Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössische Kommission für Frauenfragen EKF
Commission fédérale pour les questions féminines CFQF
Commissione federale per le questioni femminili CFQF

Eidgenössische Kommission gegen Rassismus EKR
Commission fédérale contre le racisme CFR
Commissione federale contro il razzismo CFR

Avis de droit

Protection contre la discrimination algorithmique

Impressum

Titre
Protection contre la discrimination algorithmique avec une attention particulière accordée à la discrimination raciale et la discrimination fondée sur le genre

Éditrices
Commission fédérale pour les questions féminines CFQF et Commission fédérale contre le racisme CFR

Autrices et auteur
Prof. Dre Nadja Braun Binder, MBA, professeure de droit public à l'Université de Bâle et co-responsable du Centre de recherche sur la numérisation dans l'État et l'administration e-PIAF
Prof. Dr Florent Thouvenin, professeur de droit de l'information et de la communication à l'Université de Zurich, président du comité de pilotage du Center for Information Technology, Society, and Law ITSL
Dre Stephanie Volz, directrice scientifique de l'ITSL, chargée de cours à l'Université de Zurich
MLaw Liliane Obrecht, assistante scientifique, doctorante à la Faculté de droit de l'Université de Bâle

Diffusion de la publication en ligne
Commission fédérale pour les questions féminines CFQF
Schwarztorstrasse 51
3003 Berne
ekf@apg.admin.ch
www.comfem.ch
058 462 92 75

Commission fédérale contre le racisme CFR
Inselgasse 1
3003 Berne
ekr-cfr@gs-edi.admin.ch
www.ekr.admin.ch
058 464 12 93

Édition imprimée
L'avis de droit peut être commandé gratuitement sous :
www.ekr.admin.ch > Publications > Études

Traduction
Service linguistique du DFI

Cet avis de droit a été réalisé sur mandat de la CFQF et de la CFR.

Berne, novembre 2025

Universität Basel
Juristische Fakultät

Universität Zürich[UZH]

e-PIAF
electronic Public Institutions and Administrations Research Forum

CENTER FOR INFORMATION TECHNOLOGY SOCIETY AND LAW — ITSL

Protection contre la discrimination algorithmique

avec une attention particulière accordée à la discrimination raciale
et à la discrimination fondée sur le genre

Avis de droit élaboré sur mandat de la Commission fédérale pour les questions féminines CFQF et de la Commission fédérale contre le racisme CFR

Prof. Dre Nadja Braun Binder

Prof. Dr Florent Thouvenin

Dre Stephanie Volz

MLaw Liliane Obrecht

Table des matières

Table des matières I
Liste des abréviations III
1. Introduction 1
2. Discrimination algorithmique 2
2.1. Principes 2
2.1.1. Algorithmes classiques vs algorithmes fondés sur l'IA 2
2.1.2. Champs d'application 3
2.1.3. Insignifiance du niveau d'automatisation 4
2.2. Normes internationales 4
2.2.1. Droit international public 4
2.2.2. Convention du Conseil de l'Europe sur l'IA 5
2.2.3. Excursus : Règlement européen sur l'IA (règlement UE-IA) 6
2.3. Sur la notion de discrimination 7
2.3.1. La notion de discrimination dans la Constitution 7
2.3.2. Approche intersectionnelle et discrimination multiple 10
2.3.3. Le cas particulier de la discrimination raciale et de celle fondée sur le genre 11
2.4. Discrimination directe et indirecte dans les algorithmes fondés sur l'IA 11
3. Interdiction générale de la discrimination 12
3.1. Loi sur l'égalité (LEg) 12
3.1.1. Principes inscrits dans la Constitution 12
3.1.2. Dispositions pertinentes de la LEg 13
3.1.3. Applicabilité aux systèmes algorithmiques 14
3.2. Loi sur l'égalité pour les handicapés (LHand) 15
3.2.1. Principes inscrits dans la Constitution 15
3.2.2. Dispositions légales pertinentes 15
3.2.3. Applicabilité aux systèmes algorithmiques 16
3.3. Loi fédérale sur la protection des données (LPD) 16
3.3.1. Dispositions légales pertinentes 16
3.3.2. Excursus : Traitement de données personnelles à des fins de détection de situations de discrimination 19
3.3.3. Efficacité 19
4. Dispositions particulières en droit public 20
4.1. Garanties de procédure, en particulier l'obligation de motiver 20
4.2. Autres interdictions de la discrimination 21

4.3. Efficacité des dispositions de droit public ... 22
5. Interdiction de la discrimination en droit privé ... 23
5.1. Protection de la personnalité en droit civil (art. 28 CC) ... 23
5.2. Rapports commerciaux dans les limites des principes généraux de la loi, des bonnes mœurs et de l'ordre public (art. 19 et 20 CO) ... 24
5.3. Violation de la bonne foi (art. 2 CC) ... 25
5.4. Droit du travail ... 25
5.5. Autres normes ... 27
5.6. Efficacité des dispositions de droit privé ... 27
6. Interdiction de la discrimination dans le droit pénal (art. 261bis CP) ... 27
7. Conclusion intermédiaire ... 28
8. Mesures envisageables ... 29
8.1. Dans le cadre du droit en vigueur ... 29
8.1.1. Application des normes existantes ... 29
8.1.2. Travaux de recherche supplémentaires ... 30
8.2. Mesures législatives ... 31
8.2.1. Décisions partiellement automatisées ... 31
8.2.2. Loi générale sur l'égalité de traitement ... 32
8.2.3. Nouveau principe relatif au traitement des données ... 34
8.3. Mesures particulières de droit public ... 35
8.3.1. Création de registres publics ... 35
8.3.2. Utilisation limitée à des systèmes offrant une traçabilité ... 36
8.3.3. Recommandations à l'intention des autorités administratives qui utilisent des systèmes d'IA 36
8.3.4. Interdiction des décisions entièrement automatisées dans les domaines sensibles ... 37
9. Conclusion ... 37
Remerciements ... 37
Références bibliographiques ... 39
Bases légales / normes ... 46
Documents de référence ... 49

Liste des abréviations

AELE	Association européenne de libre échange
AIPD	Analyse d'impact relative à la protection des données
AJP	Aktuelle Juristische Praxis (Zurich)
AöR	Archiv des öffentlichen Rechts (Tübingen)
ASDPO	Association suisse du droit public de l'organisation
ATF	Arrêt du Tribunal fédéral
BSK	Basler Kommentar / Commentaire bâlois
CB	Compliance Berater (Sarrebruck)
CMLR	Common Market Law Review (Alphen aan den Rijn, Pays-Bas)
digma	Zeitschrift für Datenrecht und Informationssicherheit (Zurich)
DRdA	Das Recht der Arbeit (Vienne)
DSB	Datenschutz-Berater (Francfort)
EuZ	Zeitschrift für Europarecht (Zurich)
ex/ante	Zeitschrift der juristischen Nachwuchsforschung (Zurich)
IA	Intelligence artificielle
JAR	Jahrbuch des Schweizerischen Arbeitsrechts (Berne)
JPART	Journal of Public Administration Research and Theory (Oxford)
Jusletter	Jusletter (Berne)
LS	Loseblattsammlung du Canton de Zurich
NJW	Neue Juristische Wochenschrift (Munich)
NVwZ	Neue Zeitschrift für Verwaltungsrecht (Munich)
OFK	Orell Füssli Kommentar / Commentaire Orell Füssli
PFPDT	Préposé fédéral à la protection des données et à la transparence
Plädoyer	Magazin für Recht und Politik (Zurich)
QFLR	Quid? Fribourg Law Review (Fribourg)

Questions au féminin	Revue spécialisée de la Commission fédérale pour les questions féminines (CFQF) (Berne)
RDi	Recht Digital – Die Zeitschrift zur Digitalisierung in der Rechts- und Wirtschaftspraxis (Munich)
recht	Zeitschrift für juristische Weiterbildung und Praxis
RS	Recueil systématique du droit fédéral
RSV	Recueil systématique de la législation vaudoise
SAR	Recueil systématique de la législation argovienne
SG	Recueil systématique de la législation du canton Canton de Bâle-Ville
SGK	St. Galler Kommentar / Commentaire st-gallois
SHK	Stämpflis Handkommentar / Commentaire Stämpfli
SJZ-SRJ	Schweizerische Juristen-Zeitung / Revue suisse de jurisprudence (Zurich)
sui generis	Revue juridique en accès libre (Zurich)
SZW-RSDA	Schweizerische Zeitschrift für Wirtschafts- und Finanzmarktrecht / Revue suisse de droit des affaires et du marché financier (Zurich)
TAF	Tribunal administratif fédéral
TA-SWISS	Fondation d'évaluation des choix technologiques
TF	Tribunal fédéral
UE	Union européenne
WP	White Paper
ZBJV	Zeitschrift des Bernischen Juristenvereins (Berne)
ZBl	Schweizerisches Zentralblatt für Staats- und Verwaltungsrecht (Zurich)
zsis	Zentrum für Schweizerisches und Internationales Steuerrecht (Zurich)
ZSR-RDS	Zeitschrift für Schweizerisches Recht / Revue de droit suisse (Bâle)

1. Introduction

Les systèmes algorithmiques sont utilisés dans de nombreux domaines, tant par les institutions publiques que les acteurs privés, par exemple dans l'administration fiscale, le domaine des assurances sociales, l'enseignement, le recrutement de personnel, le secteur financier ou encore sur des plateformes numériques. Ils servent en particulier à l'établissement de prévisions ainsi qu'à la génération de recommandations et de décisions entièrement ou partiellement automatisées.

Si ces algorithmes offrent des opportunités intéressantes, ils posent aussi un certain nombre de défis, dans la mesure où ils peuvent engendrer des discriminations. En Suisse, l'étude de la discrimination algorithmique n'en est qu'à ses débuts. Les premières recherches ayant mis en lumière ce phénomène datent de 2019[1]. Cantonnée jusqu'à présent aux milieux scientifiques, cette problématique n'a pas encore fait l'objet d'un débat politique, juridique ou sociétal plus large ni même donné lieu à des mesures législatives concrètes. Compte tenu de l'importance croissante des systèmes algorithmiques et des tentatives de régulation au niveau international – en particulier au sein de l'Union européenne (UE) et du Conseil de l'Europe –, la question de la protection contre la discrimination algorithmique devient prioritaire en Suisse également.

Le fonctionnement des algorithmes, en particulier ceux fondés sur l'intelligence artificielle (IA)[2], peut engendrer des cas de discrimination. En effet, il est généralement impossible de déterminer les critères que ces systèmes prennent en compte pour générer leurs résultats. Ce risque est accru par le fait que ces systèmes sont utilisés par un grand nombre d'autorités ou d'entreprises, qui s'en servent par exemple pour évaluer des candidatures. Si ces systèmes sont discriminatoires, ce sont donc tout de suite un grand nombre de personnes qui seront touchées. À l'inverse, le recours aux systèmes d'IA présente aussi un potentiel important pour détecter ou réduire les discriminations. En utilisant des données synthétiques, il est en effet possible de tester les systèmes pour vérifier s'ils prennent des décisions discriminatoires. Par exemple, il est possible de soumettre des dossiers fictifs à un système de recrutement pour voir s'il recommande exclusivement des hommes pour un poste donné. Au vu des effets d'échelle, l'implémentation de systèmes d'IA non discriminatoires aurait un fort impact sociétal. Compte tenu de ces réflexions, il devient essentiel de créer des dispositions légales claires visant à réduire le plus possible les risques de ces systèmes et à maximiser leur potentiel.

Le présent avis de droit analyse le cadre juridique suisse en matière de protection contre la discrimination algorithmique – en particulier contre la discrimination raciale et celle fondée sur le genre[3] – et examine si ce cadre permet de répondre aux défis spécifiques liés à l'utilisation des systèmes algorithmiques. L'un des principaux objectifs de l'étude est d'identifier les éventuelles mesures à prendre, notamment au niveau législatif, et de proposer des pistes concrètes pour améliorer la protection contre la discrimination algorithmique.

---

1 Voir WILDHABER / LOHMANN / KASPER, p. 459 ss

2 Concernant ces notions, voir ci-après chap. 2.1.1

3 Les deux commissions, qui ont mandaté cet avis de droit, privilégient l'usage du terme « genre » (en allemand « Geschlecht »). Elles sont conscientes que les dispositions juridiques utilisent majoritairement le terme « sexe », que nous employons uniquement lorsque le contexte légal ou biologique l'exige.

2. Discrimination algorithmique

2.1. Principes

2.1.1. Algorithmes classiques vs algorithmes fondés sur l'IA

Les algorithmes classiques sont élaborés par des humains. Concrètement, il s'agit d'une série de règles et d'instructions clairement définies qui, pour une même entrée, produiront toujours le même résultat. Ces algorithmes fonctionnent donc selon une logique déterministe et leurs résultats sont en règle générale totalement traçables. On parle aussi d'algorithmes déterministes ou fondés sur des règles.

Les algorithmes fondés sur l'IA reposent quant à eux sur des processus d'apprentissage automatique (*machine learning*). Autrement dit, ils ne suivent pas des règles fixes prédéfinies par une personne physique, mais apprennent de manière autonome à partir de grandes quantités de données (données d'entraînement). Durant cette phase d'apprentissage autonome, les algorithmes identifient les corrélations et les schémas récurrents présents dans les données d'entraînement, sans consigne explicite. Après cette phase d'entraînement, ils sont capables de prendre des décisions ou de faire des prévisions à partir de données qu'ils n'ont jamais vues auparavant. Le fonctionnement de ces algorithmes repose en règle générale sur des modèles mathématiques complexes qui traitent et pondèrent les données saisies par les utilisateurs et utilisatrices en plusieurs étapes. Si ces algorithmes sont très performants, leur fonctionnement est généralement difficile, voire impossible à comprendre pour les humains (on parle de boîte noire)[4].

Dès le départ, il a été observé que ces algorithmes fondés sur l'IA peuvent discriminer systématiquement certains groupes de personnes, par exemple en raison de leur genre, de leur origine ethnique ou de leur orientation sexuelle. Les causes de cette discrimination sont multiples[5]. Elles résident souvent dans les données d'entraînement, qui reflètent les structures dominantes et les discriminations déjà présentes dans la société, et qui peuvent conduire les systèmes d'IA à les reproduire à notre insu (biais historiques)[6]. Un tel biais peut se produire dans la reconnaissance d'images, par exemple lorsque les systèmes d'IA étiquettent les femmes se trouvant dans une cuisine comme des « femmes au foyer » alors que les hommes se trouvant dans le même contexte sont plutôt catalogués comme « cuisiniers » ou « chefs »[7]. Les traductions réalisées par des systèmes d'IA peuvent aussi présenter ce type de biais historiques. Ainsi, le terme anglais *nurse* est souvent traduit par « infirmière », même si le genre de la personne n'est pas spécifié dans le texte source[8]. Ces biais reflètent les rôles stéréotypés du passé, qui se retrouvent dans les données d'entraînement. Un autre type de biais apparaît lorsque certains groupes sont sous-représentés dans les données d'entraînement (biais de représentation)[9]. Prenons un système automatisé d'évaluation des candidatures. Ce système est entraîné pour identifier les caractéristiques des candidatures retenues les années précédentes. Si la majorité des personnes embauchées par le passé étaient des hommes, le système risque d'intégrer ce schéma récurrent et de le reproduire, ce qui

---

4 Voir à ce sujet MARTINI, Blackbox, p. 28 ss ; ORWAT, p. 97

5 Voir BRAUN BINDER et al., Studie KI, p. 40 ss ; THOUVENIN / VOLZ / HEITZ / WEINER, ch. 73 ss

6 Voir BRAUN BINDER et al., Studie KI, p. 40 ; THOUVENIN / VOLZ / HEITZ / WEINER, ch. 78

7 Voir TOM SIMONITE, Machines Taught by Photos Learn a Sexist View of Women, Wired, 21 août 2017 (https://www.wired.com/story/machines-taught-by-photos-learn-a-sexist-view-of-women/?utm_source=chatgpt.com)

8 STANOVSKY / SMITH / ZETTLEMOYER, p. 1679 ss ; voir également THOUVENIN / VOLZ / WEINER / HEITZ, ch. 78

9 Voir BRAUN BINDER et al., Studie KI, p. 40 s. ; THOUVENIN / VOLZ / HEITZ / WEINER, ch. 79

désavantagerait systématiquement les femmes[10]. Dans cet exemple, le problème vient de la base de données, qui ne permet pas au système d'IA de prendre des décisions équitables et fiables pour tous les groupes de la population. Il peut en résulter des prévisions ou des décisions biaisées qui renforcent les discriminations existantes ou créent de nouvelles inégalités.

En plus de ces différents biais dans les données d'entraînement, la discrimination algorithmique peut aussi résulter de la conception même des modèles d'IA et des critères pris en compte pour générer les résultats. Ainsi, la décision de considérer certains critères plutôt que d'autres peut conduire à une inégalité de traitement, même si les données sous-jacentes sont apparemment neutres. Prenons l'exemple d'un algorithme d'évaluation de la solvabilité qui fonde ses décisions, entre autres critères, sur le lieu de résidence de la personne qui demande un crédit. Bien que les données ne présentent aucun biais évident, le système pourrait désavantager systématiquement les personnes résidant dans certaines localités en raison de la corrélation avec le critère – en soi neutre – du lieu de résidence.

Même si les algorithmes classiques (fondés sur des règles) peuvent aussi fournir des résultats discriminatoires, les considérations ci-après se concentrent sur les discriminations émanant des systèmes fondés sur l'IA, qui posent des problèmes particuliers en matière de protection contre la discrimination algorithmique.

2.1.2. Champs d'application

L'IA est une technologie transversale, qui se prête à de nombreuses utilisations potentielles. L'attention du grand public est aujourd'hui rivée sur les systèmes d'IA « génératifs », qui sont capables de créer du texte, des images ou de la musique en très peu de temps[11]. Au-delà de ces systèmes, les algorithmes fondés sur l'IA sont aussi utilisés pour des applications plus spécifiques, parfois très sensibles, telles que l'évaluation automatisée des candidatures[12], les décisions de taxation fiscale[13] ou la reconnaissance d'images biométriques[14] ou médicales[15].

Les questionnements juridiques soulevés par ces systèmes varient en fonction du contexte d'utilisation. Les risques et les défis dépendent fortement de la finalité des systèmes, de la base de données qui a servi à leur entraînement et de leur domaine d'application. La présente étude n'a pas pour ambition de passer en revue tous les types d'application envisageables et les risques correspondants. Elle s'intéresse plutôt aux problématiques fondamentales qu'il est possible de régler par des mesures légales transversales. Il s'agit notamment d'aspects qui reviennent dans différents domaines et qu'il est possible de résoudre par l'adoption de règles générales. Bien que des règles sectorielles puissent également présenter un intérêt, leur examen ne sera pas entrepris dans le cadre de cette étude.

---

10 Voir AlgorithmWatch CH, Automatisch aussortiert? Wenn KI-Système in Bewerbungsprozessen diskriminieren, 9 octobre 2025, en particulier en ce qui concerne un logiciel de candidature d'Amazon (https://algorithmwatch.ch/de/diskriminierung-durch-ki-in-der-personalauswahl/)

11 P. ex., ChatGPT d'Open AI, Gemini de Google ou Perplexity AI

12 Voir DZIDA / GROH, p. 1917 ss ; TINHOFER, p. 174 s.

13 BRAUN BINDER / OBRECHT, Steuererklärungen, ch. 7 ss

14 BRAUN BINDER / KUNZ / OBRECHT, p. 53 ss ; SIMMLER / CANOVA, p. 107 ss

15 BATACHE, ch. 87 ss

2.1.3. Insignifiance du niveau d'automatisation

Les discriminations algorithmiques peuvent survenir non seulement lorsque les systèmes d'IA prennent des décisions sans intervention humaine (systèmes ou décisions entièrement automatisés), mais aussi lorsque les décisions sont prises par un individu en interaction avec une machine (systèmes ou décisions partiellement automatisés)[16]. Les systèmes partiellement automatisés permettent d'appuyer ou de contrôler la prise de décision humaine[17], par exemple lorsque des modèles de notation sont utilisés pour évaluer la solvabilité d'un individu[18]. Si la décision finale revient formellement à l'être humain, elle est de fait fortement influencée par la proposition du système algorithmique. Cette pratique est particulièrement problématique lorsque la personne qui prend la décision ne procède pas à une vérification critique de la proposition de l'algorithme ou lorsqu'elle lui accorde une confiance excessive (biais d'automatisation)[19].

2.2. Normes internationales

En Suisse, la protection contre la discrimination repose à la fois sur la législation nationale et sur les obligations découlant du droit international. Le présent chapitre présente les normes internationales qui peuvent entrer en ligne de compte pour la protection contre la discrimination algorithmique.

2.2.1. Droit international public

En Suisse, le droit international public et le droit national forment un ordre juridique unique (monisme). Cela signifie que les normes de droit international public sont contraignantes pour la Suisse. On distingue deux types de dispositions : celles qui, en raison de leur formulation concrète, sont directement applicables (*self-executing*) et celles qui ne le sont pas directement (*non self-executing*)[20].

La Convention européenne des droits de l'homme (CEDH) et le Pacte international relatif aux droits civils et politiques (Pacte II de l'ONU) contiennent notamment des dispositions directement applicables relatives à la protection contre la discrimination. Ces deux traités internationaux interdisent explicitement la discrimination et présentent une liste (non exhaustive) de caractéristiques qui doivent être protégées de la discrimination (p. ex. le genre ou la langue). En particulier, l'art. 26 du Pacte II de l'ONU précise que l'obligation de garantir l'égalité de traitement incombe non seulement aux autorités chargées d'appliquer le droit, mais aussi au législateur. La Suisse a également ratifié la Convention internationale sur l'élimination de toutes les formes de discrimination raciale (CERD) et la Convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes (CEDEF). La question de savoir si des droits peuvent en être directement déduits reste floue et doit être évaluée au cas par cas[21].

---

16 Sur les différents degrés d'automatisation, voir GORDON / LUTZ, p. 55 ; WEDER, p. 241 s.

17 Voir DZIDA / GROH, p. 1917 ss ; TINHOFER, p. 174 s.

18 VOIR ŠKORJANC, p. 70 ss

19 Sur le biais d'automatisation, voir ALON-BARKAT / BUSUIOC, p. 155 ; RUSCHEMEIER, Automation Bias, p. 2

20 Voir à ce sujet RHINOW / SCHEFER / UEBERSAX, ch. 3612 ss

21 Dans son message relatif à la CEDAW, le Conseil fédéral a estimé que les dispositions de la convention (à quelques exceptions près) n'étaient pas directement applicables (FF 1995 869, ici 893 ss). Dans la littérature, certains auteurs considèrent en revanche que la convention est directement applicable (voir FANKHAUSER-FEITKNECHT, p. 25 ; KÄGI-DIENER, Bedeutung, p. 44 s. ; concernant l'art. 4 CEDAW, voir KÄGI-DIENER, Impulse, p. 1462 s.).

Quoi qu'il en soit, en vertu du principe de l'interprétation conforme au droit international[22], ces conventions entrent en ligne de compte dans l'interprétation du droit interne. À noter toutefois que ni la CEDH, ni le Pacte II de l'ONU, ni la CERD, ni la CEDEF ne contiennent de dispositions concrètes relatives à la protection contre la discrimination algorithmique.

2.2.2. Convention du Conseil de l'Europe sur l'IA

La Convention-cadre sur l'intelligence artificielle, les droits de l'homme, la démocratie et l'État de droit du Conseil de l'Europe (ci-après Convention CE-IA) est le premier instrument de droit international contraignant de régulation de l'IA. Elle vise à garantir un développement et une utilisation des systèmes d'IA pleinement compatibles avec les valeurs fondamentales du Conseil de l'Europe[23] – en particulier les droits humains, la participation démocratique et les principes de l'État de droit[24]. Dans ce texte, on entend par système d'IA un « système automatisé qui, pour des objectifs explicites ou implicites, déduit, à partir d'entrées reçues, comment générer des résultats en sortie tels que des prévisions, des contenus, des recommandations ou des décisions qui peuvent influer sur des environnements physiques ou virtuels. Différents systèmes d'intelligence artificielle présentent des degrés variables d'autonomie et d'adaptabilité après déploiement »[25]. La Convention CE-IA oblige notamment les États parties à mettre en place des mesures appropriées, entre autres aux niveaux législatif et administratif, afin de réduire au maximum les risques des systèmes d'IA pour les droits humains, la démocratie et l'État de droit, tout en exploitant de manière responsable le potentiel de cette technologie[26].

La Convention CE-IA s'applique aux pouvoirs publics, mais aussi aux acteurs privés qui agissent pour leur compte (p. ex. fournisseurs de prestations informatiques ou fournisseurs de systèmes d'IA utilisés dans le domaine des poursuites pénales)[27]. Elle oblige en outre les États parties à répondre des risques et des conséquences découlant des activités menées par des acteurs privés[28], dans le respect de l'objet et du but de la convention (protection des droits humains, de la démocratie et de l'État de droit). Les États jouissent d'une grande marge de manœuvre dans la mise en œuvre.

Le Conseil de l'Europe a adopté la convention en mai 2024. En février 2025, le Conseil fédéral a annoncé son intention de la ratifier[29], ce qu'il a fait le 27 mars 2025[30].

a. Garantir la protection contre la discrimination dans l'utilisation de l'IA

Aux termes de l'art. 10 de la Convention CE-IA, chaque État partie doit adopter ou maintenir « des mesures visant à garantir le respect de l'égalité, y compris l'égalité de genre, et l'interdiction de la

---

22 Voir notamment RHINOW / SCHEFER / UEBERSAX, ch. 562 ss

23 Voir le portail du Conseil de l'Europe : Droits humains – Démocratie – État de droit (https://www.coe.int/fr/web/portal/)

24 Art. 1, par. 1, Convention CE-IA

25 Art. 2 Convention CE-IA

26 Voir le portail du Conseil de l'Europe, La Convention-cadre sur l'intelligence artificielle (https://www.coe.int/fr/web/artificial-intelligence/la-convention-cadre-sur-l-intelligence-artificielle)

27 Art. 3, par. 1, let. a, Convention CE-IA

28 Art. 3, par. 1, let. b, Convention CE-IA

29 Voir le communiqué de presse du 12 février 2025 « Réglementation de l'IA : le Conseil fédéral veut ratifier la Convention du Conseil de l'Europe » (https://www.news.admin.ch/fr/nsb?id=104110), ainsi que OFCOM, État des lieux IA 2025, p. 21 ss

30 Voir le communiqué de presse du 26 mars 2025 « La Suisse signe la convention du Conseil de l'Europe sur l'IA » (https://www.news.admin.ch/fr/nsb?id=104646)

discrimination dans les activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle, comme le prévoit le droit international et interne applicable »[31]. En outre, chaque État partie doit s'engager à adopter ou maintenir « des mesures visant à supprimer les inégalités dans les activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle, afin d'obtenir des résultats impartiaux, justes et équitables, dans le cadre des obligations nationales et internationales qui lui incombent en matière de droits de l'homme »[32]. Le rapport explicatif relatif à la Convention CE-IA précise que ces mesures peuvent être de nature réglementaire ou technique et que les États peuvent envisager des solutions en matière de gouvernance ou d'autres approches[33].

b. Application de la Convention CE-IA en Suisse

Lorsque la Suisse ratifie une convention du Conseil de l'Europe, elle s'engage à mettre en œuvre les obligations qui y sont énoncées. Concrètement, cela signifie qu'elle s'engage à respecter les normes de la convention et à adapter son ordre juridique interne afin de le rendre conforme à la convention. Elle peut le faire en créant de nouvelles lois, en adaptant des réglementations existantes ou en prenant des mesures administratives.

En février 2025, le Conseil fédéral s'est prononcé en faveur d'une application restreinte de la Convention CE-IA[34]. En d'autres termes, les obligations incomberont principalement à l'État, alors que le secteur privé sera largement épargné. En outre, lorsque des modifications légales seront nécessaires, elles devront être aussi sectorielles que possible. Une réglementation générale et intersectorielle ne sera mise en place que dans les domaines juridiques centraux relevant des droits fondamentaux (comme la protection des données).

La Suisse doit donc adapter son cadre légal afin de concrétiser l'interdiction de la discrimination algorithmique prévue à l'art. 10 de la Convention CE-IA. La mise en œuvre se limite toutefois aux obligations qui existent déjà dans le droit national. La Convention CE-IA ne contient aucune disposition exigeant d'étendre la protection contre la discrimination.

2.2.3. Excursus : Règlement européen sur l'IA (règlement UE-IA)

L'UE a adopté un règlement global sur l'intelligence artificielle. Si ce règlement n'est pas contraignant pour la Suisse, il revêt toutefois une grande importance pour les entreprises qui commercialisent des systèmes d'IA sur le marché européen et qui doivent donc respecter ces dispositions.

Le règlement UE-IA vise à prévenir le risque de discrimination algorithmique en fixant différents niveaux de risque. Ainsi, les systèmes d'IA à haut risque doivent satisfaire à des exigences particulières[35], notamment disposer d'un système de gestion des risques[36]. À l'instar de l'analyse d'impact relative à la protection des données personnelles (AIPD) prévue par la loi fédérale sur la protection des données (LPD), le règlement UE-IA prévoit concrètement l'obligation d'identifier,

---

31 Art. 10, par. 1, Convention CE-IA

32 Art. 10, par. 2, Convention CE-IA

33 Rapport explicatif Convention CE-IA 2024, ch. 75

34 Voir le communiqué de presse du Conseil fédéral du 12 février 2025 (https://www.news.admin.ch/fr/nsb?id=104110)

35 Pour la classification des systèmes d'IA comme *systèmes à haut risque*, voir l'art. 6 du règlement UE-IA ; pour plus de détails, voir KI-VO-RUSCHEMEIER, art. 6, N 1 ss ; MARTINI, Hochrisiko-KI-Systeme, § 4, N 14 ss, 40 ss

36 Art. 9 règlement UE-IA ; voir KI-VO-BRAUN BINDER / EGLI, art. 9, N 2, 15 ss ; SPINDLER, § 5, N 5 ss

d'analyser et d'évaluer les risques de ces systèmes[37]. Sur la base des risques identifiés, il faut ensuite prendre des mesures appropriées pour les réduire autant que possible et empêcher la survenance de dommages ou d'autres désavantages[38]. Contrairement à l'AIPD, qui est réalisée avant le traitement des données, la gestion des risques prévue par le règlement UE-IA couvre l'ensemble du cycle de vie des systèmes d'IA et se poursuit même après leur mise sur le marché et pendant leur utilisation[39].

Les systèmes algorithmiques fondés sur des données doivent être entraînés, validés et testés à partir de grands jeux de données[40]. Or ces jeux de données peuvent présenter des biais susceptibles d'entraîner des effets discriminatoires[41]. Pour réduire ce risque, l'art. 10 du règlement UE-IA régit la « gouvernance des données »[42], qui prévoit notamment un examen permettant de repérer d'éventuels biais susceptibles d'engendrer une discrimination (au sens du droit de l'UE)[43].

Le règlement UE-IA répond à la problématique de la traçabilité des systèmes algorithmiques en fixant des exigences en matière de transparence[44]. Pour les systèmes d'IA à haut risque, ces exigences concernent l'enregistrement des événements (art. 12), la transparence et la fourniture d'informations aux déployeurs (art. 13), le contrôle humain (art. 14) et l'exactitude (art. 15). Cette dernière exigence peut être interprétée comme une disposition visant à éviter les résultats discriminatoires[45].

2.3. Sur la notion de discrimination

2.3.1. La notion de discrimination dans la Constitution

L'art. 8, al. 2, Cst. interdit toute forme de discrimination et énumère une série de caractéristiques sur la base desquelles nul ne peut être discriminé (p. ex. le genre, l'origine, la race[46], l'âge, la langue ou la situation sociale)[47]. À noter que cette liste de caractéristiques protégées est indicative et non exhaustive[48]. Les interdictions de discrimination prévues dans les constitutions cantonales[49] n'ont de portée propre que si elles offrent une protection plus étendue que l'art. 8, al. 2, Cst.

L'interdiction de discrimination constitue l'essence même du droit fondamental à l'égalité[50]. Le principe constitutionnel de l'égalité de traitement (art. 8, al. 1) impose de traiter de manière identique ce qui est

---

37 Art. 9, par. 2, let. a et b, règlement UE-IA ; voir KI-VO-BRAUN BINDER / EGLI, art. 9, N 20, 22 ss ; SPINDLER, § 5, N 7

38 Art. 9, par. 2, let. d, règlement UE-IA ; voir KI-VO-BRAUN BINDER / EGLI, art. 9 N 2, 28 ss ; SPINDLER, § 5 N 9 ss. Pour la forme concrète des mesures, voir art. 9, par. 5, règlement UE-IA et KI-VO-BRAUN BINDER / EGLI, art. 9 N 33 ss

39 KI-VO-Braun BINDER / EGLI, art. 9, N 18 s.

40 KI-VO-Braun BINDER / EGLI, art. 10, N 1, 10 ss

41 KI-VO-Braun BINDER / EGLI, art. 10, N 15

42 KI-VO-Braun BINDER / EGLI, art. 10, N 15 ; SPINDLER, § 5, N 21 ss ; LEGNER, p. 428

43 Art. 10, par. 2, let. f, règlement UE-IA ; voir SPINDLER, § 5, N 25 ; KI-VO-BRAUN BINDER / EGLI, art. 10, N 25 ss

44 Concernant la problématique de la transparence dans l'utilisation des systèmes algorithmiques, voir KUMKAR, § 6, N 3

45 KUMKAR, § 6, N 12 ss ; LEGNER, p. 429

46 Ce terme est utilisé dans le présent avis de droit uniquement comme citation de la Cst. La « race » est un concept dépourvu de fondement scientifique qui repose sur l'attribution de caractéristiques biologiques et comportementales à un groupe de personnes (voir Mahon / Graf / Steffanini, ch. 4 ss).

47 ATF 129 I 217, consid. 2.1 ; 126 II 377, consid. 6a ; BSK BV-WALDMANN, art. 8, N 61 ; RHINOW / SCHEFER / UEBERSAX, ch. 1913 s. ; pour plus de détails, voir DIGGELMANN et al., p. 212 ss

48 KIENER / KÄLIN / WYTTENBACH, ch. 1822

49 P. ex. art. 11, al. 2, KV-ZH ; § 8, al. 2, KV-BS ; art. 10, al. 2, Cst-VD

50 BSK BV-WALDMANN, art. 8, N 13

semblable et de manière différente ce qui est dissemblable[51]. Il ne garantit donc pas une égalité purement formelle – une stricte uniformité de traitement entre les individus –, mais une égalité matérielle qui tient compte des circonstances concrètes et des besoins des personnes[52]. Le principe général d'égalité n'exclut ainsi pas toute différenciation de traitement[53]. En ce sens, il présente deux dimensions : un principe d'égalité et un principe de différenciation[54].

Le principe d'égalité de traitement remplit une double fonction : il protège les droits individuels et garantit l'État de droit[55]. Cette protection s'étend à toutes les personnes, physiques comme morales[56]. Le principe général d'égalité se retrouve à divers niveaux de l'ordre juridique ainsi qu'à tous les échelons de l'État fédéral. À l'instar de l'interdiction de discrimination, les garanties d'égalité inscrites dans les constitutions cantonales[57] n'ont de portée propre que lorsqu'elles assurent une protection plus étendue que celle garantie par la Constitution fédérale[58].

Le champ d'application de l'interdiction de discrimination est plus restreint que celui de l'égalité de traitement. En principe, seules les personnes physiques peuvent invoquer l'interdiction de discrimination (art. 8, al. 2, Cst.)[59]. Dans certaines circonstances particulières, il est toutefois possible qu'une personne morale qui poursuit un but non lucratif, par exemple religieux ou politique, se voie reconnaître la jouissance des droits fondamentaux en lien avec la poursuite de son but[60].

a. Discrimination directe

Dans sa jurisprudence relative à l'art. 8, al. 2, Cst., le Tribunal fédéral estime que la discrimination directe constitue « une forme qualifiée d'inégalité de traitement de personnes dans des situations comparables, dans la mesure où elle produit sur un être humain un effet dommageable, qui doit être considéré comme un avilissement ou une exclusion, car elle se rapporte à un critère de distinction qui concerne une part essentielle de l'identité de la personne intéressée ou à laquelle il lui est difficilement possible de renoncer »[61]. La différence de traitement repose sur l'appartenance à un groupe qui, historiquement ou dans le contexte social actuel, tend à être marginalisé ou considéré comme inférieur[62]. La discrimination directe se caractérise essentiellement par l'existence d'un lien de causalité entre la caractéristique protégée et le désavantage subi[63].

---

51 Voir la jurisprudence constante du TF : ATF 146 II 56, consid. 9.1 ; 138 I 321, consid. 3.2 ; 123 I 1, consid. 9 ; 125 I 173, consid. 6

52 SGK BV-SCHWEIZER / FANKHAUSER, art. 8, N 7

53 KIENER / KÄLIN / WYTTENBACH, ch. 1746

54 KIENER / KÄLIN/ WYTTENBACH, ch. 1740 ss

55 KIENER / KÄLIN / WYTTENBACH, ch. 1736 ss

56 ATF 142 I 195, consid. 6.1 ; 93 I 1, consid. 1 ; SGK BV-SCHWEIZER / FANKHAUSER, art. 8, N 17

57 P. ex. § 8, al. 1, KV-BS ; art. 10, al. 1, Cst-VD

58 C'est le cas p. ex. de l'art. 14, al. 2, KV-ZH, qui prévoit l'égalité d'accès aux établissements d'enseignement.

59 ATF 139 I 242, consid. 5.3 ; KIENER / KÄLIN / WYTTENBACH, ch. 1813 ss

60 KIENER / KÄLIN / WYTTENBACH, ch. 1813 ss ; SGK BV-SCHWEIZER / FANKHAUSER, art. 8, N 19

61 ATF 147 I 89, consid. 2.1 ; ATF 143 I 129, consid. 2.3.1. Voir aussi ATF 138 I 217, consid. 3.3.3 ; 132 I 49, consid. 8.1 ; 131 V 9, consid. 3.4.3 ; 130 I 352, consid. 6.1.2 ; 129 III 276, consid. 3.1 ; 129 I 217, consid. 2.1 (concernant les inégalités de traitement portant atteinte à la personnalité des salariés) ; voir aussi ATF 126 II 377, consid. 6a ; arrêt du TAF C-8867/2010 du 6.11.2013, consid. 6.2.1

62 ATF 139 I 292, consid. 8.2.1 ; voir aussi premier arrêt sur la question : ATF 126 II 377. Pour des exemples concrets, voir DIGGELMANN et al., p. 222

63 BELSER / WALDMANN, p. 293 s. ; KIENER / KÄLIN / WYTTENBACH, ch. 1853

L'interdiction de discrimination ne saurait être assimilée à une interdiction absolue de traiter différemment certaines personnes en fonction d'une caractéristique donnée[64]. Concrètement, il n'y a pas discrimination lorsque l'inégalité de traitement repose sur des motifs qualifiés, c'est-à-dire lorsqu'elle poursuit un objectif légitime et qu'elle est adéquate, nécessaire et proportionnée pour les personnes concernées, au regard du but visé[65].

b. Discrimination indirecte

Une discrimination peut aussi advenir de manière indirecte[66]. Pour le Tribunal fédéral, c'est le cas lorsqu'une « réglementation, qui ne désavantage pas directement un groupe déterminé, défavorise tout particulièrement, par ses effets et sans justification objective, les personnes appartenant à ce groupe »[67]. Le Tribunal fédéral ajoute que « l'atteinte doit revêtir une importance significative, le principe de l'interdiction de la discrimination indirecte ne pouvant servir qu'à corriger les effets négatifs les plus flagrants d'une réglementation étatique »[68].

La discrimination indirecte s'apprécie en fonction des effets réels qu'elle produit sur un groupe de personnes. Contrairement à la discrimination directe, elle ne présente pas de lien de causalité immédiat entre une caractéristique protégée et le désavantage subi. Autrement dit, une discrimination indirecte peut exister même si une personne ne possède pas la caractéristique protégée, l'élément décisif étant que, dans l'ensemble, les personnes présentant cette caractéristique subissent un désavantage disproportionné[69]. Alors que l'interdiction de la discrimination directe vise à protéger les individus, la protection contre la discrimination indirecte concerne des règles en apparence neutres qui désavantagent structurellement certains groupes.

Il convient de noter que la notion de discrimination indirecte est parfois interprétée différemment (de manière plus large) dans d'autres systèmes juridiques ou en dehors du domaine du droit. Dans ces interprétations, la discrimination fondée sur des variables corrélées à une caractéristique protégée (discrimination par procuration) est souvent qualifiée de discrimination indirecte[70]. Dans la législation suisse, cependant, la notion de discrimination indirecte n'est mentionnée qu'à l'art. 3, al. 1, de la loi sur l'égalité (LEg). Par ailleurs, le Tribunal fédéral n'a jusqu'à présent admis la discrimination indirecte que dans des cas de discrimination fondée sur le genre, en particulier dans des affaires traitant de la rémunération des professions dites féminines[71]. Quant à la question de savoir si les femmes sont désavantagées parce que les « professions féminines » sont moins bien rémunérées, le Tribunal fédéral a précisé qu'une profession est généralement considérée comme typiquement féminine lorsque la proportion de femmes y est nettement supérieure à 70 %[72]. La discrimination fondée sur le genre constitue toutefois un cas particulier, car sa reconnaissance juridique repose directement sur l'art. 8,

---

64 SGK BV-SCHWEIZER / FANKHAUSER, art. 8, N 48 ; BSK BV-WALDMANN, art. 8, N 62

65 ATF 138 I 265, consid. 4.2.1 ; 126 V 70, consid. 4c ; BV Komm-BIAGGINI, art. 8, N 22 ; BSK BV-WALDMANN, art. 8, N 8

66 BELSER / WALDMANN, p. 293 s. ; KIENER / KÄLIN / WYTTENBACH, ch. 1859 ss ; voir aussi ARIOLI, p. 1328 ss

67 BV Komm-BIAGGINI, art. 8, N 19 ; WALDMANN, Diskriminierungsverbot, p. 338 ss avec références à la jurisprudence

68 ATF 138 I 205, consid. 5.5. Voir aussi ATF 139 I 169, consid. 7.2.1 ; 138 I 205, consid. 5.5 ; 138 I 305, consid. 3.3; BSK BV-WALDMANN, art. 8, N 63

69 Pour des exemples concrets, voir DIGGELMANN et al., p. 222 s.

70 Voir notamment BUCHHOLTZ / SCHEFFEL-KAIN, p. 615 s.

71 ATF 125 II 530, consid. 2.b ; 125 II 385, consid. 3.b

72 ATF 125 II 530, consid. 2.b ; 125 II 385, consid. 3.b

al. 3, Cst., qui pose une égalité de fait entre hommes et femmes. De plus, la notion de discrimination utilisée dans la LEg est large, et la discrimination indirecte y est expressément mentionnée à l'art. 3, al. 1, comme nous venons de le dire[73]. Il n'est donc pas certain que la discrimination indirecte puisse s'appliquer aussi à d'autres caractéristiques protégées.

2.3.2. Approche intersectionnelle et discrimination multiple

Dans la doctrine, la discrimination intersectionnelle est considérée comme une forme de discrimination multiple[74]. Ce concept souligne que la discrimination ne se limite souvent pas à un seul critère, mais peut se renforcer par la combinaison de plusieurs facteurs[75]. On distingue parfois deux types spécifiques de discrimination multiple[76] : la discrimination additive et la discrimination amplificatrice. La discrimination additive désigne une situation dans laquelle une personne est discriminée pour plusieurs motifs distincts, par exemple en raison de son genre et de son appartenance religieuse[77]. La discrimination amplificatrice se produit lorsque l'interaction de plusieurs caractéristiques protégées (au moins deux) accentue le désavantage subi[78] : les critères interagissent et se renforcent mutuellement, entraînant un cumul des motifs de discrimination[79].

Outre ces deux formes, la notion de discrimination multiple s'applique également lorsque la discrimination ne repose pas sur un seul critère identifiable, mais sur l'interaction complexe de plusieurs caractéristiques déterminantes pour l'acte discriminatoire[80]. Dans ce cas, la discrimination ne résulte pas d'une simple addition ou cumulation de motifs, mais de l'interaction d'un ensemble complexe de critères qui se renforcent mutuellement et engendrent de nouvelles formes d'inégalités[81]. En droit suisse, la discrimination multiple est abordée de manière marginale, tant dans la législation (art. 5, al. 1, de la loi sur l'égalité pour les handicapés, LHand) que dans la pratique. Cette notion apparaît toutefois à plusieurs reprises dans la jurisprudence, surtout dans des affaires concernant le port du foulard[82]. Dans la doctrine, aucun auteur et aucune autrice ne considère la discrimination multiple comme une discrimination au sens strict des dispositions légales, et dans la pratique, l'existence d'une discrimination est évaluée séparément pour chaque caractéristique. Certaines voix appellent à une approche plus globale : il s'agirait de se concentrer moins sur les caractéristiques individuelles à l'origine des discriminations et davantage sur les effets des normes sociales et de la stigmatisation, ainsi que sur les différences réelles entre individus, qui sont actuellement insuffisamment prises en compte[83]. Il n'existe toutefois pas encore de propositions concrètes sur la manière de traiter juridiquement la discrimination intersectionnelle ou multiple.

---

73 À ce propos, voir ch. 3.1 ci-après
74 Pour plus de détails, voir NAGUIB, p. 235 ss
75 NAGUIB, p. 235 ss
76 NAGUIB, p. 235 ss
77 KLEBER, p. 34 s. ; NAGUIB, p. 235 ; ODERMATT, p. 7 ; PÄRLI, p. 163
78 NAGUIB, p. 236
79 NAGUIB, p. 236 ; ODERMATT, p. 7 ; PÄRLI, p. 163
80 NAGUIB, p. 236 ; PÄRLI, p. 163
81 NAGUIB, p. 235 ss ; ODERMATT, p. 8 ; SGK BV-SCHWEIZER / FANKHAUSER, art. 8, N 52
82 ATF 123 I 296 ; 132 I 167 ; voir aussi ODERMATT, p. 10 ; PÄRLI, p. 164
83 VOIR NAGUIB / PÄRLI, p. 478, qui mentionne d'autres références ; PÄRLI, p. 162

2.3.3. Le cas particulier de la discrimination raciale et de celle fondée sur le genre

Dans la présente étude, la discrimination fondée sur le genre désigne toute inégalité de traitement subie par une personne en raison de son genre, réel ou supposé.

La notion de discrimination raciale est plus difficile à cerner, car elle ne repose sur aucun concept clairement défini sur le plan biologique ou juridique[84]. Par ailleurs, le terme « race » est problématique à plus d'un titre : chargé d'une forte connotation historique, il repose sur des catégories construites qui sont dépourvues de fondement scientifique et socialement discriminatoires[85]. De manière générale, la discrimination raciale désigne toute forme de désavantage subi en raison de caractéristiques réelles ou supposées telles que l'appartenance à un groupe ethnique, la couleur de peau, l'origine ou l'appartenance religieuse[86].

2.4. Discrimination directe et indirecte dans les algorithmes fondés sur l'IA

La distinction entre discrimination directe et indirecte n'a pas été conçue à l'origine pour décrire les phénomènes de discrimination algorithmique. Si cette distinction conserve toute sa pertinence dans le contexte des systèmes d'IA, les termes « direct » et « indirect » peuvent toutefois prêter à confusion. En effet, les systèmes d'IA ne se fondent généralement pas sur une caractéristique protégée explicite, mais sur un ensemble de variables qui peuvent, dans certaines circonstances, être corrélées à une caractéristique protégée. On parle alors de variables corrélées. De plus, il est souvent impossible de déterminer quelles caractéristiques ont effectivement influencé la décision (problématique de la boîte noire). La discrimination fondée sur de telles variables corrélées (discrimination par procuration/*proxy discrimination*) est souvent, à tort, assimilée à une discrimination indirecte. Cette confusion s'explique par le fait que les systèmes d'IA n'utilisent pas directement une caractéristique protégée, mais aboutissent à une discrimination par l'intermédiaire d'une ou de plusieurs corrélations. Cette interprétation, déjà perceptible dans les débats, est toutefois erronée, car elle méconnaît la différence juridiquement décisive entre discrimination directe et discrimination indirecte. Dans la discrimination directe, les personnes sont désavantagées en raison d'une caractéristique protégée : il existe donc un lien de causalité entre cette caractéristique et le désavantage subi. En revanche, la discrimination indirecte se produit lorsqu'une règle en apparence neutre affecte de manière disproportionnée un groupe de personnes partageant une caractéristique protégée, sans que cette caractéristique soit la cause directe du désavantage. Le fait qu'un système algorithmique se fonde directement sur une caractéristique protégée ou sur une variable (ou une combinaison de variables) corrélée à une caractéristique protégée ne change rien à l'existence de ce lien de causalité. Sinon, il serait aisé de contourner une discrimination directe juridiquement prohibée simplement en recourant à des variables corrélées.

Dans le cas des systèmes d'IA, il y a donc discrimination directe même lorsqu'un système se fonde sur des variables corrélées à une caractéristique protégée. Comme dans la discrimination directe classique, il existe un lien de causalité entre la caractéristique protégée et le désavantage : les personnes concernées sont défavorisées en raison d'un critère protégé. Dans le cas d'une discrimination par des systèmes d'IA, ce lien est toutefois moins évident, car le système ne se fonde pas directement sur une caractéristique protégée, mais sur une autre variable ou une série de variables qui lui sont corrélées. La discrimination

---

84 Voir MAHON / GRAF / STEFFANINI, notamment ch. 6, 52
85 MAHON / GRAF / STEFFANINI, ch. 4
86 MAHON / GRAF / STEFFANINI, ch. 52 ; MÜLLER / SCHEFER, p. 720 ss

par procuration constitue donc une forme particulière, plus complexe, de discrimination directe. Comme il n'est généralement pas possible d'identifier les variables effectivement utilisées par le système, il faut se fonder sur le résultat pour déterminer s'il existe une discrimination. Ainsi, si un système d'IA ne propose que des hommes pour un poste donné, il s'agit d'une discrimination directe, même s'il n'est pas possible de prouver que le système a pris en compte le genre comme variable explicite. Comme il n'y a pas de lien de causalité directement observable entre la caractéristique protégée et le désavantage – mais seulement entre les variables corrélées et ce désavantage –, il faut généralement un nombre suffisant de cas pour établir l'existence d'une discrimination directe par un système algorithmique. Cela n'a toutefois rien de fondamentalement nouveau. Déjà dans le cas de la discrimination humaine, il est souvent impossible de démontrer qu'une personne s'est fondée directement sur une caractéristique protégée ; par exemple, prouver qu'une entreprise refuse d'engager des femmes parce qu'elles sont des femmes reste souvent difficile. De manière générale, la discrimination ne peut être identifiée et prouvée qu'à partir d'un échantillon significatif de situations comparables. La preuve d'une discrimination directe par un système d'IA ne nécessite toutefois pas nécessairement des cas réels. Des tests empiriques peuvent être menés, par exemple à l'aide de données synthétiques, afin de vérifier si le système discrimine certaines catégories de personnes (p. ex. en proposant uniquement des hommes pour un poste donné). Ainsi, l'utilisation de systèmes algorithmiques peut aussi représenter une chance : elle permet d'identifier les discriminations et d'ajuster les modèles pour éviter qu'elles ne se reproduisent à l'avenir.

3. Interdiction générale de la discrimination

Le présent chapitre passe en revue les normes de l'ordre juridique suisse interdisant toute forme de discrimination qui s'appliquent aux particuliers et aux autorités publiques, au premier rang desquelles les dispositions de la loi sur l'égalité (LEg) et de la loi sur l'égalité pour les handicapés (LHand). Décrites respectivement aux ch. 3.1 et 3.2, elles concrétisent le respect des principes d'égalité de traitement et d'interdiction de la discrimination inscrits dans la Constitution fédérale. Nous nous intéressons ensuite aux dispositions de la loi fédérale sur la protection des données (LPD) sur lesquelles il serait possible de se fonder pour prévenir la discrimination algorithmique lors du traitement des données personnelles (ch. 3.3.).

3.1. Loi sur l'égalité (LEg)

3.1.1. Principes inscrits dans la Constitution

La Constitution fédérale garantit l'égalité de droit et de fait entre les femmes et les hommes (art. 8, al. 3, Cst., *lex specialis*)[87], une définition de la discrimination étant donnée à l'art. 8, al. 2, Cst.[88] Les trois phrases qui composent l'art. 8, al. 3, Cst. ont chacune un contenu normatif propre[89] : la première pose le principe de l'interdiction de la discrimination comme un droit exigible ; la deuxième, à caractère programmatique, pourvoit à l'égalité de droit entre les femmes et les hommes, sans qu'un droit puisse être invoqué individuellement[90] ; la troisième renvoie à un droit que les salariés et les salariées peuvent

---

87 RHINOW / SCHEFER / UEBERSAX, ch. 1939 ss ; BSK BV-WALDMANN, art. 8, N 106

88 BSK BV-Uhlmann, art. 8, N 97

89 KIENER / KÄLIN / WYTTENBACH, ch. 1893 ss ; BSK BV-WALDMANN, art. 8, N 96

90 BELSER / WALDMANN, p. 288. Contrairement à l'al. 1, l'al. 3 consacre une égalité de droit absolue. Une inégalité de traitement entre les femmes et les hommes ne pourrait ainsi être justifiée par des raisons « objectives » (voir HÄFELIN / HALLER / KELLER / THURNHERR, ch. 779)

faire valoir[91] vis-à-vis de leur employeur, à savoir un droit opposable à un tiers, à effet direct [92] (droit à un salaire égal pour un travail de valeur égale).

La plupart des constitutions cantonales incluent elles aussi une disposition explicite destinée à garantir l'égalité salariale entre les femmes et les hommes[93]. Contrairement à l'interdiction générale de la discrimination, le principe de l'égalité salariale produit des effets directs dans les relations entre particuliers[94].

3.1.2. Dispositions pertinentes de la LEg

La LEg a pour but de promouvoir, dans les faits, l'égalité entre les femmes et les hommes (art. 1 LEg). Elle interdit la discrimination (inégalité de traitement) fondée sur le genre dans les rapports de travail (art. 3, al. 1, LEg). Les notions de « discrimination » et d'« inégalité de traitement » sont considérées comme synonymes[95]. La LEg distingue la discrimination directe de la discrimination indirecte. Selon la jurisprudence et la doctrine, la discrimination directe compte deux composantes : la discrimination directe ouverte, lorsque l'inégalité de traitement se fonde explicitement sur le critère du genre[96], la discrimination directe cachée, lorsque l'inégalité de traitement se fonde sur un critère dit « suspect » ne pouvant être rempli que par l'un des deux sexes, p. ex. la grossesse ou la maternité[97]. La discrimination directe (ouverte comme cachée) ne peut être justifiée que si le genre représente au sens de la LEg un critère essentiel pour l'exercice de l'activité professionnelle (mesures de protection des femmes enceintes, p. ex.)[98].

La LEg interdit par ailleurs explicitement la discrimination indirecte (art. 3, al. 1, LEg), mais sans toutefois en donner une définition. Le message concernant la LEg renvoie à cet égard à la jurisprudence de la Cour de justice des communautés européennes (CJCE)[99]. L'utilisation d'un critère d'apparence neutre est ainsi jugée discriminatoire si ce critère peut être rempli beaucoup plus facilement ou nettement plus souvent par les membres d'un sexe que de l'autre, sans que cela puisse être justifié objectivement (octroi d'avantages uniquement aux travailleurs et travailleuses à temps plein, p. ex.)[100]. Pour être considérée comme telle, l'inégalité de traitement doit être significative (il ne doit pas s'agir de cas

---

91 ATF 142 II 49, consid. 5.1 ; 131 I 105, consid. 3.3. Les plaintes se fondent généralement sur la LEg, qui pose le principe de l'égalité salariale.

92 ATF 125 II 3683, consid. 3 ; DIGGELMANN et al., p. 221. L'obligation de proposer un salaire égal pour un travail d'une valeur égale s'impose déjà aux employeurs publics en vertu de l'art. 8, al. 1, Cst. (voir ATF 129 I 161, consid. 3.2, et ATF 103 Ia 517, consid. 1b). L'art. 8, al. 1, Cst. ne crée pas de droit subjectif à l'égalité salariale entre les genres, il consacre le principe de l'interdiction de la discrimination salariale (ATF 141 II 411, consid. 6.1.1).

93 Art. 11, al. 3 Cst-ZH ; art. 10, als. 4 Cst-VD

94 BV Komm.-BIAGGINI, art. 8, N 8

95 Art. 3, al. 1, LEg. Voir également le commentaire sur la LEg Komm. GlG-KAUFMANN, art. 3, N 10 ; SHK GlG-WETZSTEIN / WOLFENSBERGER, art. 3, N 6

96 ATF 117 Ia 262, consid. 2a ss ; message concernant la LEg, FF 1993 I 1163, ici 1210 ; Komm. GlG-KAUFMANN, art. 3, N 18 ; SHK GlG-WETZSTEIN / WOLFENSBERGER, art. 3, N 13 ss

97 Message relatif à la LEg, p. 1210

98 ATF 117 Ia 262, consid. 2a ss ; HENSCH, ch. 2.2 ; Komm. GlG-KAUFMANN, art. 3, N 6

99 Message concernant la LEg, p. 1210 ; STEINER, p. 1286 ; voir également SHK-Wetzstein / Wolfensberger, art. 3, N 23 ss

100 Message concernant la LEg, p. 1210

isolés)[101]. Pour l'évaluer, le tribunal doit s'éloigner du cas concret et s'inspirer de l'expérience générale de la vie[102]. Il peut, le cas échéant, s'appuyer sur des probabilités et des données statistiques[103].

Dans le cas d'une discrimination indirecte individuelle, la partie plaignante doit rendre vraisemblable ou prouver qu'elle est traitée de manière significativement moins favorable qu'une personne de l'autre genre. Dans le cas d'une discrimination indirecte collective, il faut montrer que le groupe désavantagé est composé principalement de femmes (ou d'hommes), et que le désavantage est significatif. En cas de soupçon d'inégalité salariale concernant toute une catégorie professionnelle, la partie plaignante doit prouver que les tâches effectuées par le groupe en question sont des tâches typiquement féminines ou – plus rarement – typiquement masculines. Le Tribunal fédéral a renvoyé à cet égard à la jurisprudence de la CJCE, selon laquelle pour être jugée discriminatoire, la mesure doit affecter « de manière prépondérante des personnes de sexe féminin », un « nombre beaucoup plus élevé de femmes que d'hommes », ou un « pourcentage considérablement plus faible d'hommes que de femmes ». Lorsque cette proportion était connue dans les affaires jugées par la juridiction européenne, la part des femmes désavantagées par une réglementation était systématiquement supérieure à 80 %[104]. L'inégalité de traitement (différence de rémunération) doit par ailleurs être significative, un écart de 15 % étant déjà considéré comme important[105].

La LEg intègre toute une série de dispositions de procédure visant à améliorer la protection juridique en cas de discrimination fondée sur le genre. Ces dispositions viennent compléter celles du code de procédure civile et celles applicables en cas de conflits au travail, notamment pour ce qui concerne l'administration de la preuve et la conduite de la procédure : allègement du fardeau de la preuve (art. 6 LEg, selon lequel il suffit de rendre vraisemblable la discrimination considérée), protection contre le licenciement (art. 10 LEg). L'art. 7 LEg prévoit quant à lui la possibilité, pour les organisations de promotion de l'égalité entre femmes et hommes, d'intenter en leur nom propre une action collective dans certaines conditions (réglementation discriminatoire, p. ex.).

3.1.3. Applicabilité aux systèmes algorithmiques

La LEg n'établissant pas de distinction entre les décisions humaines et les décisions prises par des machines, elle s'applique également aux systèmes algorithmiques utilisés dans le cadre de processus décisionnels (recrutement, promotion, allocation de ressources, etc.). Lorsque les employeurs prennent des décisions avec l'aide d'un système algorithmique, ils sont tenus de veiller à ce qu'elles soient non discriminatoires[106].

Le mécanisme de l'allègement du fardeau de la preuve prévu dans la LEg se révèle particulièrement utile dans le cas d'une discrimination algorithmique, car il facilite la production de preuves. Ainsi, il suffit de réunir des indices de différences systématiques dans les résultats (décisions de recrutement ou de promotion, p. ex.) pour que les employeurs soient contraints d'apporter la preuve que la différence de traitement est objectivement justifiée et n'est pas fondée sur le genre.

---

101 STEINER, p. 1286
102 Voir aussi le schéma de contrôle dans CSHE LEg-Wetzstein / Wolfensberger, art. 3, N 27 ss
103 Message concernant la LEg, p. 1211
104 ATF 124 II 529, consid. 5g ; STEINER, p. 1287 s. (autres références)
105 ATF 142 II 49, consid. 6.2 ; ATF 130 III 145, consid. 4.2 ; arrêt du TF 2A.91/2007 du 25.02.2008, consid. 5.
106 HACKER, p. 1169

3.2. Loi sur l'égalité pour les handicapés (LHand)

3.2.1. Principes inscrits dans la Constitution

L'interdiction générale de discrimination à l'égard des personnes handicapées est inscrite à l'art. 8, al. 4, Cst.[107]Il ne s'agit pas ici d'un droit que les personnes concernées peuvent faire valoir directement, mais d'un mandat législatif visant à éliminer les inégalités qui frappent les personnes handicapées[108]. L'art. 8, al. 4, Cst. va ainsi moins loin que la deuxième phrase de l'art. 8, al. 3, Cst., qui pourvoit à l'égalité de droit et de fait entre les genres. Nombre de constitutions cantonales contiennent également des dispositions visant explicitement à garantir la protection des personnes handicapées contre la discrimination[109]. Comme les prescriptions établissant les principes d'égalité de traitement et d'interdiction de la discrimination, ces dispositions n'ont toutefois une signification et une portée autonomes que si la protection qu'elles confèrent aux personnes concernées va au-delà de celle accordée en vertu de la Constitution fédérale.

En précisant que des mesures doivent être prises, le mandat visant à éliminer les discriminations effectives et structurelles qui frappent les personnes handicapées considère expressément ces personnes comme appartenant à un groupe digne de protection[110].

3.2.2. Dispositions légales pertinentes

La LHand crée des conditions propres à faciliter aux personnes handicapées la participation à la vie de la société et interdit les inégalités de traitement fondées sur le handicap (art. 1 LHand). Elle s'applique principalement aux organes publics et aux particuliers qui assument des tâches publiques, mais aussi aux bâtiments et installations accessibles au public et aux rapports de travail en vertu de la loi sur le personnel de la Confédération (LPers), ainsi qu'aux particuliers qui fournissent des prestations au public (art. 6 LHand).

« Il y a inégalité lorsque les personnes handicapées font l'objet, par rapport aux personnes non handicapées, d'une différence de traitement en droit ou en fait qui les désavantage sans justification objective ou lorsqu'une différence de traitement nécessaire au rétablissement d'une égalité de fait entre les personnes handicapées et les personnes non handicapées fait défaut » (art. 2, al. 2, LHand). La loi ne précise cependant pas les exigences posées à une justification objective[111].

Comme la LEg, la LHand intègre des dispositions de procédure visant à faciliter l'application de droits. Les procédures sont ainsi gratuites dans certains cas (art. 10 LHand). Les organisations d'aide aux personnes handicapées ont par ailleurs qualité pour agir ou pour recourir en leur propre nom contre une inégalité qui affecte un nombre important de personnes (art. 9 LHand).

---

107 KIENER / KÄLIN / WYTTENBACH, ch. 1926 ss ; BSK BV-WALDMANN, art. 8, N 13, p. 124 ss
108 ATF 139 II 289, consid. 2.2 ; voir aussi KIENER / KÄLIN / WYTTENBACH, ch. 1926 ss
109 Art. 11, al. 4, Cst-ZH ; art. 8, al. 3, Cst-VD
110 BSK BV-WALDMANN, art. 8, N 126
111 NAGUIB, p. 1005

3.2.3. Applicabilité aux systèmes algorithmiques

Les dispositions de la LHand s'appliquent lorsque l'utilisation d'algorithmes dans des domaines dans lesquels les principes d'égalité de traitement et de non-discrimination sont tenus de s'appliquer (prestation de services, procédure de recrutement, attribution d'un logement, utilisation des transports publics, etc.) entraîne une discrimination des personnes en situation de handicap. Comme la LEg, la LHand n'établit pas de distinction entre les décisions humaines et les décisions prises par des machines. La décision fondée sur un algorithme touchant la personne handicapée doit induire une « différence de traitement particulièrement marquée et gravement inégalitaire » (art. 2, let. d de l'ordonnance sur l'égalité pour les handicapés, OHand).

Les dispositions de procédure sont également utiles dans le domaine de LHand pour assurer la mise en œuvre des prescriptions.

3.3. Loi fédérale sur la protection des données (LPD)

La LPD régit le traitement de données personnelles concernant des personnes physiques effectué par des personnes privées et des organes fédéraux (art. 2, al. 1, let. a et b, LPD). Le traitement des données personnelles par les autorités cantonales (et communales) est régi par les lois cantonales sur l'information et sur la protection des données[112].

3.3.1. Dispositions légales pertinentes

a. Article relatif au but de la loi et conformité aux principes de la bonne foi

Le principe du traitement égalitaire et non discriminatoire des données est inscrit dans l'article relatif au but de la LPD (art. 1 LPD), selon lequel la loi vise à protéger la personnalité et les droits fondamentaux des personnes physiques dont les données personnelles font l'objet d'un traitement. Ce principe découle par ailleurs de celui de la bonne foi applicable au traitement des données personnelles, énoncé à l'art. 6, al. 2, LPD.

b. Devoir d'informer en cas de décision automatisée

Si une décision est prise exclusivement sur la base d'un traitement automatisé des données personnelles, la personne concernée doit en être informée[113]. Le devoir d'informer inscrit à l'art. 21 LPD vaut uniquement si la décision affectant la personne est prise sans intervention humaine (décision entièrement automatisée)[114]. Cela vaut pour toutes les décisions individuelles, qu'elles soient prises par un particulier ou par un organe fédéral. La décision est considérée comme entièrement automatisée, que les

---

112 Par exemple : loi sur l'information et la protection des données (IDG) du canton de Zurich, du 12 février 2007 (LS 170.4) ; loi sur l'information et la protection des données (IDG) du canton de Bâle-Ville, du 9 juin 2010 (SG 153.260) ; loi sur la protection des données personnelles (LPrD) du canton de Vaud, du 1er septembre 2023 (RSV 172.65)

113 Art. 21, al. 1, LPD. Sont *concernées* les personnes pour lesquelles la décision automatisée a des effets juridiques ou qui sont affectées de manière significative par la décision (voir BSK DSG-HENSELER / VASELLA, art. 21, N 24 ss et 36 ss ; SHK DSG-PÄRLI / FLÜCK, art. 21, N 9 ss).

114 Un système algorithmique peut présenter différents niveaux d'automatisation : voir GORDON / LUTZ, p. 55 ; WEDER, p. 241.

algorithmes d'intelligence artificielle utilisés soient basés sur des règles ou sur l'apprentissage automatique[115].

Le législateur n'a pas défini clairement à quel moment on considère qu'une décision est prise sans intervention humaine[116]. Plusieurs auteurs et autrices se sont déjà penchés en détail sur cette question[117]. Malgré des avis parfois divergents, la doctrine considère pour l'essentiel qu'on est en présence d'une décision individuelle automatisée lorsqu'aucun être humain n'a procédé à l'analyse du contenu de la décision, à une sélection fondée sur une évaluation ou à des arbitrages entre plusieurs options[118]. L'intervention humaine dans la décision n'est pertinente que si la personne dispose de compétences particulières et du pouvoir d'influer sur le contenu de la décision[119]. Le devoir d'informer au sens de l'art. 21, al. 1 LPD ne s'applique pas si l'évaluation automatique sert uniquement à la préparation de la décision ou relève d'un système algorithmique d'aide à la prise de décision (*decision support system*) et qu'il y a une marge d'appréciation ou que le contenu de la décision est susceptible de pouvoir être contrôlé par une personne physique (*human-in-the-loop*)[120].

Outre le devoir d'informer, qui s'applique aux particuliers comme aux organes fédéraux, la LPD prévoit à l'art. 21, al. 4, si la décision individuelle automatisée émane d'un organe fédéral, l'obligation pour ce dernier de la qualifier comme telle (cela vaut en particulier pour les « décisions entièrement automatisées »)[121]. La personne concernée peut se référer à cette obligation pour demander à exercer son droit d'être entendue dans le cadre de la procédure automatisée[122].

c. Droit d'accès aux informations en cas de décision individuelle automatisée

En cas de décision individuelle automatisée, le droit d'accès prévu à l'art. 25 LPD est étendu. Les personnes concernées peuvent notamment demander à être informées de la logique sur laquelle se fonde la décision (art. 25, al. 2, let. f, LPD). Elles peuvent s'enquérir des données personnelles exploitées pour l'élaboration de la décision, également si celle-ci a été prise à l'aide d'algorithmes d'IA[123]. Il est ainsi possible de déterminer si des données liées à des caractéristiques protégées ont été utilisées et si elles ont pesé dans la décision. Le droit d'accès ne permet en revanche pas d'obtenir d'informations sur des données concernant des tiers[124].

---

115 SUTER, p. 296 s

116 En particulier, l'ordonnance ne précise pas les cas dans lesquels une décision est prise exclusivement sur la base du traitement des données personnelles (voir le message concernant la LPD, FF 2017 6565, ici 6679 s.).

117 Voir à ce sujet (sélection) BSK DSG-HENSELER / VASELLA, art. 21, N 21 ss ; SHK DSG-PÄRLI / FLÜCK, art. 21, N 6 ; BRAUN BINDER, Automatisierte Entscheidungen, p. 28 s. ; HUSI-STÄMPFLI / MORAND, ch. 442 ; OBRECHT, Verfügung, p. 40 s. ; RECHSTEINER, ch. 3 ; UTTINGER / GEISER, ch. 3.28

118 Voir SUTER, p. 297 s.

119 SHK DSG-PÄRLI / FLÜCK, art. 21, N 6

120 BSK DSG-Henseler / Vasella, art. 21, N 22 ; Husi-Stämpfli / Morand, ch. 442 ; Uttinger / Geiser, ch. 3.28

121 BRAUN BINDER, Dystopie, p. 256

122 BRAUN BINDER, Dystopie, p. 264

123 THOUVENIN / VOLZ, Rechtsrahmen, p. 633 s.

124 THOUVENIN / VOLZ, Rechtsrahmen, p. 633 s.

d. Droit au traitement par une personne physique

Les personnes concernées par une décision individuelle automatisée ont le droit de prendre position sur la décision (art. 21, al. 2, 1re phrase, LPD)[125]. Elles peuvent alors exiger qu'une personne physique prenne leur point de vue en considération et examine la décision individuelle automatisée (droit d'être entendu)[126]. Elles peuvent en outre demander que la décision individuelle automatisée soit revue par une personne physique (art. 21, al. 2, 2e phrase, LPD). Ce contrôle peut permettre l'identification d'une discrimination algorithmique et l'initiation des mesures nécessaires pour y remédier[127].

e. Analyse d'impact relative à la protection des données personnelles

Si le traitement des données est susceptible d'engendrer un risque élevé pour la personnalité ou les droits fondamentaux de la personne concernée, le ou la responsable du traitement[128] est tenu de procéder au préalable à une analyse d'impact relative à la protection des données personnelles (AIPD)[129]. Le niveau du risque s'évalue au regard de la nature, de l'étendue, des circonstances et de la finalité du traitement des données. Il peut être élevé en particulier en cas de recours aux nouvelles technologies[130]. Les circonstances du traitement des données doivent être appréciées au cas par cas[131]. L'élaboration de décisions individuelles automatisées et l'utilisation de systèmes d'IA sont considérées comme des traitements de données à haut risque[132]. L'AIPD couvre non seulement les décisions entièrement automatisées, mais aussi les systèmes algorithmiques utilisés comme aides à la décision (*decision support systems*)[133].

L'AIPD est un outil d'auto-évaluation de projets présentant un caractère sensible du point de vue du droit de la protection des données[134]. L'objectif est d'identifier et d'évaluer en détail les risques pour les personnes dont les données sont utilisées[135]. Une fois les risques identifiés, des mesures techniques et organisationnelles doivent être prises pour prévenir l'impact négatif des opérations de traitement envisagées sur les droits et libertés des personnes considérées[136]. En cas de subsistance d'un risque élevé

---

125 BSK DSG-HENSELER / VASELLA, art. 21, N 40 ; concernant les exceptions au droit d'être entendu, voir l'art. 21, al. 3, let. a et b, LPD, ainsi que l'art. 30, al. 2, PA, en relation avec l'art. 21, al. 4, LPD (critique à ce sujet BRAUN BINDER, Automatisierte Entscheidungen, p. 32 s.)

126 Art. 21, al. 2, LPD ; voir BSK DSG-HENSELER / VASELLA, art. 21, N 44 ss. ; SHK DSG-PÄRLI / FLÜCK, art. 21, N 4 et 13 ss ; HUSI-STÄMPFLI / MORAND, ch. 443 s. ; UTTINGER / GEISER, ch. 3.31

127 KÜRSTEINER, au ch. 59, voit dans le contrôle humain un potentiel de rectification des décisions algorithmiques potentiellement erronées, sans toutefois mentionner les risques associés.

128 Les *responsables* du traitement peuvent être des personnes privées ou des organes fédéraux qui, seuls ou conjointement avec d'autres, déterminent la finalité du traitement des données personnelles et les moyens utilisés (art. 5, let. j, LPD), voir aussi SHK DSG-BLONSKI, art. 22, N 5.

129 Art. 22, al. 1, LPD ; voir HUSI-STÄMPFLI / MORAND, ch. 469 ; UTTINGER / GEISER, ch. 3.36 ; SHK DSG-BLONSKI, art. 22, N 1

130 Sont notamment considérées comme *nouvelles technologies* les technologies liées à l'intelligence artificielle telles que les grands modèles linguistiques (*large language models, LLM*), l'*apprentissage profond (deep learning)* et les réseaux neuronaux (voir BSK DSG-GLATTHAAR / SCHRÖDER, art. 22, N 24).

131 HUSI-STÄMPFLI / MORAND, ch. 471

132 SHK DSG-Blonski, art. 22, N 19 ; BSK DSG-Glatthaar / Schröder, art. 22, N 42

133 Dans la mesure où les données personnelles au sens de l'art. 5, let. a, LPD subissent un traitement.

134 HUSI-STÄMPFLI / MORAND, ch. 469

135 BSK DSG-GLATTHAAR / SCHRÖDER, art. 22, N 1 ; HUSI-STÄMPFLI / MORAND, ch. 469 et 472

136 Art. 22, al. 3, LPD ; BSK DSG-GLATTHAAR / SCHRÖDER, art. 22, N 63 ss ; HUSI-STÄMPFLI / MORAND, ch. 472 s.

malgré la mise en œuvre des mesures, le préposé fédéral à la protection des données (PFPDT) doit être informé[137].

L'AIPD permet d'identifier et d'évaluer les risques d'un projet de traitement des données en amont de celui-ci et de définir les mesures visant à assurer la protection des données utilisées. La LPD ne prévoit pas d'obligation de revue ni d'actualisation régulière de l'AIPD, sauf en cas de modification d'aspects essentiels du traitement des données ou de survenance de nouveaux risques. Les systèmes d'IA étant amenés à évoluer, des risques de discrimination qui n'auraient été pas identifiés au départ risquent ainsi de ne pas non plus être identifiés et éliminés à un stade ultérieur.

3.3.2. Excursus : Traitement de données personnelles à des fins de détection de situations de discrimination

Si la législation relative à la protection des données peut contribuer à identifier et à prévenir la discrimination, elle peut aussi avoir l'effet inverse. Le Règlement général sur la protection des données (RGPD) de l'UE interdit les opérations de traitement portant sur des catégories particulières de données personnelles telles que l'origine ethnique, les convictions religieuses ou l'orientation sexuelle (art. 9 RGPD). Pour éviter les discriminations, les développeurs et développeuses et les utilisateurs et utilisatrices de systèmes d'IA tendent à exclure le traitement des données liées à ces caractéristiques (l'origine ethnique, p. ex.). Le problème est qu'il n'est alors plus possible de vérifier si un système d'IA discrimine certaines personnes au regard de ces caractéristiques protégées, pour la simple et bonne raison qu'elles ne sont pas (ou plus) disponibles. Pour éviter une telle situation, le législateur européen a prévu, dans la mesure où cela est strictement nécessaire pour assurer la détection et la correction des biais propres aux systèmes d'IA à haut risque, une disposition particulière autorisant les fournisseurs et fournisseuses de ces systèmes à traiter des données à caractère personnel (art. 10 par. 5 Règlement UE-IA).

Une telle réglementation n'est pas nécessaire en droit suisse. D'une part, la LPD n'interdit pas expressément le traitement de données personnelles sensibles. D'autre part, l'atteinte à la personnalité peut être motivée par un intérêt privé ou public prépondérant (art. 31, al. 1, LPD), ce qui est le cas lorsque les développeurs et développeuses ou les utilisateurs et utilisatrices privés de systèmes d'IA traitent certaines données personnelles dans le but de détecter et de prévenir les discriminations. Contrairement aux particuliers, les organes fédéraux ne sont en droit de traiter des données personnelles que s'il existe une base légale (art. 34, al. 1, LPD). La base légale doit être prévue dans une loi au sens formel dès lors que les données traitées sont des données sensibles (art. 34, al. 2, let. a, LPD). À supposer qu'une base légale autorise le traitement de données personnelles (particulièrement sensibles), il conviendra d'examiner si celle-ci est suffisamment robuste pour autoriser l'utilisation de ces données en vue d'évaluer le risque de discrimination d'un système d'IA.

3.3.3. Efficacité

Il est possible de s'appuyer sur les dispositions de la loi sur la protection des données pour lutter contre la discrimination algorithmique. L'AIPD en particulier peut contribuer à l'identification d'une

---

[137] Il est possible de déroger au devoir d'informer en désignant un conseiller à la protection des données ou un responsable de la protection des données (art. 23 LPD) ; HUSI-STÄMPFLI / MORAND, ch. 478 ; UTTINGER / GEISER, ch. 3.47 s.

discrimination algorithmique avant même l'utilisation d'un système d'IA et permettre la mise en œuvre de mesures préventives. En l'absence d'obligation de réévaluer le projet de traitement des données, cette analyse ne garantit cependant pas la détection des discriminations survenant ou se manifestant au moment du traitement effectif des données. Il faudrait pour cela que les décisions algorithmiques soient soumises à des contrôles réguliers.

Le devoir d'informer et, pour les organes fédéraux, l'obligation de qualifier comme telle toute décision individuelle automatisée, garantissent que le recours à des systèmes de décision exclusivement automatisés soit connu des personnes concernées. Le droit d'accès que leur confère la loi sur la protection des données leur permet également d'obtenir des informations sur la logique de ces systèmes. Si ces droits et obligations ne peuvent pas empêcher la discrimination algorithmique, ils constituent des moyens utiles pour identifier les discriminations et obtenir les informations nécessaires pour y remédier[138]. Ces moyens n'offrent cependant pas une protection complète contre la discrimination algorithmique, car les systèmes d'IA peuvent présenter des caractéristiques différentes lors de leurs phases de développement et d'application (voir le ch. 2.1). Le fait que le devoir d'informer (art. 21 LPD) et le droit d'accès (art. 25, al. 2, let. f, LPD) ne s'appliquent qu'aux décisions entièrement automatisées pose aussi un problème. Les décisions semi-automatisées présentent en effet des risques de discrimination comparables[139].

4. Dispositions particulières en droit public

L'interdiction générale de la discrimination inscrite dans la Constitution (art. 8, al. 2, Cst.) vaut pour toute action des autorités, à tous les niveaux étatiques. D'autres dispositions protègent également contre la discrimination, p. ex. les garanties de procédure prévues à l'art. 29 Cst. (ch. 4.1), ou les exigences constitutionnelles concrétisées au niveau de la loi (ch. 4.2). Ces dispositions, ainsi que leur applicabilité et leur efficacité en lien avec l'utilisation de systèmes algorithmiques, sont brièvement présentées ci-après (ch. 4.3).

4.1. Garanties de procédure, en particulier l'obligation de motiver

L'obligation, pour les autorités, de motiver leurs décisions est un corollaire du droit d'être entendu garanti par la Constitution[140]. Le contenu et l'ampleur de la motivation appropriée dépendent des circonstances propres à chaque cas.[141] Cette obligation remplit plusieurs fonctions : une fonction juridictionnelle, une fonction d'acceptation, une fonction de garantie de la sécurité juridique, mais aussi une fonction d'autocontrôle et de contrôle externe[142]. Ces deux dernières fonctions sont particulièrement importantes en matière de discrimination.

La fonction d'autocontrôle a pour effet que les autorités veillent à être en conformité avec l'ordre juridique en vigueur dans l'accomplissement des tâches qui leur incombent en vertu de la Constitution

---

138 BRAUN BINDER, Künstliche Intelligenz, p. 475

139 Pour l'administration publique, voir CHRISTEN et al., TA-SWISS 72/2020, p. 278

140 Art. 29, al. 2, Cst., voir ATF 143 IV 40, consid. 3.4.3 ; voir aussi HÄFELIN / MÜLLER /UHLMANN, ch. 1070 ss.

141 Pour plus d'informations sur ce sujet, voir WIEDERKEHR, p. 481 ss.

142 Concernant les différentes fonctions de l'obligation de motiver les décisions, voir KNEUBÜHLER, p. 94 ss ; VILLIGER, p. 159 ss

et de la législation, et à prendre des décisions conformes au droit et adaptées au cas d'espèce[143]. Les autorités sont notamment tenues à la réalisation des droits fondamentaux dans les relations qui lient les particuliers entre eux (art. 35, al. 2, Cst.), ce qui vaut également pour l'interdiction de la discrimination au sens de l'art. 8, al. 2, Cst. Lorsqu'elle est amenée à prendre une décision, la personne en charge d'un dossier peut être influencée par différents facteurs subjectifs et objectifs[144]. Procéder à un autocontrôle lui permet d'examiner si des motifs non objectifs pèsent dans sa décision[145]. En cas de contestation, la motivation permet à l'instance de recours de comprendre sur quelle base la décision a été prise et d'évaluer si le jugement rendu par l'instance précédente est bien conforme au droit[146]. L'obligation de motiver remplit en ce sens également une fonction de contrôle externe.

4.2. Autres interdictions de la discrimination

L'égalité de droit et l'interdiction de la discrimination garanties par la Constitution et le droit international sont concrétisées au niveau de la loi. Certaines dispositions de l'ordre juridique garantissent ainsi l'égalité de droit entre les genres[147]. Il y est le plus souvent explicitement précisé que les droits garantis aux femmes sont les mêmes que ceux garantis aux hommes[148]. Certaines dispositions font indirectement référence à l'égalité de droit et la renforcent de manière positive. Dans certaines circonstances, le genre peut p. ex. être retenu comme critère pour l'attribution ciblée de concessions, d'accréditations et de subventions, ou pour l'accès à des prestations et des institutions[149]. Il manque de dispositions légales comparables pour prévenir la discrimination raciale.

La loi prévoit une protection explicite contre la discrimination, p. ex. pour ce qui concerne l'accès aux prestations des autorités[150], aux institutions pédagogiques[151] ou à d'autres institutions[152], ainsi que la perception de prestations octroyées par l'État[153].

L'ordre juridique contient également des dispositions régissant la prévention de la discrimination résultant du comportement humain. Certaines visent ainsi explicitement à prévenir le harcèlement sexuel, les agressions sexuelles ou les violences sexuelles fondées sur l'appartenance de genre[154].

L'ordre juridique offre également une protection contre la discrimination raciale en prohibant certaines activités pour leur effet discriminatoire[155]. L'application de sanctions en cas de discrimination permet

---

143 HERMSTRÜWER, p. 498 ss ; KNEUBÜHLER, p. 96 ; MOSIMANN, p. 20

144 MEYLAN, p. 314 ; NINK, p. 40 ss. Les émotions font p. ex. partie des facteurs subjectifs, les conditions météorologiques des facteurs objectifs.

145 KNEUBÜHLER, p. 96 ; MOSIMANN, p. 20

146 Arrêt du TF C-6504/2011 du 03.12.2013, consid. 6.2.5 ; KNEUBÜHLER, p. 111 ss ; MOSIMANN, p. 20

147 Confédération : art. 61, al. 2, let. b, OAMal ; art. 3, al. 1, OPersTF. Cantons : § 11 al. 1 du contrat entre la FHNW et le canton BS

148 P. ex. art. 3, al. 2, let. d et g LPers, en relation avec l'art. 3 OPersTF

149 Art. 14, al. 1, let. a, LEEJ. On trouve d'autres exemples dans le droit fédéral (art. 12, al. 1, LMP ; art. 30, al. 1, let. a, ch. 5, LEHE) et dans le droit cantonal (art. 5, let. f, Cst-AG ; art. 15, al. 1, let. n, LIEPR-VD ; art. 3, al. 2, Lsubv-VD ; art. 3, al. 2, RLSubV-VD ; art. 7, al. 1, let. b, RLRNSS-VD ; § 3, al. 1, let. b, SVO-ZH).

150 Art. 4, al. 1, AOCCF-VD

151 P. ex. § 30, al. 1, HOPH-ZH

152 Accès aux crèches, p. ex. (p. ex. § 13, al. 1, let. b, TBG-BS)

153 P. ex. § 10, al. 3, let. b, WRFV-BS

154 Confédération : art. 3, al. 2, OPersTF. Cantons : LOVD-VD, en relation avec RLOVD-VD ; § 9, al. 1 et 2, PVF-ZH ; ensemble du R-UZH-SB-ZH

155 Art. 171*c*, al. 1, CPM

également de prévenir la discrimination indirecte. Certaines dispositions punissent ainsi l'incitation publique à la discrimination[156].

Les dispositions de procédure qui facilitent la prétention juridique à un traitement égalitaire et à la non-discrimination offrent elles aussi une protection indirecte contre la discrimination, tandis que la possibilité de faire appel à des autorités de conciliation peut contribuer à promouvoir l'égalité femmes-hommes.

L'égalité salariale entre les femmes et les hommes, garantie par la Constitution, est-elle aussi concrétisée au niveau de la loi[157].

4.3. Efficacité des dispositions de droit public

L'analyse de la protection contre la discrimination inscrite dans le droit international, la Constitution et la législation interne montre que diverses dispositions de droit public offrent une protection contre la discrimination raciale ou la discrimination fondée sur le genre. Il n'existe en revanche pas de dispositions concrètes visant à protéger contre la discrimination algorithmique.

Un certain nombre de problèmes surviennent en cas de recours à des algorithmes. Comme nous l'avons évoqué plus haut, l'obligation de motiver les décisions remplit entre autres une fonction d'autocontrôle et de contrôle externe, qui peuvent contribuer à la prévention ou à l'identification de situations de discrimination (voir le ch. 4.1.). L'efficacité de ces fonctions se trouve amoindrie dès lors que des systèmes algorithmiques sont utilisés, et ce pour des raisons techniques, mais aussi pour des raisons inhérentes à l'utilisation même de ces systèmes[158]. Le recours à des systèmes d'apprentissage automatique complique en particulier l'identification des facteurs ayant influencé la décision, ce qui a un impact négatif sur l'efficacité de la fonction d'autocontrôle et de contrôle externe[159]. Si ces facteurs sont malgré tout identifiés, les informations risquent d'être difficiles à décrypter, de sorte que la personne en charge du dossier ne pourra pas analyser la décision ni détecter si le système a pris en compte des paramètres discriminatoires ou non[160]. La juridiction de recours sera confrontée au même problème en cas d'action en justice[161].

Par ailleurs, le degré d'automatisation du système affecte les fonctions de différentes manières. Dans les procédures entièrement automatisées, il ne peut y avoir de phase d'auto-analyse (et donc de fonction d'autocontrôle) dans la mesure où aucune personne physique n'est impliquée dans le processus décisionnel[162]. Dans le cas de procédures partiellement automatisées, l'efficacité de la fonction d'autocontrôle peut être limitée par le biais d'automatisation (voir le ch. 2.1).

L'AIPD peut être envisagée en première intention, dans une situation de décision automatisée, pour compenser l'absence de fonction d'autocontrôle de l'obligation de motiver la décision. Elle vise à

---

156 Art. 171*c*, al. 1, CPM

157 P. ex. art. 44, al. 2, let. f, LMP ; art. 36*a* LPers

158 Concernant l'impact sur les fonctions de l'obligation de motiver les décisions dans le contexte de l'utilisation d'algorithmes d'apprentissage automatique ou basés sur des règles, voir BRAUN BINDER / OBRECHT, WP Begründung, p. 4 ss

159 BRAUN BINDER / OBRECHT, Begründung, p. 713 s. Voir le point 2.1.1 ci-dessus, pour le manque de traçabilité

160 BRAUN BINDER / OBRECHT, Begründung, p. 714 ss

161 BRAUN BINDER/OBRECHT, Begründung, p. 715

162 BRAUN BINDER / OBRECHT, Begründung, p. 713

garantir l'identification de discriminations potentielles avant même le recours à un système d'IA[163]. Cela étant, comme elle est réalisée en amont du traitement des données, elle ne compense la fonction d'autocontrôle que partiellement. Elle ne peut donc se substituer à l'autocontrôle inhérent à l'obligation de motiver les décisions administratives. D'autres mesures avant ou après la motivation de la décision sont donc nécessaires[164].

5. Interdiction de la discrimination en droit privé

Si l'interdiction générale de la discrimination n'existe pas dans le droit privé suisse, certaines dispositions légales, comme le droit du travail, visent à interdire ou à prévenir la discrimination dans des situations données. Ces normes et leur pertinence pour prévenir la discrimination algorithmique sont décrites brièvement ci-après (respectivement aux ch. 5.1 à 5.5, et au ch. 5.6).

5.1. Protection de la personnalité en droit civil (art. 28 CC)

En vertu de l'art. 28 du code civil (CC), toute personne qui subit une atteinte illicite à sa personnalité est en droit de déférer l'auteur ou l'autrice de l'acte considéré en justice et de demander réparation. La protection de la personnalité en droit civil est une protection de portée générale. L'art. 28 CC peut être invoqué en cas de discrimination par des particuliers. La notion de personnalité n'est pas définie dans la loi, elle est consacrée par la doctrine et la jurisprudence[165]. Selon la jurisprudence du Tribunal fédéral, elle comprend tout ce qui sert à individualiser une personne et qui apparaît comme digne de protection au regard des relations entre les différents individus et dans le cadre des bonnes mœurs[166]. Elle couvre donc tous les droits inaliénables inhérents à la personne humaine. La protection de la personnalité se rapproche de la protection de la dignité humaine, et rend effectifs les droits fondamentaux inhérents à la personnalité, y compris entre particuliers[167].

L'art. 28 CC contient implicitement un droit à un traitement non discriminatoire[168]. Les caractéristiques protégées par la Constitution et le droit international tels que l'origine, le genre, le handicap, l'orientation sexuelle, le mode de vie, l'âge et l'appartenance religieuse font partie des composantes essentielles de la personnalité. Ils sont généralement immuables (genre, handicap, âge, orientation sexuelle) ou ne peuvent être modifiés sans impliquer, pour les personnes concernées, des contraintes disproportionnées (convictions religieuses)[169]. Selon une partie de la doctrine, la discrimination constitue toujours une atteinte à la personnalité[170]. Bien que ce point de vue soit convaincant, il n'a pas été consacré par la jurisprudence du Tribunal fédéral. Par ailleurs, la question de savoir si l'art. 28 CC peut également couvrir les formes de discrimination indirecte n'a pas encore été tranchée[171]. L'art. 28 CC protège contre

---

163 BRAUN BINDER / OBRECHT, Begründung, p. 716
164 BRAUN BINDER/OBRECHT, Begründung, p. 714 s.
165 Message concernant la révision du CC, FF 1982 II 661, ici 682 ss ; OFK ZGB-BÜCHLER, art. 28, N 2 ; NAGUIB, p. 1005
166 ATF 95 II 491 ss ; ATF 70 II 130 ; ATF 45 II 625
167 GÖKSU, Drittwirkung, p. 89 ss, 101 ; NAGUIB, p. 1005 ; WALDMANN, Diskriminierungsverbot, p. 68
168 NAGUIB, p. 1006 ; HAUSAMMANN, p. 18 ; WEBER / HENSELER, p. 40 ; voir aussi GÖKSU, Rassendiskriminierung, ch. 214 ss
169 NAGUIB, p. 1007 ; PÄRLI, p. 129 ; PÄRLI / CAPLAZI / SUTTER, p. 133
170 GAUCH / SCHLUEP / SCHMID / REY, ch. 1111 ; GÖKSU, Rassendiskriminierung, ch. 247 ; NAGUIB, p. 1007 ; PÄRLI, p. 133
171 NAGUIB (p. 1007) est d'avis que le cas de la discrimination indirecte doit être examiné plus avant.

les atteintes individuelles à la personnalité. Le problème, dans le cas d'une discrimination indirecte, est qu'aucun lien direct ne peut être établi entre la discrimination et le critère distinctif protégé considéré. L'application de l'art. 28 CC serait envisageable dans les cas de discrimination indirecte dans lesquels un préjudice individuel et un lien avec la personnalité peuvent être démontrés (opportunités professionnelles, exclusion sociale, etc.).

Il convient de rappeler que toute atteinte à la personnalité n'est pas illicite. Certaines atteintes peuvent ainsi être motivées par le consentement de la personne concernée, ou par des intérêts publics ou privés prépondérants.
Dans le second cas, le juge doit procéder à une pesée des intérêts pour déterminer si l'intérêt invoqué pour motiver l'atteinte prime l'intérêt privé de la victime ou d'autres intérêts[172]. L'accès des hommes et des femmes à des thermes à des horaires distincts peut par exemple être justifié par un impératif de protection de la sphère intime et de la sécurité personnelle. Un intérêt purement commercial ne sera en revanche pas retenu[173].

Si l'on admet que les discriminations constituent des atteintes à la personnalité au sens de l'art. 28 CC, celui-ci s'applique alors également en cas de discrimination directe par les systèmes algorithmiques. Les cas de discrimination indirecte ne peuvent en revanche être couverts, l'art. 28 CC protégeant la personnalité des individus et n'ayant pas vocation à prévenir la discrimination de groupes de personnes.

En cas de discrimination d'un groupe de personnes, une action collective en justice peut être formée selon les dispositions du code de procédure civile (CPC). En vertu de l'art. 89 CPC, les associations et les autres organisations d'importance nationale ou régionale qui sont habilitées aux termes de leurs statuts à défendre les intérêts d'un groupe de personnes déterminé peuvent, en leur propre nom, agir pour l'atteinte à la personnalité des membres de ce groupe. Il est possible de former une action en cessation ou en suppression de l'atteinte, ou une action en constatation de violation de la personnalité, mais pas une action en dommages-intérêts ni en remise du gain. Le droit en vigueur prévoit donc à la fois des moyens matériels et des moyens procéduraux pour agir en justice en cas de discrimination algorithmique (et d'autres formes de discrimination) par des particuliers.

5.2. Rapports commerciaux dans les limites des principes généraux de la loi, des bonnes mœurs et de l'ordre public (art. 19 et 20 CO)

Si le principe de la liberté contractuelle est inscrit dans le droit des contrats, il est toutefois soumis à certaines restrictions. Ainsi, les art. 19 et 20 du code des obligations (CO) disposent que le contrat est nul s'il a pour objet une chose illicite ou contraire aux bonnes mœurs. Sont considérés comme contraires aux bonnes mœurs au sens des art. 19 et 20 CO les contrats qui transgressent la morale dominante, à savoir qui vont à l'encontre du sens commun de la décence ou des valeurs et principes éthiques inhérents à l'ordre juridique dans son ensemble[174]. Les art. 19 et 20 CO s'appliquent également aux contrats contenant des clauses discriminatoires. Un contrat qui désavantagerait des personnes en raison de leur

---

172 BSK ZGB I-MEILI, art. 28, N 49 ; NAGUIB, p. 1008
173 NAGUIB, p. 1009 (qui mentionne d'autres références)
174 ATF 136 III 474, consid. 3 ; ATF 132 III 455, consid. 4.1 ; AFT 129 III 604, consid. 5.3 ; ATF 123 III 101, consid. 2 ; ATF 115 II 232, consid. 4a ; BSK OR I-MEISE / HUGUENIN, art. 19/20, N. 32

origine ethnique, de signes extérieurs, de caractéristiques culturelles ou de leur appartenance religieuse serait déclaré nul, car considéré comme contraire aux bonnes mœurs[175].

Les art. 19 et 20 CO peuvent offrir une protection contre la discrimination découlant d'accords contractuels. Cela vaut également lorsque les contrats sont rédigés à l'aide d'une IA générative. Une clause dans un contrat de travail prévoyant, en cas de résiliation par l'employeur, un délai de préavis plus court pour les personnes d'une certaine origine, serait ainsi considérée comme nulle.

5.3. Violation de la bonne foi (art. 2 CC)

L'art. 2 CC prescrit l'exercice des droits et des obligations selon les règles de la bonne foi en droit privé. L'abus manifeste d'un droit constitue une violation du principe de la bonne foi et ne bénéficie d'aucune protection juridique en droit privé (art. 2, al. 2, CC)[176]. Il y a abus de droit lorsqu'une institution ou une compétence juridique est utilisée à des fins étrangères à son objectif ou lorsque les intérêts poursuivis par l'exercice du droit ne sont pas prépondérants par rapport aux intérêts lésés d'un tiers. Les droits subjectifs peuvent faire l'objet d'abus, mais aussi toute une institution juridique, p. ex. la liberté contractuelle[177]. Dans certaines circonstances, les discriminations peuvent également être considérées comme un abus de droit. Cela pourrait être le cas si des personnes présentant certaines caractéristiques protégées (personnes d'une certaine origine, p. ex.) se voient systématiquement proposer des conditions contractuelles moins favorables (primes d'assurance responsabilité civile automobile plus élevées, p. ex.).

La liberté contractuelle ne doit pas légitimer la programmation d'un algorithme engendrant des conditions contractuelles moins avantageuses pour des personnes présentant certaines caractéristiques. La protection offerte par l'art. 2 CC se limite cependant à la discrimination directe intentionnelle, une pratique abusive ne pouvant pas être invoquée en cas de discrimination involontaire.

5.4. Droit du travail

Outre la loi sur l'égalité, qui garantit l'égalité entre les genres, et la loi sur l'égalité pour les handicapés, qui offre une protection spécifique pour les travailleurs et travailleuses en situation de handicap, certaines dispositions générales du droit du travail permettent de prévenir la discrimination dans les rapports de travail.

Le devoir d'assistance (art. 328 CO), qui vise à garantir la protection de la personnalité des travailleurs et travailleuses, s'impose ainsi aux employeurs. En vertu de l'art. 328 CO (et de l'art. 6 de la loi sur le travail, LTr), les employeurs sont tenus de protéger la personnalité, la santé et l'intégrité des travailleurs et travailleuses : ils se doivent de prendre des mesures adaptées aux conditions de l'exploitation lorsque celles-ci sont nécessaires, applicables en l'état de la technique et raisonnablement exigibles.

Le devoir d'assistance, en droit du travail, peut offrir une protection contre la discrimination, notamment contre la discrimination arbitraire et individuelle[178]. Cette protection est néanmoins limitée, la liberté contractuelle (et par là même les clauses des contrats individuels) ayant la primauté[179]. Un employeur

---

175 HAUSAMMAN, p. 18
176 KLETT, p. 177
177 NAGUIB, p. 1009 ; KRAMER, ch. 114 (qui mentionne d'autres références sur les art. 19 et 20 CO)
178 PK OR-STREIFF / VON KAENEL/RUDOLPH, art. 328, N 2
179 ATF 129 III 276, consid. 3.1 ; PK OR-STREIFF / VON KAENEL / RUDOLPH, art. 328, N 2

peut ainsi tout à fait décider de favoriser certains employés ou certaines employées par rapport à d'autres. Une telle décision ne constitue une violation de l'interdiction de discrimination que si elle ne peut être justifiée par des motifs objectifs, désavantage clairement le travailleur concerné ou la travailleuse concernée et laisse transparaître une sous-évaluation de la personnalité du travailleur ou de la travailleuse qui soit blessante pour lui[180].

L'art. 328 CO s'applique pendant toute la durée des rapports de travail[181]. Les tribunaux ont notamment jugé discriminatoires et constituant une violation du devoir d'assistance le refus d'embaucher une personne en raison de la couleur foncée de sa peau[182] et la décision d'une entreprise de nettoyage de ne pas employer de personnes originaires des Balkans[183]. Le Tribunal fédéral a par ailleurs estimé qu'une entreprise avait également manqué à son devoir d'assistance en ne prenant aucune mesure pour empêcher qu'un employé soit régulièrement victime d'agressions racistes et doive de ce fait être mis en arrêt maladie[184].

En cas de licenciement discriminatoire, la personne concernée peut invoquer la disposition relative au congé abusif donné par l'employeur (art. 336 CO). Selon la doctrine, un licenciement pour des motifs racistes ou religieux doit être considéré comme abusif (art. 336, al. 1, let. a et b, CO)[185]. La jurisprudence a jugé abusif le licenciement donné à des travailleuses musulmanes qui refusaient d'ôter leur foulard[186]. Le licenciement signifié est valable, même s'il est abusif, mais ouvre droit à une indemnisation.

Le droit du travail prévoit en outre un principe d'égalité de traitement pour les ressortissants et ressortissantes des États de l'UE / AELE, qui peuvent invoquer l'interdiction de discrimination prévue à l'art. 9 de l'annexe I de l'accord sur la libre circulation des personnes (ALCP), en relation avec l'art. 2 ALCP.

Jusqu'à présent, le Tribunal fédéral n'a fait valoir l'application de l'art. 328 CO que dans des cas de discrimination individuelle arbitraire. Seules les discriminations à l'égard de travailleurs et travailleuses individuels ou de petits groupes de personnes ont ainsi pu être prises en compte. La question de l'application de cette disposition à des groupes plus importants ou à des catégories entières de travailleurs et travailleuses n'a pas été tranchée[187].

---

180 ATF 129 III 276, consid. 3.1

181 L'application de l'art. 328 CO dès le processus de recrutement reste controversée dans la doctrine (voir PK OR-STREIFF / VON KAENEL / RUDOLPH, art. 328, N 21, et renvoi à d'autres références). Le Tribunal fédéral l'a confirmé dans son arrêt 2C_103/2008 du 30.06.2008, sans motiver sa décision. Dans les jugements cités dans les notes de bas de pages suivantes, les juges fédéraux ont au moins reconnu l'effet *ex ante* de l'art. 328 CO, tandis que cette question n'a pas été abordée dans l'affaire jugée par le tribunal des prud'hommes de Zurich. Si l'on considère que l'art. 328 CO ne s'applique qu'à partir du moment où la signature du contrat de travail est intervenue, il faut se référer à l'art. 28 CC pour les discriminations survenues durant le processus de recrutement.

182 Voir l'arrêt du 10 octobre 2005 du Tribunal de prud'hommes de l'arrondissement de Lausanne (T 304.021563), dans : https://lexcampus.swisslex.ch/doc/unknown/47e875e1-c6e3-493c-8632-d8223784821c/citeddoc/bb46996a-5ece-4b76-8ec2-a85f5941a1a3/source/document-link (arguments soutenant la thèse de la discrimination de KURT PÄRLI)

183 Voir l'arrêt du du 13 janvier 2006 du Tribunal de prud'hommes de Zurich (2e section), affaire n° AN 050401/U 1, dans : ARV 2006, p. 23 ss (arguments soutenant la thèse de la discrimination de KURT PÄRLI)

184 Le congé donné par l'employeur en raison de l'incapacité de travail du salarié a été considéré comme abusif (ATF 4A_215/2022 du 23.08.2022).

185 PK OR-STREIFF / VON KAENEL / RUDOLPH, art. 328, N 12 ; PK OR-Streiff / von Kaenel / Rudolph, art. 336, N 3

186 Arrêt CIV 16 1317 du tribunal régional de Berne-Mittelland du 08.09.2016 ; arrêt du tribunal de district d'Arbon, JAR 1991, p. 254

187 Avis contraire : PK OR-STREIFF / VON KAENEL/RUDOLPH, art. 328, N 12

En conséquence, s'il ne fait aucun doute que l'art. 328 CO peut être invoqué en cas de discrimination directe par des systèmes algorithmiques, son application aux discriminations algorithmiques indirectes est encore une question ouverte.

5.5. Autres normes

Aux côtés des dispositions légales déjà mentionnées, d'autres normes couvrent la discrimination par des particuliers dans des cas spécifiques. La loi fédérale sur l'analyse de la génétique humaine (LAGH) interdit p. ex. la discrimination fondée sur le patrimoine génétique (art. 4 LAGH). Applicable aux autorités publiques comme aux particuliers, cette interdiction n'est pas absolue, la décision incriminée pouvant, là aussi, être motivée[188]. L'application de l'art. 4 LAGH peut être pertinente en cas d'utilisation de systèmes algorithmiques dans le cadre d'analyses génétiques.

En droit du bail, les dispositions relatives à la protection contre la résiliation du bail peuvent offrir une protection contre la discrimination. Selon l'art. 271, al. 1, CO, la résiliation donnée par le bailleur est ainsi annulable si le locataire fait valoir de bonne foi des prétentions découlant du bail. Cela peut être le cas lorsque la résiliation donnée se fonde sur des critères protégés.

5.6. Efficacité des dispositions de droit privé

L'analyse des dispositions de droit privé montre que celles-ci n'offrent pas une protection générale contre la discrimination par des particuliers. En dehors des domaines d'application de la LEg et de la LHand, cette protection ne peut être garantie que de manière indirecte, la personne concernée devant faire valoir une atteinte à la personnalité ou la violation d'une autre disposition générale du droit privé (violation de la bonne foi, comportement abusif, etc.), ce qui ne sera pas nécessairement reconnu en cas d'action en justice. Le droit privé est donc source d'insécurité juridique. Cela vaut également pour la discrimination par les systèmes algorithmiques, tout particulièrement en raison de la complexité technique inhérente à ces systèmes.

6. Interdiction de la discrimination dans le droit pénal (art. 261bis CP)

L'art. 261bis du code pénal (CP) prévoit une interdiction quasi générale de l'incitation à la haine et la discrimination fondée sur l'appartenance raciale, ethnique ou religieuse, ou sur l'orientation sexuelle. Cette disposition, approuvée en votation populaire, a été intégrée dans le code pénal pour satisfaire aux exigences découlant la Convention internationale sur l'élimination de toutes les formes de discrimination raciale (CERD). Initialement limitée à la discrimination fondée sur l'appartenance raciale, ethnique et religieuse, elle a vu son champ d'application étendu à la discrimination fondée sur l'orientation sexuelle en 2020[189].

L'art. 261bis, al. 1, CP punit l'incitation publique à la haine ou à la discrimination envers une personne ou un groupe de personnes en raison de l'appartenance raciale, ethnique ou religieuse ou de l'orientation sexuelle. Les al. 2 et 3 répriment la diffusion publique d'idéologies visant à rabaisser ou à dénigrer de façon systématique une personne ou un groupe de personnes, ainsi que l'organisation ou l'encouragement d'actions de propagande, ou la participation à de telles actions. En vertu de l'al. 4, le

---

188 Message concernant la LAGH, FF 2017 5253, ici 5657

189 Le droit pénal militaire comprend une disposition similaire (art. 171*c* CPM).

fait de rabaisser ou de discriminer publiquement une personne ou un groupe de personnes d'une façon qui porte atteinte à la dignité humaine – que ce soit par la parole, l'écriture, l'image, le geste, par des voies de fait ou de toute autre manière – ou de nier, minimiser grossièrement ou chercher à justifier un génocide ou d'autres crimes contre l'humanité est passible de sanctions pénales. En vertu de l'art. 261bis, al. 5, CP, personne ne doit se voir refuser une prestation destinée à l'usage public en raison de son appartenance raciale, ethnique ou religieuse, ou de son orientation sexuelle. Cette interdiction couvre non seulement le refus de prestation, mais aussi le refus de conclure un contrat portant sur la fourniture d'une prestation destinée à l'usage public[190].

Les dispositions de l'art. 261bis, art. 1 à 4, peuvent être invoquées en cas de diffusion de contenus interdits à l'aide de systèmes d'IA, ainsi que d'utilisation de systèmes d'IA générateurs de tels contenus. L'art. 261bis, al. 5, CP s'applique aussi lorsque la décision de fournir une prestation destinée au public ou de conclure un contrat portant sur la fourniture d'une telle prestation se fonde sur un système algorithmique.

7. Conclusion intermédiaire

Le droit suisse contient de nombreuses dispositions qui permettent de circonscrire et de prévenir la discrimination. Le principe de l'interdiction de la discrimination garanti à l'art. 8, al. 2, Cst. vaut également pour les autorités publiques. L'applicabilité de cette interdiction générale de la discrimination aux formes indirectes de discrimination n'a pas encore été clarifiée du point de vue juridique. À ce jour, il n'existe pas de jurisprudence établie ni de critères clairs permettant d'identifier les cas de discrimination indirecte et les conditions d'application de l'art. 8, al. 2, Cst. De même, la discrimination intersectionnelle est difficile à appréhender sur le plan juridique. On peut donc légitimement se demander si la réglementation actuelle est suffisante pour protéger contre la discrimination algorithmique émanant des autorités.

Le droit suisse contient en outre des dispositions pouvant être invoquées en cas de discrimination par des particuliers. Une protection relativement étendue contre la discrimination fondée sur le genre est notamment garantie dans les rapports de travail avec la loi sur l'égalité, le devoir d'assistance (art. 328 CO) et la protection contre le licenciement (art. 336 CO). Pour ce qui concerne le traitement des données, le principe de traitement non discriminatoire est certes inscrit dans la loi sur la protection des données, mais celle-ci ne constitue pas une base solide et convaincante en la matière. D'autres prescriptions permettent d'identifier d'éventuels effets discriminatoires, sans offrir de réelle protection contre la discrimination. C'est le cas p. ex. des dispositions relatives à la décision individuelle automatisée (art. 21 LPD). Les clauses contractuelles discriminatoires et les refus de conclure un contrat sur la base d'une discrimination peuvent en revanche être considérés comme des violations des art. 19 et 20 CO, et de l'art. 2, al. 2 CC. Enfin, la protection directe contre certaines formes de discrimination est inscrite dans le droit pénal : l'art. 261bis CP peut ainsi être invoqué en cas de refus de prestation par un système de décision algorithmique.

---

190 NAGUIB, p. 996

Une interdiction générale de la discrimination à l'égard des particuliers pourrait éventuellement être déduite de l'art. 28 CC (protection de la personnalité). Il reste toutefois incertain à quelles formes de discrimination algorithmique elle peut s'appliquer.

En conclusion, les discriminations dans les relations entre particuliers sont difficiles à circonscrire, qu'il s'agisse de formes de discriminations traditionnelles ou algorithmiques. Pour être efficace, la lutte contre la discrimination algorithmique ne doit pas être envisagée sous le seul angle de l'adaptation des réglementations existantes, mais d'une manière globale et indépendante de la technologie utilisée.

8. Mesures envisageables

Le risque de discrimination par les systèmes algorithmiques ne peut être réduit que si toutes les parties prenantes, du développement à l'application, sont conscientes du problème. Il est donc essentiel de les sensibiliser et de les informer des solutions possibles (littératie de l'IA, littératie des données, littératie numérique). Cela nécessite une compréhension générale du fonctionnement technique des systèmes algorithmiques et des dispositions légales relatives à la protection contre la discrimination[191], ainsi qu'une recherche interdisciplinaire approfondie.

Outre l'information des développeurs et développeuses et des utilisateurs et utilisatrices de systèmes d'IA ainsi que des personnes concernées, il convient de prendre diverses mesures pour détecter et prévenir la discrimination algorithmique. Dans les pages qui suivent, une distinction est faite entre les mesures qui peuvent être mises en œuvre dans le cadre du droit en vigueur (ch. 8.1) et celles qui nécessitent une intervention du législateur (ch. 8.2 et 8.3).

8.1. Dans le cadre du droit en vigueur

8.1.1. Application des normes existantes

Les considérations qui précèdent (voir les ch. 3 à 6) ont montré que le droit suisse connaît certes différentes dispositions qui permettraient d'identifier et de prévenir les discriminations, mais que les personnes concernées font un usage (trop) limité des possibilités existantes, car il peut être difficile pour elles de les faire valoir. À cela s'ajoute le fait que les tribunaux interprètent et appliquent parfois les normes existantes avec une certaine réserve. Cela vaut en particulier pour la protection de la personnalité en droit civil (voir ch. 5.1).

C'est dans l'interprétation et l'application du principe régissant la protection de la personnalité en droit civil (art. 28 CC) que réside le plus grand potentiel d'amélioration de la protection contre la discrimination commise par des particuliers. Depuis longtemps, la doctrine qualifie à juste titre les discriminations d'atteintes à la personnalité (voir ch. 5.1). Le nombre d'auteurs et d'autrices qui défendent cette position est certes limité, mais, pour autant que l'on puisse en juger, cette interprétation de l'art. 28 CC ne souffre d'aucune contestation expresse. Forts d'une base suffisante, les tribunaux pourraient donc suivre la doctrine en avançant des motifs valables, et le Tribunal fédéral pourrait, par sa jurisprudence, créer au titre de l'art. 28 CC une protection complète contre la discrimination émanant de particuliers. Il semble toutefois douteux que le Tribunal fédéral ose franchir le pas, car cela reviendrait

[191] BRAUN BINDER / OBRECHT, Künstliche Intelligenz, p. 38

à développer le droit sur un mode en principe réservé au législateur. S'il existe donc un potentiel pour faire évoluer le droit actuel vers une protection complète contre la discrimination en référence à l'art. 28 CC, les chances que cela se concrétise sont plutôt faibles.

Garantir une protection contre la discrimination au titre de l'art. 28 CC aurait en outre pour effet de permettre aux associations et autres organisations de lutter contre la discrimination en exerçant le droit d'action qui leur est reconnu (art. 89 CPC). Ce serait une avancée bienvenue, car la discrimination algorithmique touche régulièrement de nombreuses personnes qui ne disposent souvent ni des ressources financières ni du temps nécessaires pour intenter une action en justice et engager une procédure[192]. Le droit d'action des associations prévu à l'art. 89 CPC permettrait de faire valoir collectivement des prétentions, et ce serait particulièrement judicieux dans le cas de la discrimination algorithmique, vu le nombre de personnes concernées.

8.1.2. Travaux de recherche supplémentaires

Indépendamment de l'évolution éventuelle du droit existant, il est apparu que la protection juridique contre la discrimination due aux systèmes algorithmiques requiert de vastes recherches. De nombreuses questions apparemment évidentes au premier regard s'avèrent étonnamment complexes une fois qu'on les examine de plus près.

La discrimination algorithmique ne soulève pas seulement de nouveaux défis pour ce qui concerne la législation et l'application du droit ; elle pose également des questions de doctrine juridique[193]. Parmi celles-ci figure notamment celle du traitement juridique de la discrimination intersectionnelle (voir ch. 2.3.2), bien qu'on ignore si les systèmes d'IA sont particulièrement sujets à ce type de discrimination. Certes, de nombreux résultats des systèmes d'IA reposent sur une multitude de critères pondérés différemment (voir ch. 2.4), qui ne sont pas nécessairement des caractéristiques protégées ou des variables corrélées à une caractéristique protégée. Même s'il est tout à fait concevable que l'utilisation de systèmes d'IA conduise plus souvent à une discrimination intersectionnelle, il existe également des cas de figure où les systèmes d'IA établissent des corrélations sur la base de critères non pertinents au regard du droit de la discrimination, sans que ceux-ci ne servent de substituts à des caractéristiques protégées[194]. Vu le manque de clarté dans la manière dont la doctrine juridique traite la discrimination intersectionnelle et l'incertitude quant aux points de rattachement et aux corrélations dans les systèmes d'IA, il apparaît nécessaire d'approfondir les recherches dans ce domaine.

Il en va de même pour la distinction entre discrimination directe et discrimination indirecte en cas d'utilisation de systèmes algorithmiques. Si la forme de discrimination n'a pas d'importance, les deux étant interdites, des différences peuvent toutefois apparaître au niveau de la preuve. À suivre la conception proposée dans la présente analyse (voir ch. 2.4), la question se pose en effet de savoir à quelles exigences il faut soumettre l'administration de la preuve dans le cas d'une discrimination par procuration. Il est probable qu'on exige toujours des preuves statistiques et qu'une discrimination ne soit admise que si le système algorithmique désavantage, au final, une proportion significative de personnes présentant une caractéristique protégée, par rapport aux autres. La question du motif invoqué doit également être approfondie. Pour des raisons de sécurité juridique, il serait souhaitable de pouvoir

---

192 Voir à ce sujet aussi MÜLLER A., p. 28
193 Pour une synthèse complète, voir MÜLLER J.-L., p. 205 ss
194 MÜLLER J.-L., p. 235 s.

identifier des catégories d'objectifs pouvant justifier une discrimination pour des raisons objectives. Toutefois, l'état actuel de la recherche ne permet pas de déterminer si cela est possible compte tenu de la diversité des cas de figure possibles.

8.2. Mesures législatives

8.2.1. Décisions partiellement automatisées

En pratique, les systèmes algorithmiques sont souvent utilisés pour soutenir les décisions humaines (on parle de *decision support systems*, ou « systèmes d'aide à la décision »). Or la réglementation actuelle en matière de protection des données ne couvre que les décisions entièrement automatisées (voir ch. 3.3.1, let. b), et il n'existe ni obligation particulière d'informer ni droit particulier d'accès à l'information pour les personnes concernées lorsque les décisions sont partiellement automatisées. Cette différenciation semble discutable au regard de la protection contre la discrimination algorithmique, car le risque de discrimination ne dépend pas du degré d'automatisation (voir ch. 2.1.3). Il pourrait donc être judicieux d'étendre le devoir d'informer (art. 21 LPD) et le droit d'accès correspondant (art. 25, al. 2, let. f, LPD) prévus pour les décisions entièrement automatisées aux décisions partiellement automatisées. Il conviendrait toutefois de distinguer entre les obligations incombant aux organes publics et celles incombant aux particuliers et, dans ce dernier cas, entre le devoir d'informer et le droit d'accès.

Dans le cas des décisions prises par des organes publics, l'obligation de qualifier comme telle la décision individuelle automatisée est un moyen de garantir le droit d'être entendu des personnes concernées[195] et la motivation en bonne et due forme des décisions. Comme indiqué précédemment (voir ch. 4.1), le recours à des systèmes algorithmiques peut nuire à la mise en œuvre de l'obligation de motiver les décisions, ce qui peut à son tour augmenter le risque de discrimination algorithmique ou, à tout le moins, le risque qu'une discrimination ne soit pas identifiée. Cela vaut de manière générale pour les décisions, y compris celles qui ne sont pas entièrement automatisées. Il est donc judicieux d'étendre le devoir d'informer et le droit d'accès correspondant aux décisions partiellement automatisées afin de tenir compte effectivement de la teneur du droit d'être entendu dans le cadre de telles décisions[196].

La situation est fondamentalement différente lorsqu'il s'agit de décisions prises par des particuliers. Ceux-ci ne sont pas soumis à l'obligation d'entendre l'autre partie ni tenus de motiver leurs décisions. Dans ce contexte, on peut se demander pourquoi le droit en vigueur les oblige à informer les personnes concernées qu'une décision les concernant a été prise de manière entièrement automatisée. Cela ne s'explique probablement que par un scepticisme généralisé quant à la qualité de telles décisions[197]. Avant d'étendre le devoir d'informer en cas de décision entièrement automatisée aux décisions partiellement automatisées, il conviendrait donc d'examiner si le devoir d'informer existant atteint son objectif. Dans un premier temps, il faudrait vérifier si les personnes concernées prennent réellement connaissance de ces informations et les jugent pertinentes et utiles. Dans un deuxième temps, il faudrait examiner si les coûts liés au devoir d'informer sont, pour les entreprises, proportionnés aux avantages que l'information apporte aux personnes concernées. Là encore, on peut en douter, d'autant que la mise en œuvre du devoir d'informer requiert une analyse approfondie de tous les processus de traitement au sein d'une entreprise et des mesures nécessaires pour transmettre l'information, ce qui peut entraîner

---

195 Sur l'ensemble de la problématique, voir BRAUN BINDER, Automatisierte Entscheidungen, p. 32 s.
196 C'est également l'avis de l'OFJ dans son Analyse juridique de base, 2024, p. 54 s.
197 Voir à ce sujet SUTER, p. 109 ss ; THOUVENIN / FRÜH, p. 9

des coûts élevés. En ce qui concerne la question qui nous occupe, le simple fait d'informer sur la nature d'une décision ne constitue pas en soi une protection contre la discrimination algorithmique (ou d'une autre nature). De même, l'existence d'une discrimination ne dépend pas du fait qu'une décision a été prise uniquement ou essentiellement par un être humain, ou de manière entièrement ou essentiellement automatisée.

Il est bien plus important que les personnes concernées disposent d'un droit d'accès leur permettant, en cas de suspicion de discrimination, d'obtenir des informations complètes sur les données personnelles traitées et la logique sur laquelle repose une décision entièrement ou essentiellement automatisée. À l'instar du droit général d'accès prévu par la législation sur la protection des données, ce droit d'accès spécial pourrait être exercé en tout temps par les personnes concernées, sans qu'elles aient à s'en justifier[198]. Ainsi, les personnes concernées pourraient obtenir des entreprises les informations dont elles ont besoin pour vérifier la décision et faire valoir leurs droits, le cas échéant. Spécifique aux décisions essentiellement ou entièrement automatisées, un tel droit d'accès serait autrement plus efficace pour protéger les personnes concernées contre la discrimination algorithmique qu'une extension du devoir d'informer, tout en occasionnant des coûts nettement moins élevés pour les entreprises concernées.

En revanche, il est clair que le droit à un contrôle effectué par un être humain, prévu par la loi (*de lege lata*) pour les systèmes entièrement automatisés, ne peut être étendu aux décisions essentiellement automatisées, puisque l'être humain est déjà associé au processus décisionnel. Toutefois, pour les organes publics, les prétentions découlant du droit d'être entendu dans le cadre de décisions assistées par l'IA doivent être motivées, ce qui garantit la possibilité de les vérifier.

8.2.2. Loi générale sur l'égalité de traitement

a. Proposition

Une loi générale sur l'égalité de traitement (ou loi antidiscrimination) interdisant la discrimination directe et indirecte non seulement commise par les autorités publiques, mais aussi par les acteurs privés, permettrait d'assurer la protection la plus complète contre la discrimination algorithmique ou autre.

Une telle loi pourrait s'appuyer sur les dispositions existantes, en particulier celles de la LEg et de certaines parties de la LHand, qui couvrent déjà certaines formes de discrimination. Plusieurs États ont déjà adopté des dispositions similaires[199]. Une loi générale suisse sur l'égalité ou contre la discrimination pourrait s'en inspirer[200].

Afin d'accroître la sécurité juridique et de faciliter l'application de la loi, il serait judicieux de prévoir dans une telle loi des dispositions régissant la preuve de la discrimination. Il devrait suffire que la personne concernée rende la discrimination vraisemblable pour que l'organe public ou l'entreprise doive apporter la preuve contraire établissant l'absence de discrimination. Les procédures devraient être gratuites pour les personnes concernées, sous réserve d'abus dans l'exercice des droits. Pour des raisons

---

198 Pour ne citer qu'un commentaire, OFK DSG-STENGEL / STÄUBLE, art. 25, N°s 6 s

199 Par exemple en Allemagne : Allgemeines Gleichbehandlungsgesetz (AGG ; loi générale sur l'égalité de traitement) du 14 août 2006 ; ou en Autriche : Bundesgesetz über die Gleichbehandlung (GlBG ; loi fédérale sur l'égalité de traitement) du 17 juin 2004

200 Voir également PÄRLI / NAGUIB, p. 43 ss

de sécurité juridique, il pourrait également être judicieux de prévoir une liste non exhaustive de motifs légitimes.

En matière d'application du droit, il conviendrait de prévoir un droit d'action collective pour les organisations qui s'occupent spécifiquement de discrimination, en s'inspirant de la réglementation prévue dans la LHand. Celle-ci dispose que, dans certains domaines, les organisations d'importance nationale d'aide aux personnes handicapées existant depuis dix ans au moins peuvent faire valoir des prétentions juridiques en cas d'inégalité touchant un grand nombre de personnes handicapées (art. 9, al. 1, LHand).

Le champ d'application d'une loi générale sur l'égalité de traitement pourrait être étendu ou limité à certains domaines comme le travail, la formation ou la consommation, si la volonté politique d'adopter une loi générale venait à faire défaut. Il serait également envisageable, en vue d'une adaptation progressive du droit en vigueur, d'adopter des lois spéciales contre la discrimination fondée sur certaines caractéristiques protégées, par exemple une loi contre la discrimination raciale.

b. Avantages et inconvénients

Une loi générale sur l'égalité de traitement (ou antidiscrimination) présenterait l'avantage majeur de couvrir systématiquement et uniformément toutes les formes de discrimination. Cela renforcerait non seulement la sécurité juridique, mais pourrait également produire un signal important en établissant une interdiction générale de la discrimination, y compris celle exercée par des particuliers.

Une interdiction générale de la discrimination empiéterait certes fortement sur l'autonomie privée, en particulier sur la liberté contractuelle, mais une telle ingérence semble justifiée vu l'importance du problème. Compte tenu des intérêts concernés, une loi générale sur l'égalité de traitement (ou antidiscrimination) devrait constituer une réglementation équilibrée qui protège efficacement les personnes physiques contre la discrimination sans empiéter de manière excessive sur les rapports de droit privé.

Les effets juridiques que l'on peut en attendre relèveraient principalement d'une action défensive, en particulier un droit à la cessation et à la suppression de l'atteinte ainsi qu'un droit à la constatation d'une discrimination. L'introduction de prétentions réparatrices, en particulier le droit à des dommages-intérêts et à une réparation morale, serait également envisageable.

Les effets juridiques pourraient en outre être subordonnés à la faute. Un comportement négligent pourrait suffire pour faire aboutir une action défensive, tandis que les prétentions réparatrices ne s'appliqueraient qu'en cas de comportement intentionnel. Cette différenciation refléterait l'objectif préventif de la protection contre la discrimination et éviterait des risques de responsabilité excessifs pour les particuliers. Une telle approche garantirait également le respect du principe de proportionnalité et encouragerait le développement et le contrôle minutieux des systèmes algorithmiques.

Parallèlement, les entreprises devraient être soutenues dans la mise en œuvre de processus non discriminatoires au moyen d'aides pratiques, telles que des directives techniques pour le développement et l'utilisation de systèmes algorithmiques.

Un service de conseil pourrait en outre informer les entreprises et les personnes concernées sur les exigences légales, les mesures préventives et les conséquences juridiques potentielles. Une telle institution contribuerait à réduire les points d'incertitude, à clarifier les malentendus en amont et à

soutenir les entreprises dans le respect des règles. Il faudrait veiller à ce que ce service dispose de ressources humaines et financières suffisantes.

8.2.3. Nouveau principe relatif au traitement des données

a. Proposition

Une autre possibilité consiste à introduire dans la LPD le principe selon lequel nul ne doit être discriminé en raison du traitement de ses données personnelles. Cette nouveauté permettrait de couvrir les discriminations directes, car celles-ci sont toujours liées au traitement de données personnelles.

L'idée d'un traitement loyal et non discriminatoire des données découle déjà du but de la LPD, qui vise à protéger la personnalité et les droits fondamentaux des personnes dont les données personnelles font l'objet d'un traitement (art. 1 LPD). Le principe de la bonne foi (art. 6, al. 2, LPD) implique également l'obligation de traiter les données personnelles avec respect et sans discrimination.

b. Avantages et inconvénients

L'un des avantages d'une réglementation dans la LPD serait qu'elle ne nécessiterait qu'une modification ponctuelle, ce qui serait relativement simple à mettre en œuvre sur le plan législatif. En outre, la LPD contient déjà des mécanismes d'application efficaces et il existe, en la personne du PFPDT, une autorité qui peut agir de sa propre initiative ou sur signalement des personnes concernées en cas de violation de la loi.

Le droit d'accès prévu par la législation sur la protection des données (art. 25 LPD) constitue en outre un moyen efficace d'obtenir des informations sur le traitement des données. Ce levier est utile pour faire respecter la loi. Le PFPDT a déjà affirmé que la LPD est « directement applicable à l'IA » lorsque des systèmes d'IA traitent des données personnelles[201]. Comme le montre l'analyse des dispositions en vigueur, la LPD contient déjà des éléments permettant de protéger les personnes contre la discrimination émanant de particuliers (voir ch. 3.3). Un principe distinct régissant le traitement des données permettrait de clarifier la situation et d'assurer la sécurité juridique.

Une disposition spécifique dans la LPD présenterait toutefois l'inconvénient majeur d'être étrangère au champ d'application de cette loi et de ne pas s'inscrire dans le droit de la protection des données : elle irait au-delà de la finalité première de la LPD, à savoir la protection de la sphère privée et un certain contrôle des personnes sur le traitement des données les concernant. Y intégrer une protection contre la discrimination reviendrait à élargir considérablement le champ de la LPD et risquerait d'en diluer le but initial[202].

Les capacités limitées du PFPDT constituent un autre inconvénient. L'autorité est déjà fortement sollicitée et une compétence supplémentaire pourrait nuire à l'accomplissement de ses tâches principales. À cela s'ajoute le fait que le PFPDT est spécialisé dans la protection des personnes contre les risques liés au traitement des données personnelles et non dans l'évaluation des cas de discrimination.

---

201 Voir la communication du PFPDT du 8 mai 2025 : Update – La loi actuelle sur la protection des données est directement applicable à l'IA (https://www.edoeb.admin.ch/fr/update-loi-actuelle-applicable-ia)

202 Voir également HACKER, p. 1172

8.3. Mesures particulières de droit public

La relation de confiance particulière qui existe entre l'État et la population revêt une importance capitale dans le cadre de l'utilisation de l'IA par les autorités. Cette relation repose sur la présomption que les pouvoirs publics respectent les principes de l'État de droit, s'engagent en faveur de l'intérêt général, et respectent et protègent les droits fondamentaux de toutes les personnes. La confiance dans les institutions étatiques repose sur des fondements structurels et normatifs. L'administration dispose de pouvoirs étendus qui peuvent interférer profondément avec la situation juridique des personnes. L'utilisation de l'IA par les autorités publiques requiert par conséquent des mesures de protection particulières et supplémentaires, car une perte de confiance peut non seulement nuire à l'acceptation des nouvelles technologies, mais aussi remettre globalement en question la légitimité des décisions étatiques.

Dans ce contexte, il s'agit avant tout de créer davantage de transparence sur l'utilisation des systèmes algorithmiques dans l'administration (ch. 8.3.1) et/ou de limiter leur utilisation à des systèmes qui peuvent garantir la traçabilité (ch. 8.3.2). Certaines mesures concrètes peuvent en outre être formulées à l'intention des services administratifs qui utilisent des systèmes d'IA (ch. 8.3.3). Il convient enfin de se demander si certains domaines d'application ne se prêtent pas à une automatisation complète (ch. 8.3.4).

8.3.1. Création de registres publics

Plusieurs acteurs considèrent que la création de registres publics des systèmes algorithmiques utilisés dans l'administration fédérale est une mesure permettant de prévenir la discrimination[203]. Il ne serait pas nécessaire de répertorier tous les systèmes algorithmiques utilisés dans l'administration, mais uniquement ceux qui présentent des risques accrus pour les droits fondamentaux des personnes concernées ou qui sont susceptibles d'avoir d'autres répercussions sur les droits et obligations des justiciables[204].

Bien que n'étant pas à même d'empêcher la discrimination par des systèmes algorithmiques[205], de tels registres créent une condition préalable importante à une transparence et un contrôle social accrus : les personnes concernées peuvent savoir si un système algorithmique s'applique à elles et dans quel contexte. Sachant cela, il est possible d'identifier des effets discriminatoires et de prendre des mesures juridiques et politiques, par exemple en contestant des décisions ou en exigeant un examen indépendant du système d'IA utilisé.

En vertu du principe de légalité (art. 164, al. 1, Cst.), l'obligation de tenir un tel registre ainsi que les informations fondamentales concernant l'organisme responsable desdits registres et les systèmes couverts par ces registres devraient être réglées au niveau de la loi[206]. Les dispositions légales pourraient

---

[203] BRAUN BINDER et al., Studie KI, p. 50 ; BRAUN BINDER / OBRECHT, Transparenz, p. 1075 ss ; THOUVENIN et al., Positionspapier 2021, p. 3 ; AlgorithmWatch CH, Utilisation responsable des algorithmes et de l'IA dans l'administration publique, 25.2.2022 (https://algorithmwatch.ch/de/diskriminierung-durch-ki-in-der-personalauswahl/)

[204] Ainsi, les systèmes algorithmiques utilisés dans le cadre d'activités administratives auxiliaires, tels que les programmes de traduction basés sur l'IA, ne devraient pas être répertoriés (voir BRAUN BINDER /OBRECHT, Transparenz, p. 5).

[205] BERENDT, p. 47 s. ; OBRECHT, Questions au féminin, p. 79 s.

[206] Sur l'ensemble de cette question, voir BRAUN BINDER / OBRECHT, Transparenz, p. 1079 s.

figurer dans la LPD ou dans la Loi sur l'utilisation des moyens électroniques pour l'exécution des tâches des autorités (LMETA)[207]. Il faudrait préciser, par voie d'ordonnance, les informations de base et les références susceptibles de fournir des informations complémentaires ou des rapports d'audit[208].

8.3.2. Utilisation limitée à des systèmes offrant une traçabilité

La traçabilité requise des décisions prises avec l'aide de systèmes algorithmiques découle de l'obligation constitutionnelle de motiver les décisions[209]. Comme indiqué plus haut, la motivation des décisions est également un élément permettant de prévenir la discrimination algorithmique (voir ch. 4.1). L'impossibilité de garantir cette traçabilité crée un problème de constitutionnalité, car, dans ce cas, la personne chargée du dossier ou l'instance de recours ne disposent d'aucune base pour contrôler la décision ou la faire contrôler par un tiers.

Dans ce contexte, il est recommandé d'exclure systématiquement l'utilisation, dans l'administration publique, de systèmes d'IA ne permettant pas une traçabilité. Le législateur pourrait ainsi définir expressément que seuls les systèmes d'IA dont l'algorithme est public et retraçable par le collaborateur ou la collaboratrice en charge du dossier, la personne concernée et l'instance de recours peuvent être utilisés dans l'administration publique. Cela peut contribuer à rendre effectifs le contrôle interne et la fonction de moyens de preuve découlant de l'obligation de motiver.

8.3.3. Recommandations à l'intention des autorités administratives qui utilisent des systèmes d'IA

Les unités administratives qui souhaitent utiliser des systèmes d'IA dans le cadre de procédures administratives peuvent prendre différentes mesures pour réduire au maximum le risque de discrimination. Cela va de l'accompagnement du travail de programmation – généralement externalisé à des entreprises privées – afin d'éviter tout code de programmation discriminatoire[210] – à la formation et l'obligation pour les employés et employées d'examiner de manière critique les résultats des systèmes d'IA et, si nécessaire, de remplacer la décision correspondante par une décision humaine[211], en passant par l'utilisation ou la mise à disposition de jeux de données d'entraînement diversifiés, actuels et représentatifs[212].

Les contrôles aléatoires constituent une mesure particulièrement indiquée pour identifier et lutter contre les risques de discrimination liés aux décisions assistées par l'IA[213]. Ces contrôles ne devraient pas seulement être effectués par les tribunaux dans le cadre de procédures de recours a posteriori ; il faudrait également en faire de manière systématique au cours des processus administratifs. Ainsi, un certain

---

207 Sur l'ensemble de cette question, voir BRAUN BINDER / OBRECHT, Transparenz, p. 1078 s.

208 Pour un aperçu récapitulatif des informations pertinentes relatives aux systèmes algorithmiques dans un registre public, voir BRAUN BINDER / OBRECHT, WP Transparenz, p. 5 s.

209 L'obligation de motiver les décisions découle du droit d'être entendu prévu à l'art. 29, al. 2, Cst. Elle est concrétisée au niveau de la loi à l'art. 35, al. 1, PA. Pour les exceptions à cette règle, voir l'art. 35, al. 3, PA.

210 MARTINI, Blackbox, p. 47 ss

211 BUCHHOLTZ / SCHEFFEL-KAIN, p. 616

212 Voir l'art. 10 du règlement UE

213 BUCHHOLTZ / SCHEFFEL-KAIN, P. 617

nombre d'échantillons sélectionnés de manière aléatoire pourraient être soumis à un contrôle humain là où les procédures administratives recourent à des systèmes d'IA[214].

8.3.4. Interdiction des décisions entièrement automatisées dans les domaines sensibles

Une autre mesure de lutte contre la discrimination algorithmique consiste à limiter l'utilisation de systèmes entièrement automatisés. Il est envisageable d'y renoncer dans les domaines sensibles[215]. Cela signifie que seuls des systèmes d'aide à la décision ou de vérification des décisions pourraient être utilisés dans ces domaines. Si une analyse d'impact sur la protection des données devait montrer que le risque de discrimination est particulièrement élevé dans un domaine donné, il semblerait judicieux de réserver les décisions dans ce domaine aux êtres humains et de n'utiliser les systèmes d'IA qu'à titre d'aide à la décision.

9. Conclusion

La protection juridique contre la discrimination – qu'elle soit algorithmique ou classique – est insuffisante en Suisse. La Constitution fédérale prévoit certes une interdiction générale de la discrimination (art. 8, al. 2, Cst.), mais celle-ci ne s'applique qu'aux acteurs étatiques et non aux particuliers. Si le droit public contient diverses dispositions antidiscrimination, leur application à des systèmes algorithmiques complexes et souvent opaques n'est pas évidente. En droit privé, la protection est fragmentaire. Seule la protection des employés et employées contre la discrimination fondée sur le genre est relativement bien développée. L'application du droit pose en outre des difficultés, tant en droit public qu'en droit privé. Ces problèmes existent déjà dans les formes conventionnelles de discrimination, mais la discrimination algorithmique ajoute d'autres défis, notamment en ce qui concerne la détection et la preuve de la discrimination par procuration, très souvent utilisés par les systèmes d'IA.

Les mesures recommandées dans le présent avis de droit ne se limitent pas à la discrimination algorithmique, d'autant plus que la protection contre les formes conventionnelles de discrimination est déjà lacunaire et relativement peu développée. Pour garantir une protection effective, il est nécessaire de prévoir des règles générales et indépendantes de la technologie utilisée qui couvrent la discrimination exercée tant par les autorités que par les particuliers, idéalement sous la forme d'une loi générale sur l'égalité de traitement ou d'une loi antidiscrimination. Une telle loi permettrait de jeter les bases nécessaires pour lutter contre la discrimination. La manière dont ses dispositions générales devraient s'appliquer aux systèmes algorithmiques devrait s'appuyer sur les résultats de recherches juridiques encore en voie d'élaboration, lesquelles devraient, à leur tour, faire fond sur les acquis des recherches techniques et sociologiques, de manière à garantir une protection suffisante contre la discrimination algorithmique.

Remerciements

Les autrices et l'auteur remercient chaleureusement Grace Wittmer (MLaw) et Anna Kuhn (MLaw), toutes deux anciennes collaboratrices du centre de recherche e-PIAF, ainsi que Lars Ruflin,

---

214 HERMSTRÜWER, p. 497

215 BRAUN BINDER / OBRECHT, Künstliche Intelligenz, p. 38

collaborateur du centre de recherche e-PIAF, pour leur soutien actif dans la recherche bibliographique et l'analyse des textes législatifs pour le présent avis de droit.

Références bibliographiques

Remarque concernant les sources en ligne : les pages citées ont été consultées pour la dernière fois le 14 juillet 2025.

ALON-BARKAT SAAR / BUSUIOC MADALINA, Human-AI Interactions in Public Sector Decision Making: «Automation Bias» and «Selective Adherence» to Algorithmic Advice, JPART 2023, p. 153 ss

ARIOLI KATHRIN, Die Rechtsfigur der indirekten Diskriminierung, AJP 1993, p. 1327 ss

BATACHE DJAMILA, Künstliche Intelligenz in der Medizin aus haftungsrechtlicher Perspektive, thèse Univ. Bâle 2023, Zurich 2023

BELSER EVA MARIA / WALDMANN BERNHARD, Grundrechte II: Die einzelnen Grundrechte, 2e édition, Zurich / Bâle / Genève 2021

BERENDT BETTINA, The AI Act proposal: Towards the next transparency fallacy? Why AI regulation should be based on principles based on how algorithmic discrimination works. Dans : Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz / Frauke Rostalski (dir.), Künstliche Intelligenz – Wie gelingt eine vertrauenswürdige Verwendung in Deutschland und Europa?, Tübingen 2022, p. 31 ss

BIAGGINI GIOVANNI, BV Kommentar: Bundesverfassung der Schweizerischen Eidgenossenschaft, 2e édition, Zurich 2017 (citation : BV Komm-BIAGGINI)

BLONSKI DOMINIKA, Kommentierung des Art. 22 DSG. Dans : Baeriswyl Bruno / Pärli Kurt / Blonski Dominika (dir.), Datenschutzgesetz (DSG): Bundesgesetz vom 25. September 2020 über den Datenschutz (DSG), 2e édition, Berne 2023 (citation : SHK DSG-BLONSKI)

BRAUN BINDER NADJA, Automatisierte Entscheidungen: Perspektive Datenschutzrecht und öffentliche Verwaltung, SZW-RSDA 2020, p. 27 ss (citation : BRAUN BINDER, Automatisierte Entscheidungen)

BRAUN BINDER NADJA, Als Verfügungen gelten Anordnungen der Maschinen im Einzelfall... – Dystopie oder künftiger Verwaltungsalltag?, ZSR-RDS I 2020, p. 253 ss (citation : BRAUN BINDER, Dystopie)

BRAUN BINDER NADJA, Künstliche Intelligenz und automatisierte Entscheidungen in der öffentlichen Verwaltung, SJZ-SRJ 2019, p. 467 ss (citation : BRAUN BINDER, Künstliche Intelligenz)

BRAUN BINDER NADJA / EGLI CATHERINE, Kommentierung des Art. 10 KI-VO. Dans : Martini Mario / Wendehorst Christiane (dir.), Kommentar zur KI-VO: Verordnung über künstliche Intelligenz, Munich 2024 (citation : KI-VO-BRAUN BINDER / EGLI)

BRAUN BINDER NADJA / EGLI CATHERINE, Kommentierung des Art. 9 KI-VO. Dans : Martini Mario / Wendehorst Christiane (dir.), Kommentar zur KI-VO: Verordnung über künstliche Intelligenz, Munich 2024 (citation : KI-VO-BRAUN BINDER / EGLI)

BRAUN BINDER NADJA et al., Einsatz Künstlicher Intelligenz in der Verwaltung: rechtliche unethische Fragen, Schlussbericht vom 28. Februar 2021 zum Vorprojekt IP6.4 (https://www.zh.ch/de/news-

uebersicht/medienmitteilungen/2021/04/kuenstliche-intelligenz-in-der-verwaltung-braucht-klare-leitlini.html, citation : BRAUN BINDER et al., Studie KI)

BRAUN BINDER NADJA / KUNZ ELIANE / OBRECHT LILIANE, Maschinelle Gesichtserkennung im öffentlichen Raum, sui generis 2022, p. 53 ss

BRAUN BINDER NADJA / OBRECHT LILIANE, Transparenz über den staatlichen Einsatz algorithmischer Entscheidungssysteme: Rechtliche Erwägungen ausgehend von Forderungen nach öffentlichen Verzeichnissen, AJP 2024, p. 1069 ss (citation : BRAUN BINDER / OBRECHT, Transparenz)

BRAUN BINDER NADJA / OBRECHT LILIANE, Transparenz durch öffentliches Verzeichnis, White Paper de juin 2024 (https://ius.unibas.ch/fileadmin/user_upload/ius/03_Forschung/10_e-PIAF/WhitePaper_Verzeichnis_Juni24.pdf, citation : BRAUN BINDER / OBRECHT, WP Transparenz)

BRAUN BINDER NADJA / OBRECHT LILIANE, Die Begründung von Verfügungen beim Einsatz algorithmischer Systeme, SJZ-SRJ 2024, p. 707 ss (citation : BRAUN BINDER / OBRECHT, Begründung)

BRAUN BINDER NADJA / OBRECHT LILIANE, Begründungspflicht, White Paper de juin 2024 (https://ius.unibas.ch/fileadmin/user_upload/ius/03_Forschung/10_e-PIAF/WhitePaper_Begruendung_Juni24.pdf, citation : BRAUN BINDER / OBRECHT, WP Begründung)

BRAUN BINDER NADJA / OBRECHT LILIANE, Algorithmisch überprüfte Steuererklärungen im ordentlichen gemischten Veranlagungsverfahren, zsis) 1/2023 du 30.3.2023 (citation : BRAUN BINDER / OBRECHT, Steuererklärungen)

BRAUN BINDER NADJA / OBRECHT LILIANE, Künstliche Intelligenz in der Verwaltung – Möglichkeiten, rechtliche Rahmenbedingungen und Grenzen. Dans : SVVOR-ASDPO (dir.), Annuaire 2021/2022, Berne 2023, p. 21 ss (citation : BRAUN BINDER / OBRECHT, Künstliche Intelligenz)

BUCHHOLTZ GABRIELE / SCHEFFEL-KAIN MARTIN, Algorithmen und Proxy Discrimination in der Verwaltung: Vorschläge zur Wahrung digitaler Gleichheit, NVwZ 2022, p. 612 ss

BÜCHLER ANDREA, dans : Kren Kostkiewicz Jolanta / Wolf Stephan / Amstutz Marc / Fankhauser Roland (dir.), ZGB Kommentar Schweizerisches Zivilgesetzbuch, 4e édition, Zurich 2021 (citation : OFK ZGB-BÜCHLER)

CHRISTEN MARKUS et al., Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz. Dans : publications TA-SWISS, TA 72/2020 (https://www.ta-swiss.ch/kuenstliche-intelligenz, citation : CHRISTEN et al., TA-SWISS 72/2020)

DIGGELMANN OLIVER et al., Der Diskriminierungsbegriff im öffentlichen und grundrechtlichen Diskurs, ZSR-RDS I 2022, p. 211 ss

DZIDA BORIS / GROH NAEMI, Diskriminierung nach dem AGG beim Einsatz von Algorithmen im Bewerbungsverfahren, NJW 2018, p. 1917 ss

FANKHAUSER-FEITKNECHT VIVIAN, Uno-Frauenkonvention gilt auch für Schweizer Gerichte, Plädoyer 2009, p. 25

GLATTHAAR MATTHIAS / SCHRÖDER ANNIKA, Kommentierung des Art. 22 DSG. Dans : Blechta Gabor P. / Vasella David (dir.), Datenschutzgesetz/Öffentlichkeitsgesetz, Basler Kommentar, 4e édition, Bâle 2023 (citation : BSK DSG-GLATTHAAR / SCHRÖDER)

GÖKSU TARKAN, Rassendiskriminierung beim Vertragsabschluss als Persönlichkeitsverletzung, thèse Univ. Fribourg 2003, Fribourg 2003 (citation : GÖKSU, Rassendiskriminierung)

GÖKSU TARKAN, Drittwirkung der Grundrechte im Bereich des Persönlichkeitsschutzes, SJZ-RSJ 2002, p. 89 ss (citation : GÖKSU, Drittwirkung)

GORDON CLARA-ANN / LUTZ TANJA, Haftung für automatisierte Entscheidungen – Herausforderungen in der Praxis, SZW-RSDA 2020, p. 53 ss

HACKER PHILIPP, Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law, CMLR 2018, p. 1143 ss

HÄFELIN ULRICH / HALLER WALTER / KELLER HELEN / THURNHERR DANIELA, Schweizerisches Bundesstaatsrecht, 11e édition, Zurich / Genève 2024

HÄFELIN ULRICH / MÜLLER GEORG / UHLMANN FELIX, Allgemeines Verwaltungsrecht, 8e édition, Zurich / St-Gall 2020

HAUSAMMANN CHRISTINA, Instruments contre la discrimination en droit suisse – Un aperçu, sur mandat du Bureau fédéral de l'égalité pour les personnes handicapées et du Service de lutte contre le racisme, Berne 2008

HENSELER SIMON / VASELLA DAVID, Kommentierung des Art. 21 DSG. Dans : Blechta Gabor P. / Vasella David (dir.), Datenschutzgesetz/Öffentlichkeitsgesetz, Basler Kommentar, 4e édition, Bâle 2023 (citation : BSK DSG-HENSELER / VASELLA)

HERMSTRÜWER YOAN, Fairnessprinzipien der algorithmischen Verwaltung: Diskriminierungsprävention beim staatlichen Einsatz von Machine Learning, AöR 2020, p. 479 ss

HUSI-STÄMPFLI SANDRA / MORAND ANNE-SOPHIE, Datenschutzrecht, 2e édition, Zurich / Genève 2024

KÄGI-DIENER REGULA, Die Bedeutung internationaler Diskriminierungsverbote, insbesondere CEDAW, für die schweizerische Rechtsordnung, Questions au féminin 2009, p. 42 ss (citation : KÄGI-DIENER, Bedeutung)

KÄGI-DIENER REGULA, Impulse des CEDAW-Übereinkommens für die Gleichstellung im Erwerbsleben insbesondere in der Quotenfrage, AJP 2006, p. 1451 ss (citation : KÄGI-DIENER, Impulse)

KAUFMANN CLAUDIA, Kommentierung des Art. 3 GlG. Dans : Kaufmann Claudia / Steiger-Sackmann Sabine (dir.), Kommentar zum Gleichstellungsgesetz, 3e édition, Bâle 2022 (citation : GlG-KAUFMANN)

KIENER REGINA / KÄLIN WALTER / WYTTENBACH JUDITH, Grundrechte, 4e édition, Berne 2024

KNEUBÜHLER LORENZ, Die Begründungspflicht: Eine Untersuchung über die Pflicht der Behörden zur Begründung ihrer Entscheide, thèse Univ. Berne 1998, Berne / Stuttgart / Vienne 1998

KRAMER RUDI, Herr, Frau – und nun? Datenschutzrechtliche Aspekte bei der geschlechtlichen Anrede, DSB 2025, p. 27 ss

KÜRSTEINER THOMAS, Künstliche Intelligenz und Rechtsprechung diskutiert am Beispiel der Begründungspflicht, Revue suisse des juges « Justice - Justiz – Giustizia » 1/2024

KUMKAR LEA KATHARINA, Transparenzanforderungen an Hochrisiko- und andere KI-Systeme. Dans : Hilgendorf Eric / Roth-Isigkeit David (dir.), Die neue Verordnung der EU zur Künstlichen Intelligenz, Munich 2023, p. 109 ss

LEGNER SARAH, KI-Verordnung und algorithmische Diskriminierung, RDi 2024, p. 426 ss

MAHON PASCAL / GRAF ANNE-LAURENCE / STEFFANINI FEDERICA, La notion de « race » dans le droit suisse, étude juridique établie à la demande du Service de lutte contre le racisme du Département fédéral de l'intérieur, Neuchâtel, 30 octobre 2019 (https://www.frb.admin.ch/dam/fr/sd-web/wgWn3YKX8Q4v/La%20notion%20de%20%C2%AB%20race%20%C2%BB%20dans%20le%20droit%20suisse.pdf)

MARTINI MARIO, Hochrisiko-KI-Systeme: Risikobasierter Ansatz. Dans : Hilgendorf Eric / Roth-Isigkeit David (dir.), Die neue Verordnung der EU zur Künstlichen Intelligenz, Munich 2023, p. 51 ss (citation : MARTINI, Hochrisiko-KI-Systeme)

MARTINI MARIO, Blackbox Algorithmus, Berlin 2019 (citation : MARTINI, Blackbox)

MEILI ANDREAS, Kommentierung des Art. 28 ZGB. Dans : Geiser Thomas / Fountoulakis Christiana (dir.), Zivilgesetzbuch I, Basler Kommentar, 7e édition, Bâle 2022 (citation : BSK ZGB I-MEILI)

MOSIMANN HANS-JAKOB, Entscheidbegründung: Begründung und Redaktion von Gerichtsurteilen und Verfügungen, Zurich 2013

MÜLLER ANGELA, «Es geht nicht darum, Technologie zu verteufeln», Questions au féminin 2024, p. 26 ss

MÜLLER JÖRG PAUL / SCHEFER MARKUS, Grundrechte in der Schweiz, 4e édition, Berne 2008

MÜLLER JAN-LAURIN, Algorithmische Entscheidungssysteme im Nichtdiskriminierungsrecht. Dans : Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz / Frauke Rostalski (dir.), Künstliche Intelligenz – Wie gelingt eine vertrauenswürdige Verwendung in Deutschland und Europa?, Tübingen 2022, p. 205 ss

NAGUIB TAREK, Mehrfachdiskriminierung: Analysekategorie im Diskriminierungsschutzrecht, SJZ-SRJ 2010, p. 233 ss

NAGUIB TAREK / PÄRLI KURT, Diskriminierungsschutzrecht: Errungenschaften und Kritik. Dans : Naguib Tarek / Pärli Kurt / Copur Eylem / Studer Melanie (dir.), Diskriminierungsrecht, Handbuch für Jurist_innen, Berater_innen und Diversity-Expert_innen, Berne 2014, p. 445 ss

OBRECHT LILIANE, Öffentliche Verzeichnisse sind ideal, um über die Risiken und Chancen von KI zu diskutieren, Questions au féminin 2024, p. 78 ss (citation : OBRECHT, Questions au féminin)

OBRECHT LILIANE, Verfügung und automatisierte Einzelentscheidung – same same but different?,

ex/ante 2/2022, p. 38 ss (citation OBRECHT, Verfügung)

ODERMATT DANIELA FRANZISKA, Das Konzept der Intersektionalität im Schweizerischen Recht, QFLR 1/2022, p. 7 ss

ORWAT CARSTEN, Diskriminierungsrisiken durch Verwendung von Algorithmen, Antidiskriminierungsstelle des Bundes, Berlin 2019 (https://www.antidiskriminierungsstelle.de/SharedDocs/downloads/DE/publikationen/Expertisen/studie_diskriminierungsrisiken_durch_verwendung_von_algorithmen.pdf?__blob=publicationFile&v=3)

PÄRLI KURT, Aids: Datenschutz und Diskriminierung, digma 2003, p. 127 ss

PÄRLI KURT / CAPLAZI ALEXANDRA / SUTER CAROLINE, Recht gegen HIV/Aids-Diskriminierung im Arbeitsverhältnis, Berne 2007

PÄRLI KURT / FLÜCK NATHALIE, Kommentierung des Art. 21 DSG. Dans : Baeriswyl Bruno / Pärli Kurt / Blonski Dominika (dir.), Datenschutzgesetz (DSG): Bundesgesetz vom 25. September 2020 über den Datenschutz, 2e édition, Berne 2023 (citation : SHK DSG-PÄRLI / FLÜCK)

PÄRLI KURT / NAGUIB TAREK, Protection contre la discrimination raciale en droit privé – Analyse et proposition de solutions, étude juridique établie à la demande de la Commission fédérale contre le racisme (CFR), juin 2025 (https://www.ekr.admin.ch/pdf/EKR_Berichte_Paerli_FR_250603.pdf)

RECHSTEINER DAVID, Der Algorithmus verfügt, Jusletter du 26 novembre 2018

RHINOW RENÉ / SCHEFER MARKUS / UEBERSAX PETER, Schweizerisches Verfassungsrecht, 3e édition, Bâle 2016

RUSCHEMEIER HANNAH, Kommentierung des Art. 6 KI-VO. Dans : Martini Mario / Wendehorst Christiane (dir.), Kommentar zur KI-VO: Verordnung über künstliche Intelligenz, Munich 2024 (citation : KI-VO-RUSCHEMEIER)

RUSCHEMEIER HANNAH, The Problems of the Automation Bias in the Public Sector – A Legal Perspective, Weizenbaum Conference proceedings 2023 (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4521474, citation : RUSCHEMEIER, Automation Bias)

SCHWEIZER RAINER J. / FANKHAUSER KIM, Kommentierung des Art. 8 BV. Dans : Bernhard Ehrenzeller et al. (dir.), Die schweizerische Bundesverfassung, St. Galler Kommentar, 4e édition, Zurich 2023 (citation : SGK BV-SCHWEIZER / FANKHAUSER)

SIMMLER MONIKA / CANOVA GIULIA, Gesichtserkennungstechnologie: Die «smarte» Polizeiarbeit auf dem rechtlichen Prüfstand, Sicherheit&Recht / Sécurité&Droit 2021, p. 105 ss

ŠKORJANC ŽIGA, Automatisierte Kreditentscheidungen, CB 2020, p. 70 ss

SPINDLER GERALD, Anforderungen an Hochrisiko-KI-Systeme (ausser Transparenz). Dans : Hilgendorf Eric / Roth-Isigkeit David (dir.), Die neue Verordnung der EU zur Künstlichen Intelligenz, Munich 2023, p. 93 ss

STANOVSKY GABRIEL / SMITH NOAH A. / ZETTLEMOYER LUKE, Evaluating Gender Bias in Machine Translation, Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 28 juillet – 2 août 2019, p. 1679 ss

STEINER OLIVIER, Das Verbot der indirekten Lohndiskriminierung, AJP 2001, p. 1281 ss

STENGEL CORNELIA / STÄUBLE LUCA, Kommentierung des Art. 25 DSG. Dans : Bieri Adrian / Powell Julian (dir.), DSG Kommentar: Kommentar zum Schweizerischen Datenschutzgesetz mit weiteren Erlassen, Zurich 2023 (citation : OFK DSG-STENGEL / STÄUBLE)

STREIFF ULLIN / VON KAENEL ADRIAN / RUDOLPH ROGER, Arbeitsvertrag, Praxiskommentar zu Art. 319 – 362 OR, 7e édition, Zurich 2012 (citation : PK OR-STREIFF / VON KAENEL / RUDOLPH)

SUTER FABIENNE, Automatisierte Einzelentscheidungen im (Schweizer) Datenschutzrecht: Phänomen, Entstehungsgeschichte, Regelungsziele und Begriff, thèse Univ. Zurich 2023, Zurich 2024

THOUVENIN FLORENT / FRÜH ALFRED, Automatisierte Entscheidungen: Grundfragen aus der Perspektive des Privatrechts, SZW-RSDA 2020, S. 3 ff.

THOUVENIN FLORENT / VOLZ STEPHANIE, Diskriminierung, White Paper de juin 2024 (https://www.itsl.uzh.ch/dam/jcr:19d5ad6c-2671-431c-86d7-822e27877606/2024_ITSL_WhitePaper_Diskriminierung.pdf, citation : THOUVENIN / VOLZ, WP Diskriminierung)

THOUVENIN FLORENT / VOLZ STEPHANIE, Ein Rechtsrahmen für den Einsatz von Künstlicher Intelligenz (KI) in der Schweiz, ZBJV 2024, p. 613 ss (citation : THOUVENIN / VOLZ, Rechtsrahmen)

THOUVENIN FLORENT / VOLZ STEPHANIE / SORAYA WEINER / CHRISTOPH HEITZ, Diskriminierung beim Einsatz von Künstlicher Intelligenz (KI), Jusletter IT du 4 juillet 2024

TINHOFER ANDREAS, Algorithmenbasierte Entscheidungen und Diskriminierung, DRdA 2022, p. 170 ss

UTTINGER URSULA / GEISER THOMAS, Das neue Datenschutzrecht, Bâle 2023

VILLIGER MARK EUGEN, Die Pflicht zur Begründung von Verfügungen, ZBl 90/1989, p. 137 ss

WALDMANN BERNHARD, Kommentierung des Art. 8 BV. Dans : Waldmann Bernhard / Belser Eva Maria / Epiney Astrid (dir.), Bundesverfassung, Basler Kommentar, Bâle 2015 (citation : BSK BV-WALDMANN)

WALDMANN BERNHARD, Das Diskriminierungsverbot von Art. 8 Abs. 2 BV als besonderer Gleichheitssatz, Berne 2003 (citation : WALDMANN, Diskriminierungsverbot)

WEBER ROLF H. / HENSELER SIMON, Regulierung von Algorithmen in der EU und in der Schweiz: EuZ 2020, p. 28 ss

WEDER REGINA, Verfahrensgrundrechtliche Anforderungen an automatisierte Verwaltungsverfahren. Dans : Simmler Monika (dir.), Smart Criminal Justice, Bâle 2021, p. 237 ss

WETZSTEIN SIMONE / WOLFENSBERGER JULIANE, Kommentierung des Art. 3 GlG. Dans : Facincani Nicolas / Hirzel Monika / Sutter Reto / Wetzstein Simone (dir.), Gleichstellungsgesetz: Bundesgesetz über die Gleichstellung von Frau und Mann vom 24. März 1995, Berne 2022 (citation : SHK GlG-WETZSTEIN / WOLFENSBERGER)

WIEDERKEHR RENÉ, Die Begründungspflicht nach Art. 29 Abs. 2 BV und die Heilung bei Verletzung, ZBl 2010, p. 481 ss

WILDHABER ISABELLE / LOHMANN MELINDA F. / KASPER GABRIEL, Diskriminierung durch Algorithmen – Überlegungen zum schweizerischen Recht am Beispiel prädiktiver Analytik am Arbeitsplatz, ZSR-RDS I 2019, p. 459 ss

Bases légales / normes

Remarque : les bases légales sont présentées selon la hiérarchie des normes (Constitution – loi – ordonnance), par ordre alphabétique pour chaque niveau.

Constitution et droit international

ALCP	Accord du 21 juin 1999 entre la Confédération suisse, d'une part, et la Communauté européenne et ses États membres, d'autre part, sur la libre circulation des personnes (RS 0.142.112.681)
CC	Code civil suisse, du 10 décembre 1907 (RS 210)
CEDEF	Convention du 18 décembre 1979 sur l'élimination de toutes les formes de discrimination à l'égard des femmes (RS 0.108)
CEDH	Convention du 4 novembre 1950 de sauvegarde des droits de l'homme et des libertés fondamentales (Convention européenne des droits de l'homme) (RS 0.101)
CERD	Convention internationale du 21 décembre 1965 sur l'élimination de toutes les formes de discrimination raciale (RS 0.104)
CP	Code pénal suisse, du 21 décembre 1937 (RS 311.0)
CPC	Code de procédure civile, du 19 décembre 2008 (RS 272)
CPM	Code pénal militaire, du 13 juin 1927 (RS 321.0)
Cst.	Constitution fédérale de la Confédération suisse, du 18 avril 1999 (RS 101)
CO	Loi fédérale du 30 mars 1911 complétant le Code civil suisse (Livre cinquième : Droit des obligations) (RS 220)
LAGH	Loi fédérale du 15 juin 2018 sur l'analyse génétique humaine (RS 810.12)
LEEJ	Loi fédérale du 30 septembre 2011 sur l'encouragement des activités extrascolaires des enfants et des jeunes (RS 446.1)
LEg	Loi fédérale du 24 mars 1995 sur l'égalité entre femmes et hommes (RS 151.1)
LEHE	Loi fédérale du 30 septembre 2011 sur l'encouragement des hautes écoles et la coordination dans le domaine suisse des hautes écoles (RS 414.20)
LHand	Loi fédérale du 13 décembre 2002 sur l'élimination des inégalités frappant les personnes handicapées (RS 151.3)

LMETA	Loi fédérale du 17 mars 2023 sur l'utilisation des moyens électroniques pour l'exécution des tâches des autorités (RS 172.019)
LMP	Loi fédérale du 21 juin 2019 sur les marchés publics (RS 172.056.1)
LPD	Loi fédérale du 25 septembre 2020 sur la protection des données (RS 235.1)
LPers	Loi du 24 mars 2000 sur le personnel de la Confédération (RS 172.220.1)
LTr	Loi fédérale du 13 mars 1964 sur le travail dans l'industrie, l'artisanat et le commerce (RS 822.11)
OPersTF	Ordonnance du 27 août 2001 sur le personnel du Tribunal fédéral (RS 172.220.114)
OSAMal	Ordonnance du 18 novembre 2015 sur la surveillance de l'assurance-maladie sociale (RS 832.121)
PA	Loi fédérale du 20 décembre 1968 sur la procédure administrative (RS 172.021)
Pacte II ONU	Pacte international du 16 décembre 1966 relatif aux droits civils et politiques (RS 0.103.2)

Canton de Zurich (ZH)

Cst-ZH	Constitution du Canton de Zurich du 27 février 2005 (LS 101 ; RS 131.211)
HOPH-ZH	Hochschulordnung der Pädagogischen Hochschule Zürich vom 22. November 2017 (LS 414.410)
PVF-ZH	Personalverordnung der Zürcher Fachhochschule vom 16. Juli 2008 (LS 414.112)
R-UZH-SB-ZH	Reglement zum Schutz vor sexueller Belästigung an der Universität Zürich vom 1. März 2007 (LS 415.116)
SVO-ZH	Submissionsverordnung vom 28. Juni 2023 (LS 720.11)

Canton de Bâle-Ville (BS)

Cst-BS	Constitution du Canton de Bâle-Ville du 23 mars 2005 (SG 111.100 ; RS 131.222.1)
TBG-BS	Gesetz betreffend Tagesbetreuung von Kindern vom 8. Mai 2019 (SG 815.100)

Vertrag FHNW-BS	Vertrag zwischen den Kantonen Aargau, Basel-Landschaft, Basel-Stadt und Solothurn über die Fachhochschule Nordwestschweiz vom 9. November 2004 (SG 428.100)
WRFV-BS	Verordnung über die Wohnraumförderung vom 17. Juni 2014 (SG 861.520)

Canton d'Argovie (AG)

Cst-AG	Constitution du Canton d'Argovie, du 25 juin 1980 (SAR 110.000 ; RS 131.227)

Canton de Vaud (VD)

AOCCF-VD	Arrêté du 2 mars 1990 concernant les offices de consultation conjugale ou familiale (RSV 850.455.1)
Cst-VD	Constitution du Canton de Vaud du 14 avril 2003 (RSV 101.01 ; RS 131.231)
LIEPR-VD	Loi du 23 janvier 2007 sur l'intégration des étrangers et sur la prévention du racisme (RSV 142.52)
LOVD-VD	Loi du 26 septembre 2017 d'organisation de la prévention et de la lutte contre la violence domestique (RSV 211.12)
Lsubv-VD	Loi du 22 février 2005 sur les subventions (RSV 610.15)
RLOVD-VD	Règlement du 10 octobre 2018 d'application de la loi du 26 septembre 2017 d'organisation de la prévention et de la lutte contre la violence domestique (RSV 211.12.1)
RLRNSS-VD	Règlement du 18 décembre 2019 d'application de la loi du 11 décembre 2018 sur les ressources naturelles du sous-sol (RSV 730.02.1)
RLSubV-VD	Règlement du 22 novembre 2006 d'application de la loi du 22 février 2005 sur les subventions (RSV 610.15.1)

Normes internationales

Convention CE-IA	Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit, du 5 septembre 2024
Règlement UE-IA	Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle et modifiant les règlements (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 et (UE) 2019/2144 et les directives 2014/90/UE, (UE) 2016/797 et (UE) 2020/1828 (règlement sur l'intelligence artificielle)

Documents de référence

Remarque : les documents sont classés par date de publication (des plus récents aux plus anciens).

Office fédéral de la communication (OFCOM), État des lieux sur la règlementation de l'intelligence artificielle, rapport du 12 février 2025 à l'attention du Conseil fédéral (https://www.bakom.admin.ch/fr/intelligence-artificielle, citation : État des lieux IA 2025)

Office fédéral de la justice (OFJ), Analyse juridique de base dans le cadre de l'état des lieux sur les approches de régulation en matière d'intelligence artificielle, du 31 août 2024 (https://www.bj.admin.ch/bj/fr/home/staat/gesetzgebung/kuenstliche-intelligenz.html, citation : Analyse juridique de base 2024)

Comité sur l'intelligence artificielle (CAI), Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit, rapport explicatif CM(2024)52-addfinal, du 17 mai 2024 (https://search.coe.int/cm?i=091259488025fb07, citation : Rapport explicatif Convention CE-IA)

Message du 15 septembre 2017 concernant la loi fédérale sur la révision totale de la loi fédérale sur la protection des données et sur la modification d'autres lois fédérales, FF *2017* 6565 (https://www.fedlex.admin.ch/eli/fga/2017/2057/fr, citation : Message LPD 2017)

Message du 5 juillet 2017 concernant la loi fédérale sur l'analyse génétique humaine, FF *2017* 5253, https://www.fedlex.admin.ch/eli/fga/2017/1501/fr, citation : Message LAGH 2017)

Message du 24 février 1993 concernant la loi fédérale sur l'égalité entre femmes et hommes (loi sur l'égalité) et l'arrêté fédéral relatif à une modification de l'ordonnance concernant l'attribution des offices aux départements et des services à la Chancellerie fédérale, FF *1993* 1163 (https://www.fedlex.admin.ch/eli/fga/1993/1_1248_1163_987/fr, citation : Message LEg 1993)

Message et projet d'arrêté du 23 août 1995 relatif à la Convention de 1979 sur l'élimination de toutes les formes de discrimination à l'égard des femmes, FF *1995* IV 869 (https://www.fedlex.admin.ch/fr/fga/index/1995, citation : Message CEDEF 1995)

Message du 5 mai 1982 concernant la révision du code civil suisse (Protection de la personnalité : art. 28 CC et 49 CO), FF *1982* II 661 (https://www.fedlex.admin.ch/eli/fga/1982/2_636_661_628/fr, citation : Message CC 1982)