les contrôles “système” ou “individuels”

Les contrôles “système”



Les vérifications “individuelles”



Human in
the loop



Contestation
ex post

Les trois finalités du contrôle humain

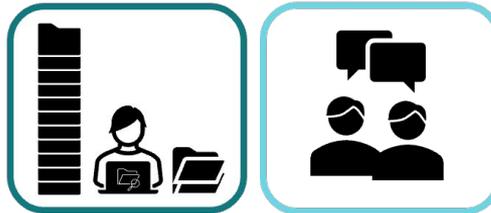
Détecter les erreurs

- biais et erreurs aléatoires
- détecter les discriminations
- vérifier la performance prédictive



Due process

- Dignité humaine
- Etre jugé par un autre humain
- Participer à la décision
- “Réversibilité des rôles”

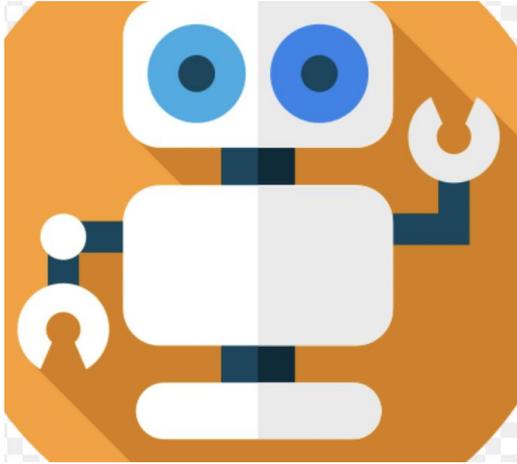


Responsabilité

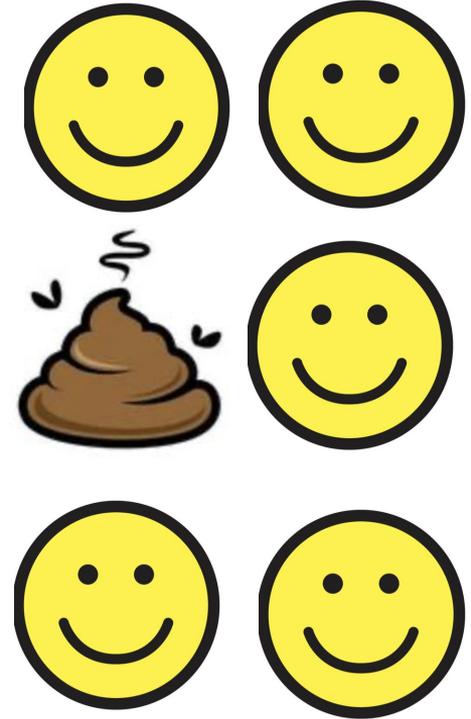
- Fixer la responsabilité
- Justifier une décision
- Responsibility gap
- Démontrer sa conformité



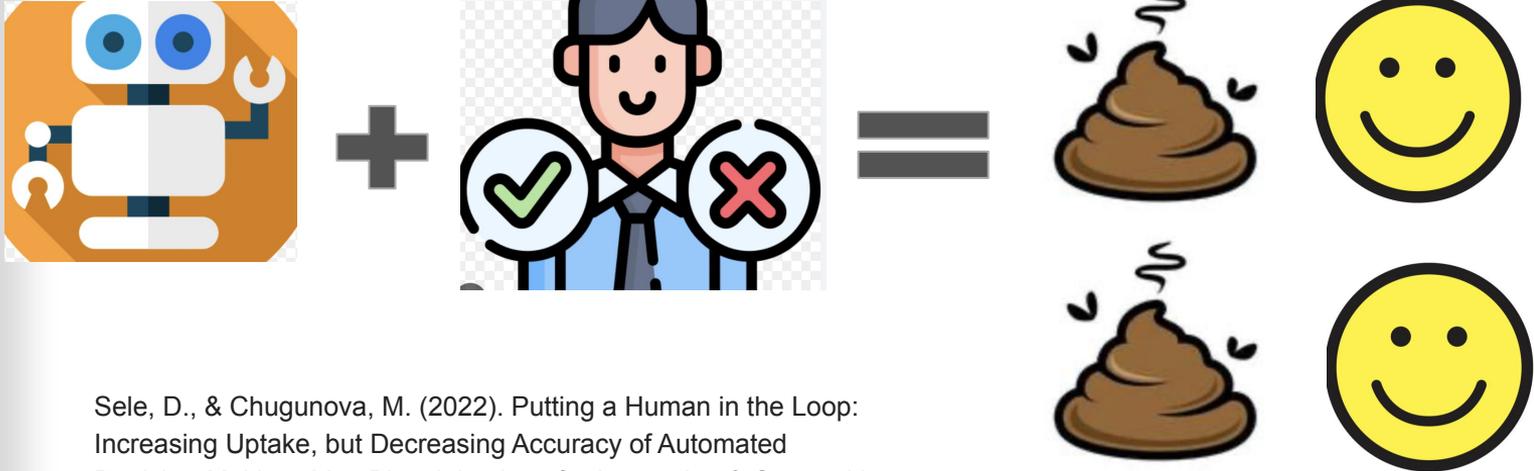
Un algorithme peut se tromper



=



Un humain peut augmenter les erreurs



Sele, D., & Chugunova, M. (2022). Putting a Human in the Loop: Increasing Uptake, but Decreasing Accuracy of Automated Decision-Making. *Max Planck Institute for Innovation & Competition Research Paper*, (22-20).

A. Z. HUO, « A right to a human decision », *Virginia law review*, vol. 106, 2020, p. 611.

A quoi sert l'humain?



Détecter les erreurs

- biais et erreurs aléatoires
- détecter les discriminations
- vérifier la performance prédictive

Due process

- Dignité humaine
- Etre jugé par un autre humain
- Participer à la décision
- “Réversibilité des rôles”

L'approche "due process"



Solove, Daniel J. and MATSUMI, Hideyuki, AI, Algorithms, and Awful Humans (October 16, 2023). 96 Fordham Law Review (forthcoming 2024), Available at SSRN: <https://ssrn.com/abstract=4603992>



Due process : une valeur de procédure

Principe constitutionnel américain applicable aux décisions prises par l'Etat

Plusieurs litiges concernant les décisions fondées en partie sur des scores algorithmiques (Loomis v. Wisconsin, Houston Teachers)

Etendre les principes de *due process* à certaines décisions algorithmiques du secteur privé (Williams, 2022)

Due process: synonyme d'un processus de décision respectueux des droits

Due process: trois éléments clés

La participation active de la personne affectée par la décision

L'existence d'un décisionnaire humain

La "rationalité" de la décision

RS Summers, Evaluating and Improving Legal Processes - A Plea for "Process Values", 60 Cornell L. Rev. 1, 12 (1974)

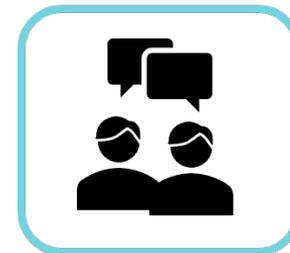
Henry J. Friendly, "Some Kind of Hearing", 123 U. Pa. L. Rev. 1267 (1975).

Martin H. Redish and Lawrence C. Marshall, Adjudicatory Independence in the Values of Procedural Due Process, 95 Yale L. J. 455 (1986).

Une participation

La personne doit pouvoir:

- Présenter ses arguments
- Avoir accès aux informations détenues par le décideur pour pouvoir les commenter



Principes qui recourent “procès équitable”,

La participation devient une valeur intrinsèque

Participation: avant ou après la prise de décision?

Frederick Mostert, ‘Digital due process’: a need for online justice, *Journal of Intellectual Property Law & Practice*, Volume 15, Issue 5, May 2020, Pages 378–389, <https://doi.org/10.1093/jiplp/jpaa024>

Kaminski, Margot E. and Urban, Jennifer M., The Right to Contest AI (November 16, 2021). *Columbia Law Review*, Vol. 121, No. 7, 2021, Available at SSRN: <https://ssrn.com/abstract=3965041>.

Un décisionnaire humain

Matthews v. Eldridge: une décision mineure ne nécessite pas forcément un décisionnaire humain

Pourquoi un humain pour les décisions importantes?

Réversibilité de rôles



A. Z. HUO, « A right to a human decision », *Virginia law review*, vol. 106, 2020, p. 611.

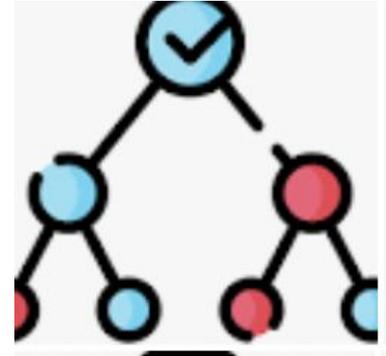
K. Brennan-Marquez et S. Henderson, "Artificial Intelligence and Role-Reversible Judgment, 109 J. of Crim. L. & Criminology 137 (2019)

La rationalité

Un cheminement logique entre les entrées et la sortie

Un défi pour machine learning

Une explication n'est pas nécessairement une justification



RS Summers, Evaluating and Improving Legal Processes - A Plea for "Process Values", 60 Cornell L. Rev. 1, 12 (1974)

Clement Henin, Daniel Le Métayer. A Framework to Contest and Justify Algorithmic Decisions. AI and Ethics, Springer, 2021, 1, pp.463-476.

Schauer, Frederick. "Giving Reasons." Stanford Law Review, vol. 47, no. 4, Stanford Law Review, 1995, pp. 633-59, <https://doi.org/10.2307/1229080>.

Conclusions



1. Distinguer les finalités du contrôle humain, et adapter les modalités (et KPIs) du contrôle humain à chaque finalité
2. L'objectif "contrôle d'erreurs"
3. Tester l'équipe humain/machine par rapport aux KPIs liés à chaque finalité (erreurs, procédure équitable, responsabilité)

Pour plus d'informations...



Fichier principal



Le contrôle humain
des systèmes
algorithmiques -
mémoire original
HDR W. Maxwell -
5.pdf (1.18 Mo)



LE CONTRÔLE HUMAIN DES SYSTÈMES ALGORITHMIQUES - UN REGARD CRITIQUE SUR L'EXIGENCE D'UN "HUMAIN DANS LA BOUCLE"

Mémoire original pour présenter l'habilitation à diriger des recherches
de l'Université Panthéon-Sorbonne, soutenue le 25 novembre 2022 par
Winston Maxwell*

Jury:

Célia Zolynski, Professeure, Université Panthéon Sorbonne, Présidente du Jury
Brunessen Bertrand, Professeure, Université de Rennes, Rapporteur
Nicolas Curien, Professeur, Conservatoire National des Arts et Métiers, Rapporteur
Judith Rochfeld, Professeure, l'Université Panthéon-Sorbonne, Garante
Jean-Yves Ollier, Conseiller d'Etat
Céline Castets-Renard, Professeure University of Ottawa

Origin : Files produced by the author(s)

Licence : CC BY NC ND - Attribution - NonCommercial - NoDerivatives

HAL : tel-04010389

MERCI !