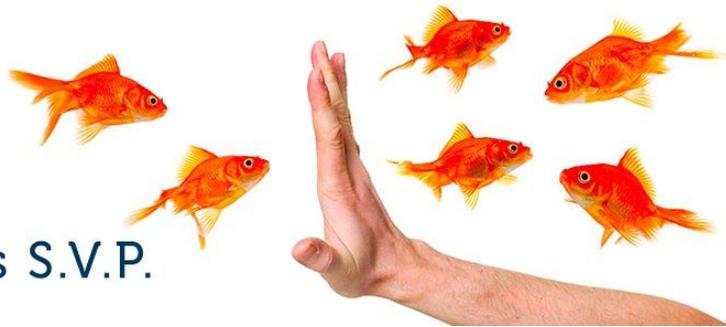


## Algorithmes : contrôle des biais S.V.P.



[CLIQUEZ ICI POUR DÉCOUVRIR  
NOTRE RAPPORT](#)

Paris, le 10 mars 2020 - La stratégie de la Commission européenne en matière de données et d'intelligence artificielle, rendue publique fin février, pousse toujours plus loin la réflexion sur une IA responsable, centrée sur l'humain. Mais l'éthique de l'IA soulève immanquablement une question cruciale : celle des algorithmes, de leur neutralité, ou au contraire de leurs biais - susceptibles d'induire des discriminations dans les décisions humaines. Dans ce contexte, l'Institut Montaigne publie aujourd'hui un rapport sur la lutte contre les biais algorithmiques – condition *sine qua non* d'une IA européenne éthique et performante.

*“Le sujet est majeur : le numérique et l'IA ne pourront pas se développer en France et en Europe s'ils sont porteurs de discriminations massives. Aujourd'hui pourtant, les biais algorithmiques restent majoritairement traités sous un prisme américain.”*, déclarent les deux présidents du groupe de travail en charge de ce rapport, Anne Bouverot, présidente du conseil d'administration de Technicolor et présidente de la Fondation Abeona et Thierry Delaporte, directeur général délégué, Capgemini.

### Nos “non-propositions”

Face à ces enjeux, il faut être clair : nous ne recommandons **ni une loi contre les biais des algorithmes commune à tous les secteurs d'activités, ni un contrôle systématique par l'Etat de l'absence de biais dans tous les algorithmes.**

Il existe d'ores et déjà de nombreux textes s'attaquant aux discriminations. Ceux-ci s'appliquent au monde physique comme au monde numérique et sont de nature à limiter le risque de biais. Par ailleurs, le règlement général sur la protection des données (RGPD) a montré que l'usage des données personnelles est bien trop répandu pour qu'une autorité administrative puisse en contrôler l'intégralité avant leur utilisation.

Afin de minimiser efficacement les risques de discriminations, nous insistons sur la nécessité de responsabiliser les acteurs qui développent et utilisent les algorithmes. Nous formulons ainsi quatre recommandations prioritaires.

## Nos recommandations

- **Tester la présence de biais dans les algorithmes comme l'on teste les effets secondaires des médicaments**

C'est *in fine* par le test de l'absence de biais qu'il est possible de créer la confiance dans le caractère équitable des algorithmes. Tester l'équité a un coût, mais les développeurs et acheteurs d'algorithmes devront intégrer cette contrainte, et recourir à des tests fonctionnels ou de performance pour s'assurer de l'absence de biais. Dans certains cas où la création de ces bases de données est difficile ou problématique, l'Etat pourrait la prendre en charge.

- **Promouvoir une équité active, plutôt que d'espérer l'équité en ne mesurant pas la diversité**

Un algorithme peut introduire des biais contre les femmes, même si l'on a explicitement exclu le genre des variables utilisées — il est en effet facile de déduire cette information à partir d'autres informations, comme le fait d'acheter en ligne des soins pour femme. Lutter contre les discriminations suppose de passer d'une approche qui attend l'équité par l'ignorance de la diversité à une équité active, c'est-à-dire accepter que l'équité d'un algorithme s'obtient en incluant les variables protégées comme le sexe, l'âge ou encore la religion et en testant l'indépendance du résultat vis-à-vis de ces variables.

Cependant, la collecte et l'utilisation de ces données doivent donc être strictement encadrées ! Afin d'éviter les dérives, une telle collecte serait limitée aux tests de la présence de biais, entreprise par un acteur tiers, et restreinte à un échantillon des utilisateurs concernés. Par ailleurs, une telle approche devrait faire l'objet d'une analyse d'impact déclarée à la CNIL de manière préalable. La nature des algorithmes testés devra justifier la collecte de telles données (voir ci-dessous).

- **Être plus exigeant pour les algorithmes ayant un fort impact sur les personnes (droits fondamentaux, sécurité, accès aux services essentiels)**

Nous définissons un algorithme à fort impact comme un algorithme dont les décisions génèrent un ou plusieurs des types d'impacts suivants : le déni d'accès à des services essentiels, l'atteinte à la sécurité des personnes, la restriction des droits fondamentaux.

Pour ces algorithmes, nous recommandons un cadre *ad hoc* prévoyant des obligations de transparence en ce qui concerne les données utilisées et les objectifs fixés à l'algorithme, ainsi qu'un droit de recours contre la décision prise. La création d'un tel cadre ne nécessite pas une nouvelle loi sur les biais des algorithmes mais plutôt la mise en œuvre de bonnes pratiques dans les entreprises et administrations, l'usage de dispositifs juridiques existants et l'ajout au cas par cas de dispositions dans des législations sectorielles.

- **Assurer la diversité des équipes de conception et de déploiement des algorithmes**

Définir le comportement équitable de l'algorithme permet d'en modifier profondément les impacts économiques et sociétaux. Il est plus que jamais nécessaire d'assurer une pluridisciplinarité dans la prise de ce type de décisions qui ne peut être le seul fait d'experts techniques. Au-delà de la diversité professionnelle, il est désormais clair que des équipes socialement diverses sont mieux armées pour prévenir les biais, pour éviter de reproduire des discriminations.

Intégrer une diversité de profils, de compétences, d'expériences, d'âges, de genres dans les équipes de conception, de production et de pilotage des algorithmes doit devenir une norme pour prévenir les biais algorithmiques.

*Et pour aller plus loin, nous proposons de ...*

- Former les techniciens et ingénieurs aux risques de biais, et améliorer la connaissance citoyenne des risques et opportunités de l'IA.
- Mettre à disposition des bases de données de test publiques pour permettre aux entreprises d'évaluer les biais de leur méthodologie.
- Soutenir l'émergence de labels pour renforcer la confiance du citoyen dans les usages critiques et accélérer la diffusion des algorithmes bénéfiques.
- Donner à la puissance publique une capacité d'audit des algorithmes à fort impact.

**CLIQUEZ ICI POUR DÉCOUVRIR  
NOTRE RAPPORT**

Nous vous attendons sur [Twitter](#), [Facebook](#) et sur [Instagram](#).  
Inscrivez-vous à notre [Newsletter](#).

**Contact presse** : Lara Oliveau, responsable de la communication et des relations presse  
[06 99 79 43 62](tel:0699794362) ; [loliveau@institutmontaigne.org](mailto:loliveau@institutmontaigne.org)

**À propos de l'Institut Montaigne :**

*Think tank indépendant créé en 2000, l'Institut Montaigne est une plateforme de réflexion, de propositions et d'expérimentations consacrée aux politiques publiques en France et en Europe. Ses travaux sont le fruit d'une méthode d'analyse et de recherche rigoureuse et critique, ouverte sur les comparaisons internationales. L'Institut Montaigne, association à but non lucratif pionnière en France, réunit des chefs d'entreprise, des hauts fonctionnaires, des universitaires et des personnalités issues d'horizons divers. Ses financements sont exclusivement privés, aucune contribution n'excédant 1,5 % d'un budget annuel de 5,9 millions d'euros (estimation 2019). À travers ses publications et les événements qu'il organise, l'Institut Montaigne souhaite jouer pleinement son rôle d'acteur du débat démocratique.*