

Protection contre la discrimination algorithmique

Septembre 2023

La prise de décision automatisée s'imisce de plus en plus dans notre vie quotidienne. Elle intervient, par exemple, lorsque des algorithmes évaluent nos candidatures, traitent nos déclarations d'impôts, établissent des diagnostics médicaux par l'intermédiaire d'Intelligences artificielles, prévoient le taux de criminalité, évaluent les risques de récidive des délinquant·e·s ou estiment les chances d'insertion des personnes réfugiées sur le marché du travail. Ces systèmes algorithmiques émettent des prévisions, dispensent des conseils ou prennent des décisions (appelés dans ce cas systèmes de décision automatisée, ADM). Ils peuvent également produire du contenu relatif à des décisions qui concernent des personnes et peuvent ainsi engendrer des discriminations.

En Suisse, la Constitution prévoit une interdiction de discrimination ([art. 8 al. 2 Cst.](#)). Celle-ci peut être définie comme une inégalité de traitement subie par une personne en raison de caractéristiques qui lui sont propres et qui ne peut être justifiée par un motif objectif. Actuellement, le cadre juridique helvétique n'accorde toutefois pas une protection suffisante ni efficace contre la discrimination, qui peut découler de l'utilisation de systèmes algorithmiques. Dans ce papier de position, nous présentons les enjeux liés à la discrimination algorithmique. Nous exposons également nos revendications visant à établir une meilleure protection contre la discrimination en Suisse.

Sommaire

/ Comment les algorithmes peuvent-ils discriminer ?	2
/ Pertinence sociale et exemples de discrimination algorithmique	2
/ Causes de la discrimination algorithmique	4
/ Caractéristiques de la discrimination algorithmique	6
/ Pourquoi le cadre juridique actuel ne protège-t-il pas suffisamment contre la discrimination algorithmique ?	7
/ Comment la protection contre la discrimination peut-elle être renforcée ?	8

Comment les algorithmes peuvent-ils discriminer ?

/ Pertinence sociale et exemples de discrimination algorithmique

Les algorithmes peuvent entraîner des discriminations dans l'**accès aux prestations sociales**. C'est par exemple le cas d'un [algorithme](#) mis en place aux Pays-Bas, qui classait les mères célibataires comme étant particulièrement susceptibles de commettre une fraude à l'aide sociale. Des études ont montré que certaines caractéristiques (par exemple parentalité, femme, jeune, ne parlant pas couramment le néerlandais, éprouvant des difficultés à trouver du travail) augmentaient le score de risque de fraude d'une personne. En 2019, il a été également révélé que l'administration fiscale néerlandaise avait employé un [algorithme fondé sur l'auto-apprentissage](#) pour établir des profils de risque afin de détecter des demandes d'allocations familiales potentiellement illicites. Sur la base d'une simple suspicion, laquelle reposait sur l'indicateur de risque élaboré par le système, l'administration fiscale a sanctionné de nombreuses familles. Des dizaines de milliers d'entre elles, souvent à faible revenu ou appartenant à des minorités ethniques, ont été sommées à tort de rembourser des années d'allocations familiales perçues. Ayant dû contracter des dettes écrasantes, beaucoup ont sombré dans la pauvreté, et plus de mille enfants ont été placés dans des familles d'accueil. Les autorités de protection des données ont finalement conclu que le traitement des données par ce système algorithmique était discriminatoire.

Dans le **domaine de l'éducation**, des algorithmes utilisés par exemple pour calculer les notes d'examen peuvent conduire un **accès inégal aux études supérieures**. Au Royaume-Uni, des examens, dont les résultats sont décisifs pour intégrer les universités britanniques, ont été annulés en raison de la pandémie de COVID-19. Pour pallier cette situation, le gouvernement a calculé les résultats finaux à [l'aide d'un algorithme](#). Le système s'appuyait entre autres sur l'historique des résultats enregistrés par les écoles : ainsi, de nombreux élèves ayant des résultats supérieurs à la moyenne issus d'écoles moins bonnes se sont vu attribuer des notes trop basses, tandis que les élèves moins doué·e·s mais provenant de meilleurs établissements (en particulier d'écoles privées) ont plutôt reçu de meilleures notes. L'algorithme a donc désavantagé les élèves issus de milieux à faibles revenus.

Dans le **domaine professionnel**, l'utilisation d'algorithmes peut entraîner un **accès inégal aux offres d'emploi**. Une [recherche](#) menée par AlgorithmWatch a révélé que sur Facebook, l'affichage ou non des postes vacants sur le fil d'actualité dépendait du genre des utilisateur·trice·s. L'annonce pour la conduite de camions est plutôt affichée aux hommes, celle pour la garde d'enfants plutôt aux femmes. Une autre [étude](#) a montré que l'algorithme employé par le réseau social discriminait les personnes en recherche d'emploi plus âgées, puisqu'il permettait de cibler des groupes d'âge spécifiques. Certaines offres d'emploi n'étaient pas visibles par les personnes plus âgées, qui étaient donc exclues du processus de candidature. Par ailleurs, les algorithmes peuvent engendrer des conséquences discriminatoires dans les procédures de recrutement. L'un des [exemples](#) les plus connus est un logiciel de candidature que Amazon

aurait développé, censé accroître l'efficacité des processus de recrutement. Durant la période d'essai, il a été constaté que l'algorithme discriminait les femmes en écartant les CV de celles-ci. Les tentatives cherchant à mettre fin à cette discrimination sont restées vaines.

Dans le **domaine médical**, les systèmes algorithmiques mènent parfois à **un accès inéquitable aux soins**. Aux États-Unis, des prestataires de soin de santé employaient un [algorithme](#) leur permettant d'identifier les patient·e·s qui pourraient bénéficier d'un suivi supplémentaire. Ce système discriminait les personnes racisées, qui se voyaient moins souvent proposé les soins supplémentaires pourtant nécessaires.

Dans le **domaine de la police prédictive (predictive policing) et de la justice**, la discrimination algorithmique peut également entraîner des répercussions lourdes sur les personnes concernées. Lors de la phase d'entraînement d'un [algorithme](#) de reconnaissance faciale, développé par les autorités judiciaires des États-Unis, les données fournies au système représentaient majoritairement des hommes à la carnation claire. Le système rencontrait donc des difficultés à reconnaître les visages à la peau plus foncée, ce qui engendrait de faux positifs. En outre, comme le démontre [l'exemple](#) du logiciel COMPAS, employé par le système pénitentiaire américain, l'utilisation d'algorithmes en police prédictive pour prédire le risque de récidive de condamné·e·s peut provoquer une surévaluation du risque présenté par un groupe de personnes particulier (*over-policing*). Une [étude](#) a révélé que les personnes noires étaient deux fois plus susceptibles d'être considérées comme potentiellement récidivistes, et ce à tort. En Suisse, un système similaire est également utilisé ; d'après une [étude](#), il est moins bien testé scientifiquement.

Dans le **domaine de la migration**, les systèmes algorithmiques peuvent également entraîner des effets discriminatoires. Aux Pays-Bas, un [algorithme de profilage](#) aurait été utilisé afin d'évaluer les risques associés aux demandeur·euse·s de visa à court terme dans le pays et dans l'espace Schengen. À l'aide de variables telles que la nationalité, le genre et l'âge, le système aurait établi le profil de risque de millions de personnes ayant déposé une demande de visa. Les individus classés « à haut risque » faisaient automatiquement l'objet d'un examen approfondi, un score pouvant démarrer des enquêtes et entraîner des retards de traitement considérables. Entre 2015 et 2020, le Royaume-Uni a également utilisé un [algorithme](#) similaire, qui permettait d'examiner les demandes de visa. Le logiciel devait identifier la nationalité des demandeur·euse·s : si la nationalité de ces personnes était évaluée comme « suspecte », celles-ci se voyaient attribuer un score de risque plus élevé et étaient en conséquence soumises à des contrôles plus stricts par les autorités.

Cette liste non exhaustive d'exemples met en évidence la diversité des domaines concernés. Bien que l'utilisation des systèmes algorithmiques augmente rapidement, **la société et les personnes concernées ignorent où, par qui et dans quels buts ces algorithmes sont employés. Elles ne peuvent que spéculer.** Mais ce qui est évident : plus les algorithmes sont utilisés, plus les cas de discriminations peuvent se multiplier.

/ Causes de la discrimination algorithmique

Les systèmes algorithmiques ne sont ni neutres ni objectifs. Les modèles de discrimination structurelle déjà existants dans une société s'y retrouveront. Mais comment la discrimination algorithmique peut-elle se produire exactement ?

Les algorithmes utilisent des **données** spécifiques qui ne représentent jamais entièrement la réalité. Celles-ci ne sont pas neutres : elles reflètent les relations et les rapports de forces existant dans la société. Elles se différencient par leur qualité, par exemple si elles sont actuelles ou obsolètes, et si elles sous-représentent ou surreprésentent un groupe d'individus. Les données, collectées dans le passé, ne peuvent que refléter celui-ci. Par exemple, si elles servent à produire des prédictions sur l'avenir, cela repose sur l'hypothèse que l'avenir ressemble au passé. Dans le cas d'algorithmes basés sur l'apprentissage autonome, les données utilisées influencent particulièrement les résultats produits. Imaginons par exemple qu'au sein d'une entreprise, moins de femmes occupent des positions de cadres que des hommes. Si un algorithme de recrutement est entraîné avec les données représentant les effectifs actuels, le système peut reproduire cette réalité. Un résultat similaire peut aussi être observé si les données ne mentionnent pas explicitement le genre des employé·e·s : l'algorithme peut trouver des corrélations avec les prénoms ou les loisirs pouvant être représentatifs du genre « féminin ». Toutefois, il est essentiel de rappeler que les humains jouent un rôle considérable dans l'organisation des données. Ce sont en effet eux qui préparent, nettoient et annotent les données, afin que l'algorithme autoapprenant puisse les utiliser. Ils influent également sur la collecte de données, c'est-à-dire celles qui sont retenues, quand et comment. Finalement, ils décident de l'allocation ou non des ressources qui réduisent le risque de discrimination, en garantissant que les ensembles de données sont représentatifs.

Un **modèle** algorithmique ne décrit qu'une partie de la réalité, qui est simplifiée pour être traduite en un simple problème mathématique. Les humains décident des critères et des aspects jugés pertinents pour un problème devant être résolu à l'aide du système. Au travers de ces choix, ils introduisent leurs hypothèses, leurs attentes et leurs préjugés dans le système, qui reproduit ceux-ci lorsqu'il génère des résultats. L'algorithme peut ainsi potentiellement causer des effets discriminatoires. Par exemple, lorsque les candidat·e·s qualifié·e·s pour un poste doivent être triés automatiquement selon leur CV, des humains doivent tout d'abord définir les critères qui permettent de déterminer la « qualification » des personnes qui postulent. Toutefois, ces paramètres peuvent se référer directement ou par substitution à des caractéristiques protégées contre la discrimination. C'était le cas du service de l'emploi autrichien, qui employait un [algorithme](#) calculant la probabilité de réinsertion des personnes au chômage sur le marché du travail. Ce système retirait des points lorsqu'un individu devait s'occuper de personnes à charge, mais seulement s'il s'agissait d'une femme. Lorsque les hommes avaient des enfants à charge, aucun point négatif n'était attribué. Étant donné que la garde des enfants est souvent confiée aux femmes, la parentalité des hommes influe moins sur leur situation professionnelle, réalité que l'algorithme a reproduite. Or, si cette pratique

caractéristique du passé est intégrée dans un modèle prenant des décisions sur le futur, l'algorithme n'entretient pas seulement les inégalités actuelles, mais provoque de nouvelles formes massives de discrimination.

Bien que la question des biais dans les modèles et dans les données d'entraînement soit primordiale, la discrimination peut également être due à d'autres causes. Même un système contrôlé pour qu'il ne provoque pas de discrimination et qu'il ne perpétue pas de préjugés peut être utilisé de manière discriminatoire. Par exemple, si un logiciel de reconnaissance faciale est développé afin de reconnaître toutes les personnes, mais qu'il est principalement utilisé pour contrôler de façon ciblée les personnes racisées, l'utilisation de ce système algorithmique peut entraîner des répercussions racistes.

Les **humains décident** également des **objectifs** pour lesquels le système algorithmique doit être utilisé. Ils identifient d'abord un problème devant être résolu, comme choisir une personne qualifiée parmi tous les candidat·e·s pour un poste. Ils peuvent, en second lieu, choisir d'employer un système de décision automatisée (ADM) pour mener le processus de recrutement, auquel cas ils déterminent les objectifs que l'algorithme doit suivre. Lorsque les humains précisent des intentions telles qu'une réduction des coûts et un accroissement du rendement, mais qu'ils ne mentionnent pas l'équité ou l'égalité de traitement, des effets discriminatoires peuvent se produire. Les algorithmes disposent pourtant du **potentiel** de réduire l'influence des préjugés et donc de **réduire la discrimination**. Cependant, pour qu'ils se comportent de telle manière, ce but doit être explicitement défini.

L'**utilisation** d'un un système algorithmique pour la prise de décision peut entraîner des conséquences discriminatoires. Si une administration ou une entreprise décide d'employer un système afin d'accomplir une tâche spécifique, elle le fait dans un contexte particulier qui conditionne l'utilisation du logiciel. Il s'agit par exemple des circonstances dans lesquelles les opérateur·trice·s voudraient ou pourraient contredire les recommandations proposées par l'algorithme. Les personnes qui se font conseiller par des algorithmes peuvent, au fil du temps, devenir plus dépendantes du système et moins critiques à son égard. Ainsi, un outil développé pour qu'il *assiste* à la prise de décision peut, dans la pratique, être utilisé comme un instrument *prenant* les décisions. En théorie, le système interagit bel et bien avec un humain (principe appelé « *human-in-the loop* »), mais cette personne cède ses responsabilités de supervision et son autorité de décision à l'algorithme. Il est donc capital que des processus adéquats permettant de réduire ces « [bais d'automatisation](#) » soient mis en place. De plus, les utilisateur·trice·s devraient suivre des formations favorisant la compréhension de ces systèmes, ainsi que la remise en question des recommandations et des prévisions produites par ceux-ci.

Finalement, les risques d'injustice et de discrimination sont particulièrement élevés lorsque les algorithmes sont utilisés dans des domaines où il existe déjà un **déséquilibre de pouvoir**, par exemple entre les personnes postulant pour un poste et l'entreprise, entre les salarié·e·s et les employeur·euse·s, entre les suspect·e·s et la police, entre les migrant·e·s et les autorités

chargées de surveiller les frontières, entre les personnes bénéficiant de l'aide sociale et l'administration, entre les élèves et le corps enseignant, ou encore entre les individus et les plateformes de réseaux sociaux. Dans de telles circonstances, le développement et l'utilisation d'un algorithme pourraient profiter particulièrement au groupe de personnes disposant déjà du pouvoir et des ressources. En revanche, l'autre groupe serait, à son insu, davantage exposé aux décisions et aux résultats produits par le système ; il pourrait ainsi subir les conséquences discriminatoires découlant potentiellement de l'utilisation de ce système.

/ Caractéristiques de la discrimination algorithmique

La **difficulté à identifier la discrimination algorithmique** est rencontrée tant par les personnes concernées que par les utilisateurs·trice·s. Elle constitue l'un des enjeux majeurs de la lutte contre cette forme de discrimination. Les personnes ne savent souvent pas qu'un système est utilisé ou comment il fonctionne. La discrimination peut être **systemique**, c'est-à-dire qu'elle est ancrée dans le système et touche ainsi un grand nombre de personnes (effet d'échelle), qui peinent à la reconnaître et à se considérer comme victimes.

Les **boucles de rétroaction** peuvent renforcer les « biais » déjà existants. Lorsqu'un système ADM utilisé en police prédictive propose d'intensifier la présence de patrouilles dans certains quartiers, la police y constatera plus de délits qu'ailleurs. Pourtant, un nombre similaire d'infractions commises dans d'autres zones pourraient rester indétectées. Ces résultats consolident à leur tour les prédictions du système.

La discrimination algorithmique peut découler des **proxys**, c'est-à-dire des variables qui sont mesurées à la place d'autres, par exemple lorsque ces dernières sont trop difficiles à relever ou qu'elles ne peuvent pas être observées. Les variables de substitution permettent de déterminer certaines caractéristiques spécifiques (et parfois protégées contre la discrimination), sans que celles-ci aient été demandées ou recensées. Par exemple, la donnée « 30 ans d'expérience professionnelle » indique que la personne en question doit être âgée de 45 ans au minimum. Le code postal permet de formuler des hypothèses sur le statut socioéconomique ou sur l'origine d'un individu. En conséquence, les variables proxys peuvent exercer une influence déterminante sur la décision prise par un système ADM.

La discrimination algorithmique peut résulter de **l'association de différentes caractéristiques protégées contre la discrimination**, par exemple lorsque des personnes sont discriminées à la fois en raison de leur genre et de leur origine. Dans ce cas, il est possible que l'effet discriminatoire se renforce et qu'il devienne difficile de déterminer quel facteur joue quel rôle. De plus, les systèmes fondés sur l'apprentissage automatique identifient eux-mêmes des schémas et peuvent donc déterminer **de nouvelles caractéristiques engendrant des conséquences discriminatoires**. Un système reconnaissant des visages sur la base de vidéos pourrait attribuer différents résultats selon des mouvements des yeux ou de la gestuelle dans ses résultats, ce qui pourrait potentiellement désavantager les personnes en situation de handicap. En outre, l'obtention des résultats concernant les systèmes ou le fonctionnement de ces

derniers, ainsi que l'identification des effets théoriquement discriminatoires, se révèlent souvent complexes (problème de la « **boîte noire** »). Par **manque d'éléments objectifs et concrets**, aucun débat ne peut avoir lieu.

Pourquoi le cadre juridique actuel ne protège-t-il pas suffisamment contre la discrimination algorithmique ?

La Constitution suisse prévoit un principe d'égalité, selon lequel « **tous les êtres humains sont égaux devant la loi** », ainsi qu'une interdiction de discrimination ([art. 8 al. 1 et al. 2 Cst.](#)). Cette dernière peut être définie comme une inégalité de traitement fondée sur des caractéristiques essentielles et qui ne peut être justifiée par un motif objectif. Il y a inégalité de traitement lorsqu'une personne subit un traitement défavorable en comparaison avec celui d'autres individus ou groupes d'individus, bien qu'elle se trouve dans une situation identique ou comparable. L'article 8, alinéa 2 de la Constitution mentionne des caractéristiques biologiques (telles que la « race », le sexe, l'âge, la déficience corporelle, mentale ou psychique), culturelles ou sociales (l'origine, la langue, la situation sociale, le mode de vie, les convictions religieuses, philosophiques ou politiques). Cette énumération n'est volontairement pas exhaustive, car de nouveaux groupes subissant de l'exclusion sont continuellement reconnus et de nouveaux mécanismes de discrimination peuvent voir le jour. Les groupes sociaux stigmatisés sont généralement protégés contre la discrimination. L'article 8 de la Constitution contient deux autres alinéas, qui définissent une égalité entre l'homme et la femme ([art. 8 al. 3 Cst.](#)) et prévoient des mesures en vue d'éliminer les inégalités qui frappent les personnes handicapées ([art. 8 al. 4 Cst.](#)).

Lacune 1 : En Suisse, l'interdiction de discrimination ne concerne en principe que les **acteurs étatiques** et ne s'applique pas aux particuliers. L'État est tenu indirectement d'empêcher la discrimination entre personnes privées, une obligation découlant de l'interdiction de discrimination prévue par la Constitution ([art. 8 al. 2 Cst.](#) en lien avec [art. 35 al. 3 Cst.](#)) Toutefois, il n'existe aucune loi générale interdisant spécifiquement la discrimination par des acteurs privés alors que **l'utilisation d'algorithmes et de systèmes ADM progresse rapidement dans l'ensemble de la société**. Il faut un cadre juridique qui donne des moyens de lutter contre la discrimination liée aux systèmes algorithmiques utilisés par des acteurs privés.

Lacune 2 : Actuellement, la protection contre la discrimination n'est pas suffisante pour faire face aux **caractéristiques spécifiques de la discrimination algorithmique** (notamment : effet d'échelle, boucles de rétroaction, variables proxys).

Lacune 3 : La discrimination algorithmique pose de nouveaux enjeux pour l'**application** de l'interdiction de discrimination. D'une part, il est **difficile d'identifier les personnes concernées**. Celles-ci ne sont généralement pas conscientes de la discrimination ou ne peuvent l'être, car elles ne savent pas comment le système fonctionne, les données de comparaison font défaut ou la discrimination se révèle systémique. Le cadre juridique existant repose sur une

identification précise de la personne discriminée ; il n'est donc que partiellement adapté à cette situation. D'autre part, les personnes concernées ne disposent souvent pas du temps ou des moyens nécessaires pour faire recours, et des obstacles pratiques ou formels leur barrent la route.

Comment la protection contre la discrimination peut-elle être renforcée ?

Afin que la protection contre la discrimination s'adapte aux enjeux liés à la discrimination algorithmique, nous exposons ci-après nos revendications :

- / **Le domaine d'application de l'interdiction de discrimination devrait être élargi, afin que les personnes concernées par la discrimination algorithmique soient efficacement protégées.** Lorsque des systèmes algorithmiques entraînent des conséquences discriminatoires, celles-ci peuvent potentiellement concerner un grand nombre de personnes. En effet, ces systèmes produisent toutes leurs prédictions ou leurs décisions selon le même modèle. Des risques sont courus non seulement au niveau individuel mais aussi au niveau structurel. Les conséquences pour l'ensemble de la société peuvent donc être importantes, comme l'illustrent les exemples mentionnés ci-dessus.
- / **L'interdiction de discrimination devrait également s'étendre aux caractéristiques relevant de la discrimination algorithmique, telles que les variables proxys.** L'utilisation de systèmes algorithmiques peut causer des inégalités de traitement injustifiées, qui ne correspondent pour autant à aucune des catégories identifiables et protégées par l'interdiction de discrimination. Une différence de traitement fondée sur des catégories prétendument neutres peut néanmoins avoir un effet discriminatoire, étant donné que la catégorie neutre (comme le code postal) peut servir de proxy pour des catégories protégées par le droit de la discrimination (comme le statut social).
- / **Une interdiction générale de la discrimination pour les privés devrait être introduite.** À quelques exceptions près, il n'existe en Suisse pas d'interdiction de discrimination s'appliquant aux acteurs privés. Cette lacune est importante, puisque ce sont en grande partie les acteurs privés qui mettent au point et qui utilisent les systèmes algorithmiques. Ceux-ci exercent des influences notables sur la vie d'autres personnes, et les retombées qu'ils entraînent peuvent être observées dans divers contextes, tels l'attribution de logements, le calcul des primes d'assurances, la vérification de la solvabilité ou le recrutement d'employé·e·s. Une loi-cadre sur l'égalité de traitement pourrait permettre de protéger contre la discrimination par des privés.

- / La protection contre la discrimination doit également être garantie lorsque celle-ci survient de manière systémique et que les personnes concernées ne peuvent être clairement identifiées. Il en va de même lorsqu'un système automatisé produit, par hasard ou selon sa propre logique, une prédiction ou une décision qui engendre des conséquences discriminatoires.
- / L'interdiction de discrimination devrait également inclure la discrimination fondée sur l'association de plusieurs caractéristiques protégées (aussi appelée discrimination intersectionnelle), afin qu'elle protège également contre les effets discriminatoires que celle-ci peut engendrer.
- / Les entreprises et les autorités qui mettent en place, fournissent ou utilisent des systèmes algorithmiques devraient être contraintes de conduire régulièrement des analyses d'impact, ainsi que des audits standardisés et indépendants. Ceux-ci permettraient d'identifier les risques de discrimination et de violation des droits fondamentaux qui peuvent découler de l'utilisation de ces systèmes (plus précisément des bases de données, de l'usage envisagé, du modèle ou du contexte d'application), ainsi que d'adopter des mesures adéquates. Dans le cas de systèmes employés par l'administration publique, les résultats des audits et des évaluations d'impacts devraient être publiés dans des registres accessibles au public. Cette responsabilité permettrait d'instaurer et de garantir une transparence et une supervision (p. ex. par la communauté scientifique ou par la société civile) publiques. Les acteurs privés devraient être tenus de garantir la transparence des systèmes employés dans des domaines particulièrement délicats, ainsi que celle des résultats des audits et des évaluations d'impact.
- / De meilleurs moyens pour améliorer l'application de l'interdiction de discrimination sont nécessaires. Il faut d'abord venir à bout des obstacles pratiques et formels pouvant décourager des personnes à faire recours : le fardeau de la preuve, entre autres, doit être réduit pour les personnes concernées. En effet, la discrimination algorithmique, camouflée par le manque de transparence du système et par d'autres facteurs, n'est souvent pas directement identifiable. Le fardeau de la preuve devrait donc incomber en premier lieu à la partie accusée, chargée de démontrer qu'il ne s'agit pas de discrimination et qu'elle a fait tout son possible pour empêcher une inégalité de traitement, notamment par le biais du devoir de diligence, d'analyses d'impact et d'audits. En outre, une procédure de recours collectif devrait être intégrée à la législation et être accessible aux acteurs privés et publics ; elle permettrait de réduire les obstacles et les coûts pour les personnes concernées, ainsi que de démontrer que la discrimination algorithmique touche un grand nombre de personnes. Dans ce cas de figure, un droit de l'action collective devrait également être proposé aux organisations de la société civile. Compte tenu de la problématique de la « boîte noire », les actions collectives peuvent aussi jouer un rôle central pour traiter les

cas de discrimination algorithmique.

- / **La recherche et le débat fondé sur des preuves autour de la discrimination algorithmique et de son impact social doivent être encouragés.** Les autorités doivent acquérir des compétences et être sensibilisées aux thèmes et aux enjeux de cette forme de discrimination. Il est central que les problèmes liés à la discrimination algorithmique ne soient pas considérés comme purement techniques, mais comme nécessitant une approche interdisciplinaire. Finalement, l'implication des personnes concernées, également dans l'élaboration du cadre juridique, doit être garantie.

Contact / AlgorithmWatch CH / Fabrikstrasse 17 / CH-8005 Zurich / +41 78 411 94 54