



## L'IA une solution contre le sexisme ordinaire ? Une start-up Niçoise créée par une femme y travaille.

À l'occasion de la journée internationale des droits des femmes, Daphné Marnat dévoile sa solution pour outiller les concepteurs d'IA éthique.

L'intelligence artificielle est entraînée à partir de jeux de données qui comportent de nombreux biais possiblement discriminant, même dans les corpus les plus « propres » (issus de Wikipedia, reddit par exemple). Les biais se logent au cœur de nos cultures, même si nous souhaitons sincèrement les éviter, pas facile de les identifier, encore moins pour les machines qui apprennent sur nos productions de textes, d'images.



Les déclarations d'intention sur l'intelligence artificielle pleuvent ([rapport Villani](#)) mais peu d'outils technologiques sont produits pour permettre aux data scientists d'agir concrètement.

**Unbias** est né de la volonté de passer à l'action et de donner les moyens techniques au développement d'une IA éthique. Une anthropologue s'est associée avec un data scientist pour créer un modèle d'apprentissage capable de détecter des biais sexistes dans les phrases et de les remplacer par des mots ou des syntaxes moins discriminants.

Le modèle peut débiaiser un autre modèle ou « machine » comme un traducteur automatique, un agent conversationnel (chat-bot). [Or les entreprises s'équipent de plus en plus de chat-bots](#) pour

gérer leur relation client ou leurs ressources humaines. L'idée est d'agir au cœur de ce discours pour changer profondément nos éléments de langage. La machine peut nous aider à mettre en œuvre le changement que la plupart des grands acteurs institutionnels et économiques déclarent initier (cf. [Politique RSE](#) des entreprises), non sans mal.

## Les algorithmes sont-ils vraiment sexistes ?

La question mérite d'être posée tant les discriminations sexistes ou raciales liées aux algorithmes sont pointées du doigt (des IA discriminantes chez [Microsoft](#), [Google](#), [Amazon](#)).

Oui, une IA peut avoir des préjugés sexistes tout simplement parce qu'elle apprend en nous imitant. Les corpus d'apprentissage sont biaisés, comme ceux qui conçoivent les machines. Il y a peu de diversité parmi la communauté des data scientists tant du point de vue du genre, que de l'origine sociale et ethnique. Selon l'UNESCO, 22% des professionnels de l'IA à travers le monde sont des femmes.

Dans ce contexte, difficile pour eux d'avoir une réflexivité suffisante pour construire des modèles neutres du point de vue de la discrimination.

## Comment remédier à ce problème ?

**Combattre le sexisme dans l'IA, c'est diversifier les équipes et les données pour que la machine apprenne la diversité.**

Unbias est créée par une femme, la féminisation des métiers de l'IA est un premier pas, celles d'intégrer les sciences humaines pour produire des jeux d'apprentissage de qualité est en un autre.

**Unbias crée des jeux de données hautement qualifiés** qui sauront apporter un meilleur apprentissage à la machine du point de vue des biais. Aujourd'hui ceux qui créent les corpus d'apprentissage (Google, OpenIA) s'attachent à créer le plus « gros » corpus que le précédent, mais peu à leur qualité intrinsèque. Une logique (masculine ? :-)) qui a ses limites : le volume de données est impressionnant ([GP-3](#)) mais pas toujours pertinent en termes de [qualité d'apprentissage](#) (on ne fait qu'amplifier les biais dans les exemples répétés) et très coûteux à utiliser (le coût carbone du calcul des machines est loin d'être neutre- [source](#)).

**Limiter les biais, c'est savoir les repérer et les redresser.**

Comprendre le sexisme est la solution pour alerter de l'existence d'un biais ou l'éviter. Or le comprendre n'est pas si simple, c'est pourquoi Unbias invite les sciences humaines (linguistes, philologues, sociologues) dans la création de ses modèles et corpus d'apprentissage.

Nos recherches ont permis de repérer quatre types de sexisme dans les corpus des agents conversationnels notamment : hostile - celui qui insulte, dévalorise un genre, dominant - celui qui instaure un rapport de force, une supériorité entre les genres, essentialiste - celui qui associe un genre à une activité, un métier inutilement, suggestif - celui qui peut provoquer une réception discriminante par la sur-représentation d'un genre dans une phrase. Ce qui est un cas récurrent en langue française cf. la règle du masculin l'emporte sur le féminin.

Si le sexisme hostile est facile à identifier, bien qu'encore trop présent sur les jeux de données notamment issus du web, le sexisme ordinaire est quant à lui plus difficile à limiter tant il infuse nos cultures (cf. [rapport annuel Haut conseil de l'égalité sur le sexisme en France](#)). Or notre modèle a appris à le corriger au cœur des corpus qu'on lui donne à analyser.

En cas de maladie des enfants, les **mères parents** peuvent prendre un congé garde d'enfants.

Qui ~~est le chef de~~ **prend les décisions pour la** famille ?

The nurse entered the rooms and he gave me my medication.

Google Traduction : \_\_\_\_\_

L'**infirmière infirmier** est entrée dans la pièce et **elle il** m'a donné mes médicaments.

---

### À propos d'Unbias

**Unbias** est une start up co-fondée par une anthropologue et un data scientist, dirigée une femme.

**Unbias** est incubé à l'incubateur Telecom Paris Eurecom Entrepreneurs Sophia Antipolis



par

### Biographie Daphné Marnat

Daphné Marnat experte en sciences sociales est une entrepreneure qui depuis 10 ans accompagne les entreprises dans leur déploiement technologique du point de vue de leurs utilisateurs (études des données, des conditions d'appropriation de la technologie). Unbias est sa deuxième création d'entreprise. Elle milite pour une technologie au service de l'humain, pour la place des femmes dans l'entrepreneuriat et la technologie. Spécialiste de la donnée, elle défend une donnée qualitative non pas issue de la quantité, mais du sens, de la vraie intelligence. Daphné vit à Nice, développe Unbias à Sophia Antipolis.

**Site internet :** [unbias.fr](http://unbias.fr)

**Contact presse :** Marie-Philippine Méchet [presse@unbias.fr](mailto:presse@unbias.fr) – 06 60 65 20 65