

## [InternetActu.net-7July2021-Utiliser l'IA pour détecter ses propres biais ?](#)

Posted: 07 Jul 2021 12:00 AM PDT

Pour les ingénieurs, bien souvent, la question des biais algorithmiques n'est qu'un problème technique à corriger. Le fait qu'une IA ne soit ni neutre, ni loyale, ni équitable n'est finalement qu'une question de modélisation à ajuster, de données à corriger, de calculs à améliorer...

Pour remédier aux discriminations, il suffirait finalement de calculer des mesures correctrices proportionnelles au niveau de discrimination, une **discrimination positive** en quelque sorte. Ce n'est peut-être pas si simple...

**Kate Crawford (@katecrawford) le disait très bien** : quelle correction appliquée ? La question est bien plus compliquée qu'une correction relative à des problèmes de physique, comme de corriger les turbulences d'un avion ou le freinage d'une voiture. Modéliser la société n'est pas la même chose que modéliser des problèmes de physique, **disait déjà le physicien Pablo Jensen...**

Reste que la question de la correction des biais des systèmes fait de plus en plus l'objet d'une attention forte des autorités. Et trouver les préjugés de l'IA est devenue une activité en plein essor pour les startups et les grands noms de la technologie, rapporte le journaliste Cade Metz (@cademetz, qui vient de publier **Genius Makers**, Penguin, 2021) **dans le New York Times**.

Le National Institute of Standards and Technology a publié récemment **une proposition** détaillant la manière dont les entreprises peuvent lutter contre les préjugés de leurs systèmes. Fin 2019, **les régulateurs de l'État de New York ont ouvert une enquête** contre United Health Group accusé d'avoir utilisé un algorithme dans des hôpitaux qui donnait la priorité aux patients blancs sur les patients noirs (et ce n'est pas le seul exemple des problèmes que l'usage de l'IA en médecine pose, cf. notre article « **En médecine, l'IA est en plein essor, mais pas sa crédibilité** »). Plus de 100 millions de dollars ont été investis au cours des 6 derniers mois dans des entreprises explorant les questions éthiques liées à l'intelligence artificielle, estime **PitchBook**, un cabinet de recherche qui suit les activités financières des entreprises. **Software Alliance (@BSAnews)** a proposé récemment **un cadre détaillé (.pdf)** pour lutter contre les préjugés de l'IA en pointant le fait que certaines technologies automatisées nécessitaient une surveillance humaine régulière. Les grandes entreprises du numérique travaillent toutes sur ces sujets et déploient des outils dédiés.

Le problème, c'est qu'il n'y a pas de solution simple pour lutter contre les biais.

Pour nombre d'entreprises, construire une technologie équitable consiste surtout à ignorer les problèmes de discrimination qu'elle crée (ce qu'on appelle, d'une manière assez paradoxale, « *l'équité par inconscience* »). L'idée est simple : plus on apporte de données – et notamment des données les plus diverses possibles –, plus l'équité viendra. Pour la Software Alliance cependant, tout comme pour nombre de spécialistes du sujet, cet argument ne tient pas la route. Le problème n'est pas la masse de données, mais bien leur qualité, leur diversité bien sûr, qui n'est pas nécessairement assurée, mais aussi leur validité.

Le problème du problème, c'est que nous ne savons pas à quel point le problème des biais est grave, estime **Jack Clark (@jackclarksf)**, cofondateur d'**Anthropic**, membre du laboratoire **Human-Centered Artificial Intelligence de Stanford (StanfordHAI)** et auteur de l'**AI Index**

