

LA PROTECTION JURIDIQUE CONTRE LA DISCRIMINATION ALGORITHMIQUE EN EUROPE

Cadres actuels et
lacunes substantantes



En collaboration avec
le Centre interfédéral pour l'égalité des chances (Unia), Belgique,
le Médiateur de la non-discrimination (YVV), Finlande,
et la Commission pour la citoyenneté et l'égalité de genre (CIG), Portugal

Raphaële Xenidis

Cofinancé
par l'Union européenne



UNION EUROPÉENNE

COUNCIL OF EUROPE



CONSEIL DE L'EUROPE

Cofinancé et mis en œuvre
par le Conseil de l'Europe

LA PROTECTION JURIDIQUE CONTRE LA DISCRIMINATION ALGORITHMIQUE EN EUROPE :

Cadres actuels et
lacunes subsistantes

Raphaële Xenidis

Conseil de l'Europe

Ce rapport a été produit avec le soutien financier de l'Union européenne et du Conseil de l'Europe. Son contenu relève de la responsabilité exclusive de son autrice. Les opinions exprimées dans ce rapport ne peuvent en aucun cas être considérées comme reflétant l'opinion officielle de l'Union européenne ou du Conseil de l'Europe.

La reproduction d'extraits (jusqu'à 500 mots) est autorisée, sauf à des fins commerciales, à condition que l'intégrité du texte soit préservée, que l'extrait ne soit pas utilisé hors contexte, ne fournisse pas d'informations incomplètes ou n'induisse pas le lecteur en erreur quant à la nature, à la portée ou au contenu du texte. Le texte source doit toujours être crédité comme suit : « © Conseil de l'Europe, année de la publication ». Toute autre demande concernant la reproduction/traduction de tout ou partie du document doit être adressée à la Division des publications et de l'identité visuelle du Conseil de l'Europe (Publications and Visual Identity Division - F-67075 Strasbourg Cedex ou publishing@coe.int).

Toute autre communication relative au présent document doit être adressée à l'Unité « Discours de haine, crimes de haine et intelligence artificielle » de la Division des programmes d'inclusion et de lutte contre la discrimination du Conseil de l'Europe, F-67075 Strasbourg Cedex, France.

Courriel : anti-discrimination@coe.int

Le texte original en anglais est celui du Conseil de l'Europe. Cette traduction française est publiée en accord avec le Conseil de l'Europe, mais relève de la seule responsabilité des traducteurs-rices. L'unité éditoriale de la Division des publications et de l'identité visuelle du Conseil de l'Europe n'a corrigé aucune erreur typographique ou grammaticale dans cette publication française.

Design page de couverture : Division des publications et de l'identité visuelle (DPIV), Conseil de l'Europe
Mise en page de la version anglaise : Jouve, Paris
Mise en page de la version française : MCI Benelux S.A.
Traduction : LingvoHouse Translation Services Ltd
Révision juridique de la traduction : Diana Nunes, consultante
Images : Shutterstock

Table des matières

ABRÉVIATIONS	5
REMERCIEMENTS	6
INTRODUCTION	7
1. CADRES JURIDIQUES EXISTANTS	9
1.1 Définitions	9
1.2 La discrimination algorithmique en Europe	10
1.3 Le recours aux systèmes d'IA et de PDA dans les administrations publiques et les secteurs d'intérêt	12
1.4 Cadres juridiques existants	16
1.4.1 Régir les systèmes d'IA et de PDA	16
1.4.2 Instruments de lutte contre la discrimination	30
1.4.3 Instruments de protection des données	36
1.4.4 Jurisprudence de la Cour européenne des droits de l'homme et de la Cour de justice de l'Union européenne	39
1.5 Principaux acteurs concernés	42
2. LACUNES JURIDIQUES	47
2.1 Justification de l'adoption de systèmes d'IA et de PDA	47
2.2 Manque de transparence	48
2.3 Problématiques d'accès à la justice	50
2.4 Application des dispositions existantes et futures	52
2.4.1 Coopération institutionnelle	52
2.4.2 Lacunes conceptuelles	53
2.4.3 Restrictions pour la collecte et le traitement des données	55
2.5 « Dérisquer » les systèmes d'IA	55
2.6 Les normes techniques, les organismes de promotion de l'égalité et la question de l'harmonisation	56
RÉSUMÉ – UTILISATION DE L'IA DANS L'ADMINISTRATION PUBLIQUE BELGE ET RECOMMANDATIONS POLITIQUES SUR LA DISCRIMINATION ALGORITHMIQUE	61
RÉSUMÉ – UTILISATION DE L'IA DANS L'ADMINISTRATION PUBLIQUE FINLANDAISE ET RECOMMANDATIONS POLITIQUES SUR LA DISCRIMINATION ALGORITHMIQUE	65
RÉSUMÉ – UTILISATION DE L'IA DANS L'ADMINISTRATION PUBLIQUE PORTUGAISE ET RECOMMANDATIONS POLITIQUES SUR LA DISCRIMINATION ALGORITHMIQUE	70
RÉFÉRENCES	75

Abréviations

IA	Intelligence artificielle
PDA	Prise de décision automatisée
OSC	Organisations de la société civile
CAI	Comité du Conseil de l'Europe sur l'intelligence artificielle
CEN	Comité européen de normalisation
CENELEC	Comité européen de normalisation électrotechnique
AIDF	Analyses d'impact sur les droits fondamentaux
RGPD	Règlement général sur la protection des données
GPAI	IA à usage général
HUDERIA	Évaluation des risques et impacts des systèmes d'IA du point de vue des droits de l'homme, de la démocratie et de l'État de droit
TUE	Traité sur l'Union européenne
TFUE	Traité sur le fonctionnement de l'Union européenne

Remerciements

Ce rapport a été préparé dans le cadre du projet « Défense de l'égalité et de la non-discrimination par les organismes de promotion de l'égalité concernant l'utilisation de l'intelligence artificielle (IA) dans les administrations publiques », financé par l'Union européenne via l'Instrument d'appui technique et cofinancé par le Conseil de l'Europe. Le projet est mis en œuvre par le Conseil de l'Europe en coopération avec la Commission européenne, le Centre interfédéral pour l'égalité des chances (Unia, Belgique), le Médiateur contre les discriminations (Finlande) et la Commission pour la citoyenneté et l'égalité des genres (Portugal).

Ce rapport a été rédigé par Raphaële Xenidis (maîtresse de conférences en droit européen à Sciences Po, Faculté de droit, France). Les synthèses analytiques des contextes nationaux ont été tirées des travaux d'Ine van Zeeland (consultante nationale du Conseil de l'Europe) pour la partie belge, d'Emeline Banzuzi (consultante nationale du Conseil de l'Europe) pour la partie finlandaise et d'Eduardo dos Santos (consultant national du Conseil de l'Europe) pour la partie portugaise.

Le Conseil de l'Europe souhaite exprimer sa gratitude aux institutions bénéficiaires du projet et à la Commission européenne pour leur engagement soutenu tout au long du processus de rédaction, et en particulier à : Nele Roekens, Nadine Brauns et Thelma Raes (Centre interfédéral pour l'égalité des chances, Unia, Belgique); Tiina Valonen et Ville Rantala (Médiateur contre les discriminations, YVV, Finlande); Carla Peixe, Ana Martinho Fernandes, Alexandra Andrade et Susana Miguel (Commission pour la citoyenneté et l'égalité des genres, CIG, Portugal) et Massimiliano Santini (Groupe de travail «Réforme et investissement», Secrétariat général, Commission européenne) ainsi qu'à Menno Ettema, Sara Haapalainen, Ayça Dibekoğlu et Delfine Gaillard (Département anti-discrimination, Conseil de l'Europe).

Le rapport a également bénéficié des contributions d'expert-e-s nationaux-ales d'entités publiques et privées basées en Belgique, en Finlande et au Portugal, qui ont participé à des entretiens fin 2024 et début 2025. Leurs perspectives ont enrichi les parties nationales du rapport.





Introduction

La hausse de l'intégration de l'intelligence artificielle (IA) dans diverses applications au sein de nombreux secteurs différents amène d'importantes problématiques vis-à-vis de la protection des droits fondamentaux. Parmi ces questions, la discrimination algorithmique apparaît comme une préoccupation particulièrement urgente. La recherche empirique a démontré qu'en plus de refléter les inégalités sociales existantes à grande échelle, les biais algorithmiques les exacerbent également, en particulier dans les domaines où la prise de décision algorithmique est prédominante, tels que la police, l'emploi, l'éducation et les assurances. Pendant longtemps, dans le contexte européen, l'intersection entre le droit de la non-discrimination et le droit relatif à la protection des données a servi de base juridique principale pour traiter les cas de discrimination algorithmique.

Cependant, les évolutions juridiques, institutionnelles et politiques récentes ont considérablement remodelé le paysage de la gouvernance de l'IA en Europe. En 2024, le Conseil de l'Europe et l'Union européenne ont adopté des instruments juridiques historiques : la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit (Convention-cadre sur l'IA) et le règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées en matière d'intelligence artificielle (règlement européen sur l'IA). Ces instruments établissent des cadres de gouvernance complets

conçus pour protéger les droits fondamentaux, dont le droit à la non-discrimination, dans le contexte de l'IA et des systèmes de prise de décision automatisée (PDA). L'adoption de ces cadres devrait avoir des incidences juridiques et institutionnelles substantielles au niveau national, notamment compte tenu du fait que d'ici 2026, les États membres seront tenus d'adapter leur législation et leurs structures administratives nationales sur les obligations énoncées dans le règlement européen sur l'IA.

Le règlement européen sur l'IA souligne la nécessité de développer les technologies d'IA et leurs cadres réglementaires dans le respect des valeurs de l'UE consacrées à l'article 2 du traité sur l'Union européenne (TUE), ainsi que des libertés et droits fondamentaux garantis par les traités et de la Charte des droits fondamentaux, conformément à l'article 6 du traité sur l'Union européenne. Il affirme explicitement que l'IA doit être une technologie centrée sur l'humain. De même, la Convention-cadre du Conseil de l'Europe sur l'IA précise son objectif, qui consiste à veiller à ce que toutes les activités du cycle de vie de l'IA soient pleinement compatibles avec les droits humains, la démocratie et l'État de droit.

Dans ce rapport, nous verrons comment ces instruments juridiques émergents contribuent au renforcement des dispositifs de protection contre la discrimination algorithmique en Europe, et nous évaluerons les lacunes qui continuent de les fragiliser.

Parallèlement, deux nouvelles directives de l'UE concernant les organismes de promotion de l'égalité visent à renforcer le rôle et les capacités de ces institutions dans toute l'Europe (Union européenne 2024a, 2024b). Les organismes de promotion de l'égalité, ainsi que les institutions nationales de défense des droits humains et les médiateurs, devraient jouer un rôle central dans la gouvernance des systèmes d'IA et de PDA, notamment en veillant au respect des normes relatives aux droits fondamentaux, dont le principe de non-discrimination. Ainsi, ce rapport vise à retracer et synthétiser ces évolutions afin de fournir aux acteurs concernés, parmi lesquels les organismes de promotion de l'égalité, des outils qui leur permettent de s'adapter à un paysage juridique en mutation et, au besoin, d'intervenir dans les évolutions en cours.

Le rapport est structuré en deux grandes parties :

- ▶ une évaluation des cadres juridiques existants;
- ▶ un état des lieux des lacunes subsistantes.

La première partie propose des définitions (1) et des exemples de déploiement et d'utilisation des systèmes d'IA et de PDA par les administrations publiques et dans d'autres secteurs d'intérêt au niveau européen (2). Après avoir expliqué comment la discrimination algorithmique se produit (3), nous recenserons les cadres juridiques, procéduraux et de gouvernance actuels liés à la lutte contre la discrimination, à l'égalité et à l'IA en Europe (4). Cette partie se terminera par une réflexion sur les rôles actuels et potentiels des parties prenantes concernées face à la discrimination algorithmique (5).

La deuxième grande partie présente les lacunes juridiques en ce qui concerne ces cadres et leur application. Ce rapport se limite principalement aux systèmes d'IA et de PDA utilisés par les administrations publiques en Europe, et au second plan, utilisés dans les secteurs privés tels que les grandes agences de travail intérimaire.



1. Cadres juridiques existants

1.1. Définitions

Tout au long de ce rapport, les termes « IA » et « systèmes de PDA » sont employés pour décrire les opérations de systèmes fondés sur l'apprentissage et sur des règles, qui peuvent à la fois perpétuer et amplifier la discrimination. Selon l'article 2 de la Convention-cadre du Conseil de l'Europe sur l'IA, un système d'IA est « 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 influencer sur des environnements physiques ou virtuels ». Dans ce texte, il est établi que « différents systèmes d'intelligence artificielle présentent des degrés variables d'autonomie et d'adaptabilité après déploiement » (Conseil de l'Europe 2024a). Cette définition fait écho à la définition proposée à l'article 3, paragraphe 1, du règlement européen sur l'IA¹. Dans le domaine des techniques d'IA, l'utilisation de l'apprentissage automatique en particulier s'est généralisée pour identifier ce que l'on appelle les corrélations entre

1. L'article 3, paragraphe 1, du règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées en matière d'intelligence artificielle définit un système d'IA comme « un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels ».

les points de données afin de faire des prédictions et des recommandations ou de noter, classer et hiérarchiser des éléments ou des personnes.² En revanche, certains systèmes de PDA ne possèdent pas le niveau d'adaptabilité et d'autonomie requis pour être considérés comme des systèmes d'IA, mais il reste néanmoins important de les étudier dans le présent rapport étant donné leur potentiel discriminatoire. Par exemple, les systèmes fondés sur des règles peuvent fausser la classification d'individus et automatiser les processus décisionnels basés sur cette classification, entraînant des conséquences potentiellement discriminatoires. L'expression « systèmes de PDA » couvre un large éventail de systèmes de décision, allant des systèmes entièrement automatisés qui n'impliquent aucune intervention humaine, tels que ceux visés à l'article 22 du Règlement général sur la protection des données (RGPD) de l'UE, aux systèmes semi-automatisés ou d'aide à la décision, qui ont à la fois recours à une intervention humaine et à une intervention automatisée et sont très largement utilisés dans la pratique.³ Ce rapport porte à la fois sur les systèmes d'IA et de PDA.

1.2. La discrimination algorithmique en Europe

Dans la prise de décision algorithmique, la discrimination survient lorsque les systèmes automatisés désavantagent injustement certains individus ou groupes. Plusieurs facteurs clés contribuent à la discrimination algorithmique.

Premièrement, les biais présents dans les données d'entraînement et d'évaluation peuvent conduire à des décisions discriminatoires. D'abord, les pratiques de collecte et de production de données biaisées produisent des ensembles de données non représentatifs, mais lorsque l'IA et les systèmes de PDA axés sur les données traitent ou sont entraînés à partir de données historiques qui reflètent des préjugés, des inégalités ou des exclusions de certaines catégories ethniques, de genre ou socio-économiques, par exemple, ces pratiques ont également tendance à reproduire et/ou amplifier ces biais. Par exemple, un algorithme de sélection entraîné à partir de données historiques d'embauche pourrait favoriser injustement les candidats masculins en reproduisant des décisions d'embauche antérieures biaisées par des stéréotypes discriminatoires à l'égard des femmes ou des groupes minoritaires.

Deuxièmement, les choix de conception faits par les développeurs ou développeuses, y compris avec la sélection des caractéristiques cibles et des prédicteurs, le type de modèle, les indicateurs d'équité et les seuils, peuvent refléter des biais discriminatoires. Si les développeurs ne tiennent pas compte des inégalités systémiques, les systèmes d'IA et de PDA pourraient prendre par inadvertance des décisions qui nuisent de manière disproportionnée à certains groupes. Par exemple, les services de protection sociale d'Europe ont de plus en plus recours à des systèmes de PDA

2. Selon Binns, « la majorité des utilisations de l'apprentissage automatique dans le secteur public relèvent de la catégorie supervisée. Cette supervision signifie qu'il est nécessaire d'indiquer à l'algorithme d'apprentissage ce que le décideur veut prédire ou classer, contrairement aux méthodes non supervisées, qui sont conçues pour découvrir une structure latente dans un ensemble de données ». Voir Binns R. (2020), « Algorithmic Decision-Making : a guide for Lawyers », *Judicial Review*, vol. 25, no 2, p. 3-4.
3. Pour une taxonomie de ces différents systèmes de PDA, voir Palmiotto F. (2024), « When is a decision automated? A taxonomy for a fundamental rights analysis », *German Law Journal*, Vol. 25, No. 2, pp. 210-36, accessible via : <https://doi.org/10.1017/glj.2023.112>, consulté le 6 novembre 2025.

pour anticiper les cas de fraude chez les bénéficiaires de prestations sociales. Il a été démontré que la catégorie même du « risque » et les prédicteurs opérationnalisés par ces systèmes de PDA reflètent en réalité des stéréotypes à l'encontre des personnes vivant dans la pauvreté qui sont profondément ancrés dans les pratiques, les formes et les institutions des services sociaux (Dubois 2021). Ce problème reflète l'omniprésence de la discrimination dite par variables de substitution, à savoir la discrimination basée sur le traitement algorithmique de variables en apparence neutres qui cryptent de manière invisible les inégalités.

Troisièmement, le manque de transparence et de responsabilisation des systèmes d'IA et de PDA actuellement utilisés rend difficile la surveillance ou la contestation des cas de discrimination algorithmique. De nombreux systèmes d'IA et de PDA, en particulier ceux basés sur des modèles complexes d'apprentissage automatique, fonctionnent comme des « boîtes noires », ce qui signifie que leurs processus décisionnels sont opaques. Ce manque de transparence fait qu'il est difficile de déterminer quand et comment la discrimination se produit, et empêche la mise en œuvre de contrôles efficaces et la responsabilisation de ces systèmes. Par exemple, le scandale causé aux Pays-Bas en 2021 par un algorithme utilisé pour prédire les cas de fraude de la part d'allocataires percevant des prestations familiales a montré combien il était difficile pour les personnes accusées à tort de comprendre comment et pourquoi elles avaient été classées comme fraudeuses, et encore plus de contester cette décision (Amnesty International 2021).

En outre, l'inadéquation des tests et de la surveillance opérés par les fournisseurs et les dépoyeurs de systèmes d'IA et de PDA, tant dans le secteur public que privé, renforcent la probabilité que des discriminations surviennent sans prévenir. Des processus de test et de surveillance doivent être mis en place avant, pendant et après leur déploiement. La participation des organisations de la société civile (OSC) concernées, des organismes de promotion de l'égalité et des institutions de défense des droits humains garantirait l'efficacité de ces procédures.

Enfin, la rapidité et le nombre des systèmes d'IA et de PDA risquent de propager la discrimination à un niveau systémique. Les promesses d'efficacité et de réduction des coûts avancées par le secteur ont encouragé de nombreuses organisations, dont des administrations publiques, à rationaliser la prise de décision en déployant des systèmes d'IA et de PDA. Pourtant, si l'on n'investit pas suffisamment de ressources dans la prévention et l'atténuation de la discrimination algorithmique, par exemple en formant des assistants sociaux et en investissant dans des données représentatives, des mécanismes d'audit et de test appropriés et des processus de responsabilisation, les systèmes d'IA et de PDA sont forcément discriminatoires. Par conséquent, l'argumentaire du secteur, selon lequel les systèmes d'IA et de PDA permettront une réduction drastique des coûts, doit être remis en question : il se pourrait en effet que de veiller à la légalité du fonctionnement des applications d'IA et de PDA, c'est-à-dire l'absence de discrimination, grâce à des mesures de garantie et des investissements adéquats, s'avère finalement coûteux.

1.3. Le recours aux systèmes d'IA et de PDA dans les administrations publiques et les secteurs d'intérêt

Les systèmes D'IA et de PDA sont de plus en plus utilisés dans le secteur privé comme dans le secteur public en Europe. Différents travaux ont illustré des exemples d'utilisation dans de nombreux domaines, tels que la santé, la finance et la banque, les transports et les ressources humaines. Les administrations publiques adoptent ou expérimentent progressivement des systèmes d'IA et de PDA dans des domaines tels que la fiscalité, le transport, la sécurité sociale, les migrations, la justice et la police. Les applications peuvent être développées en interne ou achetées auprès de sociétés privées externes. Au-delà de tâches administratives bénignes, les principaux usages de l'IA et des systèmes ADM dans ces domaines concernent la détection de fraude et la surveillance, qui présentent des risques élevés de discrimination. D'autres applications répandues sont les chatbots et les assistants virtuels, qui peuvent également présenter des caractéristiques discriminatoires, mais probablement avec un impact moins préjudiciable. Cela dit, l'adoption de l'IA, comparée à celle des systèmes ADM, semble différer : tandis que les applications d'IA développées en interne en sont peut-être encore à des stades expérimentaux, les systèmes ADM semblent être déployés plus largement.

Défis liés à la cartographie du recours à des systèmes d'IA et de PDA

L'exercice d'inventaire mené dans cette partie s'accompagne d'une mise en garde importante : cartographier l'utilisation des systèmes d'IA et de PDA dans les administrations publiques est une tâche difficile. La recherche met en avant le manque d'informations claires et cohérentes sur la façon dont les administrations publiques utilisent les systèmes d'IA et de PDA. Dans certains pays, comme la Finlande, les administrations publiques sont soumises à des obligations d'information et doivent publier des informations sur le recours à des systèmes de PDA sur leur site web et informer les personnes concernées de cette utilisation lors de la prise de décision. Cependant, les bases de données qui enregistrent systématiquement les applications d'IA et de PDA sont encore sous-développées.⁴ Ainsi, les parties prenantes concernées, telles que les organismes de promotion de l'égalité, les institutions nationales de défense des droits humains et les médiateurs rencontrent des difficultés dans l'exercice de leur mandat pour enquêter sur les cas de discrimination algorithmique, les surveiller, les identifier et les contester, ainsi que pour accompagner les victimes. Ces obstacles compromettent l'application effective du droit à la non-discrimination. D'autres stratégies déployées pour obtenir des informations (par exemple, les demandes au titre de la loi britannique sur la liberté d'information, les enquêtes parlementaires) se heurtent également à des écueils et des contraintes et ne permettent pas toujours d'avoir une vision suffisante de l'utilisation des systèmes d'IA et de PDA par les administrations publiques. En France, par exemple, les demandes d'accès à l'information formulées par les OSC pour clarifier un système de notation des risques employé par la Caisse nationale des allocations familiales (CNAF) a entraîné la publication limitée du code source des anciens modèles (par opposition au modèle actuel) et des listes de variables utilisées supprimées (La Quadrature du

4. Cela changera probablement après la mise en œuvre du règlement européen sur l'IA.

Net 2023). Parmi les autres problématiques, on peut citer le manque de transparence sur le recours à des systèmes d'IA et de PDA qui caractérise les questions et les demandes de renseignements en elles-mêmes : si les administrations publiques ne signalent pas l'utilisation de ce type de systèmes, il devient difficile de savoir où se renseigner en premier lieu.

Emploi

Parmi les utilisations signalées, on peut citer des outils de profilage utilisés par les agences publiques pour l'emploi pour prédire les chances d'emploi ou de chômage, qui sont déployés pour aider les agentes et les agents à allouer les aides aux personnes demandeuses d'emploi. Certains de ces outils de profilage ont été jugés discriminatoires. En Autriche, par exemple, il a été démontré qu'un prototype développé par l'agence autrichienne pour l'emploi, appelé « algorithme AMS » et conçu pour prédire les perspectives d'emploi des demandeur.ses d'emploi pour aider les agent.es à décider de l'allocation des aides, générait des discriminations, notamment à l'encontre des mères célibataires et des demandeur.ses d'emploi issu.es de l'immigration (Allhutter et al. 2020). Des systèmes similaires ont été déployés depuis 2018 par France Travail, agence française pour l'emploi, en vue d'évaluer les risques de fraude, de prévoir les désengagements et d'évaluer l'employabilité des bénéficiaires (La Quadrature du Net 2024). Des risques similaires de discrimination à l'encontre de personnes vivant dans des conditions socialement et économiquement précaires, mais aussi pour des raisons de sexe, de race⁵ ou d'appartenance ethnique, et de handicap, ont été signalés. D'autres préoccupations concernent le manque de conseils et de formation adéquats dispensés aux agent.es qui utilisent des outils d'IA et de PDA, et la manière dont les données personnelles des demandeur.ses d'emploi sont utilisées dans les applications de profilage. Les agences publiques pour l'emploi peuvent également déployer d'autres types d'outils, tels que des algorithmes d'appariement pour recommander des offres d'emploi aux personnes demandeuses d'emploi.

Domaine répressif

Les systèmes d'IA et de PDA sont également régulièrement utilisés par les autorités policières. Or, la recherche démontre que les technologies de reconnaissance faciale peuvent présenter des biais discriminatoires.⁶

5. Tous les êtres humains appartenant à la même espèce, le Comité des ministres du Conseil de l'Europe rejette les théories fondées sur l'existence de différentes « races ». Toutefois, dans ce document, le terme « race » est employé pour garantir que les personnes qui sont généralement et à tort perçues comme « appartenant à une autre race » ne soient pas exclues de la protection assurée par la législation et de la mise en œuvre des politiques.

6. L'OSC Liberty a critiqué une application de reconnaissance faciale utilisée par la police Pays de Galles du Sud, entre autres au titre de discriminations fondées sur le sexe et/ou l'origine ethnique, car l'application produisait un taux plus élevé de correspondances positives pour les visages féminins et/ou pour les visages noirs et ceux des minorités ethniques. Voir la décision ultérieure de la Cour d'appel du Royaume-Uni dans l'affaire R (Bridges) contre Chief Constable of South Wales Police ([2020] EWCA Civ 1058), qui souligne que l'« analyse d'impact sur l'égalité de la police du Pays de Galles du Sud était manifestement inadéquate et fondée sur une erreur de droit (aucune reconnaissance du risque de discrimination indirecte) » et que son « approche ultérieure de l'évaluation d'éventuelles discriminations indirectes découlant du recours à la reconnaissance faciale automatisée est défectueuse ».

Même lorsque les systèmes d'IA et de PDA ne sont pas eux-mêmes signalés comme étant biaisés, leur déploiement est critiqué en raison du ciblage et de la surveillance disproportionnés des minorités et du profilage ethnique opéré.⁷ Souvent, les preuves de ce type d'utilisations restent circonstancielles et le manque d'informations transparentes sur les systèmes utilisés et la manière dont ils sont utilisés empêche la réalisation d'enquêtes systématiques sur les risques qu'ils présentent pour les droits fondamentaux, dont le droit à la non-discrimination. Comme indiqué ci-dessous (voir la partie 1.4.1.), le règlement européen sur l'IA accorde de larges exceptions à la police et aux autorités répressives en ce qui concerne l'utilisation de technologies de reconnaissance traitant des images telles que des visages et d'autres données biométriques.

Secteur social

Plusieurs branches des systèmes de protection sociale et de sécurité sociale en Europe expérimentent ou déploient actuellement des systèmes d'IA et de PDA (Autriche, Belgique, Danemark, France, Pologne, Portugal, Slovaquie, Suède, Pays-Bas). En France, le système utilisé par la Caisse d'allocations familiales vise à prédire les risques de fraude et d'erreur chez les personnes allocataires dans le but d'aider les agent.es à cibler les contrôles (La Quadrature du Net 2023). Des risques de discrimination ont été signalés par les OSC et une affaire a été portée devant le Conseil d'État français, notamment au titre de discriminations fondées sur le sexe, la situation de famille, l'âge et le handicap (voir Conseil d'État 2024). Aux Pays-Bas, un système similaire a conduit à la démission du gouvernement en 2021 après qu'il a été jugé discriminatoire à l'égard de bénéficiaires en raison de leur « race », de leur origine ethnique et de leur nationalité.⁸ Dans le contexte de l'éducation, l'Agence exécutive néerlandaise de l'éducation a utilisé un algorithme de profilage des risques pour favoriser la détection des fraudes en relation avec l'obtention de bourses d'études (Algorithm Audit 2024a). L'agence a reconnu que « les étudiants issus de l'immigration hors Europe se voyaient attribuer un score de risque plus élevé et étaient plus souvent manuellement choisis pour une visite à domicile », créant ainsi des discriminations indirecte (Algorithm Audit 2024b). En Belgique, le système OASIS a été abandonné en 2023 après que des chercheurs ont révélé des risques de discrimination fondée sur la pauvreté. Ces systèmes regroupent des données provenant de différentes administrations (par exemple en matière de fiscalité, d'emploi, de prestations familiales, de pensions de retraite), parfois combinées à des données collectées par des entreprises privées (par exemple, sur la consommation d'énergie), afin d'établir un profil des personnes utilisatrices et de déterminer des indices de risque.

7. Cela se produit par exemple lorsque les forces de police conservent dans leurs systèmes des images de correspondances erronées sur une durée disproportionnée, ou lorsque la surveillance accrue de lieux largement habités par des communautés racisées alimente les outils prédictifs et renforce la surveillance de masse de ces communautés.

8. Le système SyRI a également été considéré comme ayant porté atteinte de manière disproportionnée au droit à la vie privée des utilisateurs finaux parce qu'il traitait des données à caractère personnel provenant de diverses agences gouvernementales.

Immigration et nationalité

Les systèmes d'IA et de PDA sont également utilisés par certaines administrations publiques en Europe pour contribuer à la prise de décisions dans le domaine de l'immigration, par exemple en matière de nationalité, d'asile ou de résidence. Les utilisations signalées comprennent, par exemple, l'identification et l'évaluation de la langue, la détection des fraudes liées aux documents d'identité, la gestion des dossiers, les interactions avec les migrantes et migrants, notamment par le biais de chatbots, les technologies de prévisions migratoires et de surveillance des frontières (réseau européen des migrations et Organisation de coopération et de développement économiques 2022; réseau européen des institutions nationales des droits de l'homme 2024; McGregor et Molnar 2023). Bien que ces technologies soient prétendument utilisées pour améliorer l'efficacité de la gestion des migrations, elles peuvent avoir un impact négatif sur les droits des migrantes et migrants en renforçant les discriminations, en suscitant des inquiétudes quant à la protection de la vie privée et des données, et en décourageant ces personnes en quête de protection (réseau européen des institutions nationales des droits de l'homme, 2014). Selon le Rapporteur spécial des Nations Unies sur les formes contemporaines de racisme, de discrimination raciale, de xénophobie et de l'intolérance qui y est associée, les « frontières numériques » englobent un éventail d'infrastructures de gouvernance qui « s'appuient de plus en plus sur l'apprentissage automatique, les mégadonnées, les systèmes décisionnels algorithmiques automatisés, l'analyse prédictive et les technologies numériques connexes » (Rapporteur spécial des Nations Unies, 2021). Les technologies utilisées comprennent « les systèmes d'identification, les systèmes de reconnaissance faciale, les capteurs au sol, les drones de vidéosurveillance aérienne, les bases de données biométriques et même les processus décisionnels en matière de visas et d'asile » (ibid.).

Autorités fiscales

Bien que l'on dispose généralement de moins d'informations sur l'utilisation des systèmes d'IA et de PDA par les autorités fiscales, plusieurs pays européens expérimentent actuellement ces technologies, principalement à des fins de détection des fraudes. Par exemple, des systèmes d'IA et de PDA sont utilisés pour signaler les dossiers suspects en vue de la réalisation d'une enquête plus approfondie. Les autorités peuvent être réticentes à révéler des informations sur ces systèmes de peur que les personnes utilisatrices puissent « déjouer » le système. Il existe des risques de discrimination car les systèmes de détection des fraudes peuvent impacter de manière disproportionnée les individus sur la base de motifs de discrimination protégés, tels que leur statut socioéconomique, leur origine migratoire ou leur situation géographique, même lorsque ces caractéristiques ne sont pas explicitement utilisées comme facteurs de risque dans ces systèmes. Dans l'affaire SyRI citée plus haut, aux Pays-Bas, les autorités fiscales soupçonnaient 26 000 foyers de fraude aux prestations sociales. L'algorithme avait signalé de manière disproportionnée les minorités ethniques et les ressortissant.es non néerlandais.es.

Secteur privé

Dans le secteur privé, les systèmes d'IA et de PDA sont utilisés par les agences de travail intérimaire. Ils servent notamment à la sélection et au profilage des personnes candidates à un emploi, à l'association des candidatures aux profils, ou encore à la rédaction et à la traduction des offres d'emploi. Ces utilisations peuvent présenter des risques de discrimination à des degrés divers. De même, le recours à des systèmes d'IA et de PDA est très répandu dans les domaines de la finance et du commerce, comme le marketing, le service client (via les chatbots, par exemple), la tarification dynamique et la surveillance des opérations. Par exemple, en Finlande, un client s'est vu refuser un crédit en ligne par une société de services financiers qui a utilisé un système de notation fondé sur des données telles que le sexe, la langue, l'âge et le lieu de résidence pour évaluer le risque de défaut de paiement d'un crédit (Algorithm Watch 2018). Le tribunal national finlandais chargé des questions de non-discrimination et d'égalité (2017) a conclu que cela équivalait à une discrimination directe fondée sur de multiples motifs protégés.

1.4. Cadres juridiques existants

Dans cette partie, nous analyserons la réglementation existante, les lois, la jurisprudence, les procédures, les politiques, institutions, etc., et examinerons comment elles traitent les problématiques de discrimination découlant du recours aux systèmes d'IA et de PDA.

1.4.1. Régir les systèmes d'IA et de PDA

Conseil de l'Europe

Au niveau du Conseil de l'Europe, plusieurs instruments abordent la question de la discrimination algorithmique ou peuvent être utilisés à cette fin.

Premièrement, la Convention-cadre du Conseil de l'Europe sur l'IA, adoptée en 2024 mais qui n'a pas encore été ratifiée pour entrer en vigueur, vise à « garantir que les activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle soient pleinement compatibles avec les droits humains, la démocratie et l'État de droit ». Dans son préambule, la Convention-cadre sur l'IA reconnaît que l'IA peut « promouvoir... l'égalité des sexes et l'autonomisation de toutes les femmes et filles ». Dans le même temps, elle fait état des préoccupations concernant « les risques de discrimination dans les cadres numériques, en particulier ceux impliquant des systèmes d'intelligence artificielle, et par le fait que ceux-ci sont susceptibles de créer des inégalités ou de les aggraver, y compris les inégalités vécues par les femmes et les personnes en situation de vulnérabilité, en ce qui concerne la jouissance de leurs droits humains et leur participation pleine, égale et effective aux affaires économiques, sociales, culturelles et politiques ».

À l'article 10 sur l'égalité et la non-discrimination, la Convention-cadre intime chaque partie à « adopte[r] ou maint[enir] des mesures visant à garantir le respect de l'égalité, y compris l'égalité de genre, et l'interdiction de la discrimination dans

les activités menées dans le cadre du cycle de vie⁹ des systèmes d'intelligence artificielle, comme le prévoit le droit international et interne applicable ». Les parties s'engagent également à « adopter ou à 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 ». Ces obligations exigent que les États membres réexaminent et, si nécessaire, réforment leur législation pour s'assurer que leurs lois sur la non-discrimination englobent la discrimination algorithmique. Comme le confirme le point 77 de l'exposé des motifs, la Convention prévoit également une obligation pour les États d'adopter des mesures pour parer aux inégalités structurelles et historiques en ce qui concerne les activités menées dans le cadre du cycle de vie des systèmes d'IA. Les points 72-73 à 40 de l'exposé des motifs confirment que l'interdiction de la discrimination algorithmique énoncée dans la Convention repose sur une approche de la discrimination algorithmique fondée sur les droits humains qui intègre les cadres mondiaux et régionaux des droits humains. Les points 75-76 à 40 de l'exposé des motifs dressent une liste de sources et de types de biais algorithmiques connus. Parmi les « différentes façons dont la partialité peut être incorporée intentionnellement ou par inadvertance dans les systèmes d'intelligence artificielle à différents stades de leur cycle de vie », l'exposé des motifs cite :

- ▶ développement (« en raison des stéréotypes ou des biais conscients ou inconscients des développeurs »), modélisation (« biais potentiels intégrés dans le modèle sur lequel les systèmes sont construits »);
- ▶ données (ensembles de données inexacts ou insuffisamment représentatifs aux stades de l'entraînement, de l'agrégation ou de l'évaluation);
- ▶ déploiement (« biais introduits lorsque ces systèmes sont mis en œuvre dans le monde réel »);
- ▶ interprétation (« biais d'automatisation ou de confirmation, par lesquels les humains peuvent accorder une plus grande confiance aux machines et aux artefacts technologiques ou dans des situations où ils sélectionnent des informations qui confirment leurs propres opinions, en ignorant dans les deux cas leur propre jugement potentiellement contradictoire et en validant les résultats algorithmiques sans les remettre en question »);
- ▶ biais technique (« problèmes liés à l'application de l'apprentissage automatique entraînant des biais supplémentaires qui ne sont pas présents dans les données utilisées pour former le système ou prendre des décisions »);

9. Voir points 14-15 du rapport explicatif de la Convention-cadre : « Cette référence au cycle de vie garantit une approche globale des risques liés à l'IA et des effets négatifs sur les droits de l'homme, la démocratie et l'État de droit, en tenant compte de toutes les étapes des activités liées aux systèmes d'intelligence artificielle. » Voici quelques exemples d'activités pertinentes : « (1) planification et conception, (2) collecte et traitement des données, (3) développement de systèmes d'intelligence artificielle, y compris l'élaboration de modèles et/ou l'adaptation de modèles existants à des tâches spécifiques, (4) essais, vérification et validation, (5) fourniture/mise à disposition des systèmes, (6) déploiement, (7) exploitation et suivi, et (8) mise hors service. »

- biais social (« incapacité à prendre correctement en compte les inégalités historiques ou contemporaines dans la société dans les activités menées dans le cadre du cycle de vie de l'intelligence artificielle »).

Pour faciliter le respect des obligations exposées dans la Convention, le Conseil de l'Europe a mis au point la méthodologie HUDERIA, qui vise à garantir « une approche uniforme de l'identification, de l'analyse et de l'évaluation des risques et de l'impact [des systèmes d'IA] sur la jouissance des droits de l'homme, le fonctionnement de la démocratie et le respect de l'État de droit. » (Conseil de l'Europe 2022a). Cette méthodologie consiste en une analyse des risques fondée sur le contexte, avec un processus de mobilisation des parties prenantes, une évaluation des risques et des impacts et un plan d'atténuation, et exige un examen itératif. Bien que l'approche diffère de la classification fondée sur le risque adoptée en vertu du règlement européen sur l'IA (voir ci-dessous la partie sur l'Union européenne), la méthodologie HUDERIA introduit des éléments similaires à travers une « approche graduelle et différenciée des mesures d'identification, d'évaluation, de prévention et d'atténuation des risques et des impacts, qui tienne compte de la gravité et de la probabilité d'occurrence des impacts négatifs sur les droits humains, la démocratie et l'État de droit, ainsi que des facteurs contextuels pertinents » (Conseil de l'Europe 2024b). Par exemple, la détermination des niveaux de risque doit tenir compte de l'ampleur des préjudices potentiels, de leur gravité, de leur réversibilité et de leur probabilité. Les « questions zéro », qui concernent la pertinence et l'adéquation des systèmes d'IA et de PDA pour effectuer certaines tâches, doivent également être prises en considération, avec la conséquence potentielle que des décisions de ne pas développer ou déployer de tels systèmes pourraient être prises si les risques de violations des droits humains l'emportent sur les avantages potentiels. La méthodologie HUDERIA d'évaluation des risques et de l'impact vise à assurer une compatibilité transparente « avec les pratiques de conformité existantes suivies par le secteur » (Conseil de l'Europe 2022a) et peut favoriser la prévention de la discrimination algorithmique par les acteurs publics et privés. Bien que l'article 16 de la convention prévoie la mise en place d'un cadre de gestion des risques et de l'impact, le respect de la méthodologie HUDERIA n'est pas obligatoire pour satisfaire aux obligations de la Convention. Bien que la méthodologie HUDERIA ne soit pas un instrument juridiquement contraignant et précise qu'elle ne fournit pas d'orientations interprétatives par rapport à la Convention, elle constituera un outil particulièrement utile pour identifier, évaluer et atténuer les risques de discrimination algorithmique.

En outre, diverses recommandations ont été publiées pour demander aux États parties d'aborder la discrimination algorithmique dans leurs cadres législatifs et procéduraux. Par exemple, l'Assemblée parlementaire, dans sa résolution 2343 (2020) pour « Prévenir les discriminations résultant de l'utilisation de l'intelligence artificielle », a appelé les États membres à :

revoir leur législation anti-discrimination et la modifier si nécessaire afin de veiller à ce qu'elle couvre tous les cas où une discrimination directe ou indirecte, y compris une discrimination par association, peut être causée par l'utilisation de l'IA, et à ce que les plaignants aient pleinement accès à la justice. À cet égard, il s'agit de veiller tout particulièrement à garantir la présomption d'innocence et à ce que les victimes

de discrimination ne soient pas soumises à une charge de la preuve disproportionnée (Conseil de l'Europe, Assemblée parlementaire, 2020a).

En outre, la résolution demandait aux États membres de « veiller à ce que les organismes de promotion de l'égalité soient pleinement habilités à traiter les problèmes concernant les principes d'égalité et de non-discrimination résultant d'une utilisation de l'IA ». Elle exigeait également que les gouvernements soient invités à « informer à l'avance le parlement de l'utilisation de ces technologies [d'IA et de PDA] » et à « à rendre obligatoire l'inscription systématique dans un registre public de chaque utilisation de ces technologies par les autorités ». Il est à noter que nombre de ces recommandations détaillées ne sont pas prises en compte dans la Convention-cadre sur l'IA récemment adoptée. Les travaux en cours du Comité directeur sur l'anti-discrimination, la diversité et l'inclusion et l'égalité des sexes portant sur une recommandation du Comité des Ministres sur l'égalité dans l'IA pourraient remédier à ces lacunes (Conseil de l'Europe, à paraître).

D'autres instruments publiés par les institutions du Conseil de l'Europe peuvent être utilisés pour lutter contre la discrimination algorithmique. Par exemple, la Recommandation générale n°1 sur la dimension numérique de la violence à l'égard des femmes adoptée par le Groupe d'expert.es sur la lutte contre la violence à l'égard des femmes et la violence domestique (GREVIO) aborde spécifiquement les formes discriminatoires de violence en ligne telles que le harcèlement sexuel en ligne, le harcèlement en ligne et facilité par la technologie et les dimensions numériques de la violence psychologique (Conseil de l'Europe 2021a).

Des recommandations ou instruments plus transversaux axés sur des sujets connexes méritent également d'être mentionnés :

- ▶ Recommandation 2102 (2017) de l'Assemblée parlementaire du Conseil de l'Europe sur « La convergence technologique, l'intelligence artificielle et les droits de l'homme » (Conseil de l'Europe, Assemblée parlementaire 2017);
- ▶ La Déclaration du Comité des Ministres sur les risques présentés par le suivi numérique et les autres technologies de surveillance pour les droits fondamentaux (Conseil de l'Europe 2013);
- ▶ La Recommandation du Commissaire aux droits de l'homme sur « Décoder l'intelligence artificielle : 10 mesures pour protéger les droits de l'homme » (Conseil de l'Europe, Commissaire aux droits de l'homme 2019);
- ▶ La Déclaration du Comité des Ministres sur les capacités de manipulation des processus algorithmiques (Conseil de l'Europe 2019a);
- ▶ Recommandation CM/Rec (2020) 1 du Comité des Ministres aux États membres sur les impacts des systèmes algorithmiques sur les droits de l'homme (Conseil de l'Europe 2020);
- ▶ Recommandation CM/Rec(2022)13 du Comité des Ministres aux États membres sur les impacts des technologies numériques sur la liberté d'expression (Conseil de l'Europe 2022b);
- ▶ La note d'orientation adoptée par le Comité directeur sur les médias et la société de l'information, « Modération du contenu. Meilleures pratiques en vue de la mise en place de cadres juridiques et procéduraux efficaces pour

les mécanismes d'autorégulation et de corégulation de la modération de contenu » (Conseil de l'Europe 2021b);

- La Recommandation de suivi du Commissaire aux droits de l'homme sur « Les droits de l'homme dès la conception de l'IA : une protection durable des droits humains à l'ère de l'intelligence artificielle » (Conseil de l'Europe, Commissaire aux droits de l'homme 2023).

Ces différents instruments offrent des orientations importantes en matière de respect des droits humains, dont le droit à la non-discrimination, dans le contexte de l'utilisation des systèmes d'IA et de PDA. Toutefois, leur caractère non contraignant et leurs restrictions d'applicabilité limitent leur efficacité dans la pratique.

Union européenne

Au niveau de l'UE, la discrimination algorithmique fait l'objet de plusieurs dispositions du règlement européen sur l'IA (Union européenne 2024c).¹⁰ Le règlement européen sur l'IA vise à compléter les cadres existants de lutte contre la discrimination et de protection des données. Il adopte une approche fondée sur les risques, avec un certain nombre de pratiques interdites et de systèmes à haut risque qui sont soumis à des exigences spécifiques. Le règlement repose sur la reconnaissance que « [S]i l'IA peut être utilisée à de nombreuses fins positives, elle peut aussi être utilisée à mauvais escient et fournir des outils nouveaux et puissants à l'appui de pratiques de manipulation, d'exploitation et de contrôle social. De telles pratiques sont particulièrement néfastes et abusives et devraient être interdites, car elles sont contraires aux valeurs de l'Union [...] y compris le droit à la non-discrimination. » (considérant 28).

D'autres systèmes sont considérés comme présentant un risque élevé parce qu'ils « peuvent porter atteinte au droit à la dignité et à la non-discrimination et sont contraires aux valeurs d'égalité et de justice » (considérant 31). Une troisième catégorie de systèmes d'IA, considérés comme à faible risque, sont beaucoup plus faiblement réglementés. En d'autres termes, « [l']ampleur de l'incidence négative du système d'IA sur les droits fondamentaux protégés par la Charte est un critère particulièrement pertinent lorsqu'il s'agit de classer un système d'IA en tant que système à haut risque » (considérant 48). Cela dit, le considérant 27 du règlement rappelle que les sept principes suivants, définis par le Groupe d'experts de haut niveau sur l'IA, sont applicables à tous les systèmes, quel que soit leur niveau de risque : « action humaine et contrôle humain; robustesse technique et sécurité; respect de la vie privée et gouvernance des données; transparence; diversité, non-discrimination et équité; bien-être sociétal et environnemental; et responsabilité ». Dans ce contexte, « [l]a diversité, la non-discrimination et l'équité renvoient au fait que les systèmes d'IA sont développés et utilisés de manière à inclure des acteurs divers et à promouvoir l'égalité d'accès, l'égalité de genre et la diversité culturelle, tout en évitant les effets discriminatoires et les biais injustes, qui sont interdits par le droit de l'Union ou le droit national » (ibid.). Ce principe est censé s'appliquer à tous les systèmes d'IA mis sur le marché de l'UE, quelle que soit leur catégorisation des risques.

10. Voir également l'exposé des motifs et le Pacte sur l'IA qui l'accompagne, disponibles à l'adresse suivante : <https://digital-strategy.ec.europa.eu/fr/policies/ai-pact>, consulté le 10 novembre 2025.

Systèmes d'IA interdits

La liste des systèmes interdits, particulièrement pertinents dans le contexte de la discrimination, est énoncée à l'article 5 du règlement européen sur l'IA. L'article 5 a) interdit les « techniques subliminales [...] ou délibérément manipulatoires ou trompeuses [qui ont] pour objectif ou effet d'altérer substantiellement le comportement d'une personne ou d'un groupe de personnes en portant considérablement atteinte à leur capacité de prendre une décision éclairée [et] cause ou est raisonnablement susceptible de causer un préjudice important à cette personne ». L'article 5 b) interdit les systèmes d'IA qui exploite « les éventuelles vulnérabilités dues à l'âge, au handicap ou à la situation sociale ou économique spécifique d'une personne physique ou d'un groupe de personnes donné » de manière à altérer leur comportement ou de leur porter préjudice.

L'article 5 c) interdit les systèmes de notation sociale qui évaluent ou classent des personnes ou des groupes « en fonction de leur comportement social ou de caractéristiques personnelles ou de personnalité connues, déduites ou prédites » dans le but de les traiter de manière défavorable dans des contextes sociaux sans rapport avec eux ou de manière injustifiée ou disproportionnée. L'article 5 d) interdit l'utilisation d'un « système d'IA pour évaluer les risques des personnes physiques pour mener des évaluations des risques des personnes physiques visant à évaluer ou à prédire le risque qu'une personne physique commette une infraction pénale, uniquement sur la base du profilage d'une personne physique ou de l'évaluation de ses traits de personnalité ou caractéristique ». L'article 5 e), interdit les systèmes de reconnaissance faciale sur la base d'un moissonnage d'images provenant d'Internet ou d'images de vidéosurveillance.

L'article 5 f) interdit l'utilisation de systèmes de reconnaissance des émotions sur le lieu de travail et dans les établissements d'enseignement. Selon l'article 5, points g) et h), « l'utilisation de systèmes de catégorisation biométrique qui catégorisent individuellement les personnes physiques sur la base de leurs données biométriques afin d'arriver à des déductions ou des inférences concernant leur race, leurs opinions politiques, leur affiliation à une organisation syndicale, leurs convictions religieuses ou philosophiques, leur vie sexuelle ou leur orientation sexuelle » est interdite, de même que « l'utilisation de systèmes d'identification biométrique à distance en temps réel dans des espaces accessibles au public à des fins répressives », sauf si cela s'avère strictement nécessaire.

Ces interdictions couvrent des risques importants de discrimination algorithmique par une interdiction générale. Pourtant, les critiques ont mis en évidence des lacunes importantes en ce qui concerne les exceptions générales accordées aux services de police et de répression, en particulier en ce qui concerne la reconnaissance faciale en direct et la surveillance biométrique visées à l'article 5, points g) et h) (droits numériques européens 2024). La Commission européenne a publié un ensemble de lignes directrices sur les pratiques interdites en matière d'IA, qui clarifient la portée de certaines dispositions du règlement européen sur l'IA. Ces lignes directrices stipulent ce qui suit :

Même lorsqu'un système d'IA n'est pas interdit par le règlement européen sur l'IA, son utilisation pourrait être interdite ou illégale au titre d'un autre droit primaire ou

dérivé de l'Union (par exemple, au titre du non-respect des droits fondamentaux dans un cas donné, comme l'absence de base juridique pour le traitement des données à caractère personnel requis par la législation sur la protection des données, ou en cas de discrimination interdite par le droit de l'Union, etc.) (Commission européenne 2025a, paragraphe 43).

Systèmes d'IA à haut risque

L'annexe III du règlement européen sur l'IA énumère les IA dites à haut risque utilisées dans les domaines de la biométrie, des infrastructures critiques (infrastructures numériques critiques, circulation routière ou approvisionnement en eau, gaz, chauffage ou électricité), de l'éducation et de la formation professionnelle, de l'emploi, de la gestion des travailleuses, des travailleurs et de l'accès au travail indépendant, de l'accès aux services privés essentiels et aux services publics et prestations essentiels, des services répressifs de l'immigration, de l'asile et de la gestion des contrôles aux frontières, ainsi que de la justice et des processus démocratiques. Ces systèmes à haut risque sont soumis aux exigences légales spécifiques énumérées ci-dessous. L'article 6, point 3, prévoit qu'un système d'IA utilisé dans une zone énumérée à l'annexe III « n'est pas considéré comme à haut risque lorsqu'il ne présente pas de risque important de préjudice pour la santé, la sécurité ou les droits fondamentaux des personnes physiques, y compris en n'ayant pas d'incidence significative sur le résultat de la prise de décision ». Il s'agit notamment des systèmes qui accomplissent « une tâche procédurale étroite » ou une « tâche [purement] préparatoire » et des systèmes « destiné[s] à améliorer le résultat d'une activité humaine préalablement réalisée » ou « à détecter les constantes en matière de prise de décision ou les écarts par rapport aux constantes habituelles antérieures ou des écarts par rapport à des schémas décisionnels antérieurs et [qui ne sont] pas destiné[s] à remplacer ou influencer l'évaluation humaine préalablement réalisée, sans examen humain approprié ». Cette disposition est susceptible d'exclure du champ d'application de l'article 6 un éventail potentiellement vaste de systèmes d'IA utilisés dans des domaines à haut risque et pourrait être utilisée à mauvais escient pour éviter de se conformer aux exigences énoncées par le règlement européen sur l'IA concernant les systèmes à haut risque (voir section 2.5.). Toutefois, ces systèmes restent soumis à certaines obligations : les fournisseurs qui considèrent un système d'IA comme n'étant pas à haut risque doivent néanmoins documenter l'évaluation avant que ce système ne soit mis sur le marché ou mis en service et doivent fournir cette documentation aux autorités nationales compétentes sur demande. Ils doivent également enregistrer le système d'IA en question dans la base de données de l'UE prévue à l'article 71 du règlement européen sur l'IA.

Modèles d'IA à usage général

Les modèles d'IA à usage général, définis à l'article 3, paragraphe 63, comme des modèles d'IA « qui présente[nt] une généralité significative et [sont] capable[s] d'exécuter de manière compétente un large éventail de tâches distinctes, indépendamment de la manière dont [les modèles sont] mis sur le marché, et qui peu[ven]t être intégré[s] dans une variété de systèmes ou d'applications en aval », sont spécifiquement réglementés par l'article 53 du règlement européen sur l'IA. Les fournisseurs

doivent élaborer et présenter une documentation technique et des instructions d'utilisation, se conformer à la législation de l'UE sur le droit d'auteur.e et publier un résumé détaillant les données utilisées pour l'entraînement du système. Les modèles d'IA à usage général d'une certaine capacité et d'une certaine échelle sont classés comme présentant un « risque systémique » en vertu de l'article 51 du règlement européen sur l'IA, et sont soumis à des exigences supplémentaires en vertu de l'article 55 du règlement européen sur l'IA. Les fournisseurs des systèmes de cette catégorie doivent, entre autres, procéder à des évaluations de modèles et à des essais contradictoires, mais également surveiller et signaler les incidents graves. De plus, lorsque un modèle d'IA à usage général est intégré dans un système d'IA, ce système est également soumis aux exigences légales applicables à sa catégorie de risque (inacceptable, élevé, limité ou minimal).

Systèmes d'IA présentant des risques limités

En revanche, les utilisations présentant des risques plus limités sont uniquement soumises à des exigences de transparence. Par exemple, les déployeurs de chatbots doivent informer les utilisatrices et utilisateurs qu'ils interagissent avec un système d'IA et les déployeurs de modèles d'IA à usage général produisant du texte, des images, des vidéos ou du son doivent indiquer que leurs résultats ont été générés artificiellement, conformément à l'article 50 du règlement européen sur l'IA. Les autres applications d'IA sont considérées comme ne présentant aucun risque ou des risques minimes, et ne sont pas réglementées.

Calendrier de conformité

Le règlement européen sur l'IA est un règlement, de sorte qu'aucune transposition n'est nécessaire de la part des États membres : les dispositions s'appliquent directement. Cela dit, il y a un certain nombre d'étapes à accomplir pour que le règlement entre en vigueur. L'interdiction des pratiques établie à l'article 5 est entrée en vigueur en février 2025 et la Commission a publié des orientations sur ces interdictions (Commission européenne, 2025a). En 2025, le Bureau de l'IA de l'UE a publié un code de bonnes pratiques pour les modèles d'IA à usage général, y compris ceux qui présentent des risques systémiques (Commission européenne 2025b), et devrait élaborer des lignes directrices pour la classification des systèmes d'IA à haut risque, ainsi qu'un modèle pour les évaluations d'impact des systèmes d'IA à haut risque sur les droits fondamentaux (voir ci-dessous la partie sur les évaluations d'impact sur les droits fondamentaux). D'ici août 2025, les États membres devraient avoir désigné les autorités nationales compétentes (au moins une autorité notifiante et au moins une autorité de surveillance du marché) chargées de superviser la mise en œuvre du règlement européen sur l'IA au niveau national. En 2026, le Bureau européen de l'IA devrait se prononcer sur les normes techniques élaborées par les organismes de normalisation de l'UE, le Comité européen de normalisation (CEN) et le Comité européen de normalisation électrotechnique (CENELEC) en ce qui concerne les systèmes de gestion des risques et la prévention des biais dans les systèmes d'IA à haut risque (voir ci-dessous la partie sur le CEN-CENELEC). Enfin, les règles relatives aux systèmes à haut risque devraient entrer en vigueur en août 2026, ou en août 2027 pour les systèmes d'IA utilisés comme produits, ou leurs composants de sécurité, qui

doivent faire l'objet d'une évaluation de conformité réalisée par un tiers en vertu de la législation d'harmonisation de l'Union exposée à l'annexe I du règlement européen sur l'IA. Pour les systèmes d'IA à haut risque déjà mis sur le marché ou mis en service au plus tard en août 2026 et faisant l'objet d'« importantes modifications de leur conception » ou destinés à être utilisés par les pouvoirs publics, les fournisseurs et les déployeurs doivent prendre les mesures nécessaires pour se conformer aux exigences et obligations du règlement européen sur l'IA au plus tard en août 2030 (article 111, paragraphe 2). Toutes ces mesures contribueront à équiper les États membres pour prévenir et corriger les discriminations algorithmiques.

Gouvernance des données

Certaines dispositions spécifiques du règlement européen sur l'IA ont une pertinence directe pour la lutte contre la discrimination algorithmique, notamment parce qu'elles énoncent des exigences spécifiques pour les systèmes à haut risque. L'article 10 sur les données et la gouvernance des données établit des critères de qualité pour l'entraînement, la validation et les essais des systèmes à haut risque énumérés à l'annexe III. En particulier, l'article 10, paragraphe 2, points f) et g), dispose que les fournisseurs de systèmes d'IA et de PDA procèdent à un « examen permettant de repérer d'éventuels biais qui sont susceptibles de porter atteinte à la santé et à la sécurité des personnes, d'avoir une incidence négative sur les droits fondamentaux ou de se traduire par une discrimination interdite par le droit de l'Union, en particulier lorsque les données de sortie influencent les entrées pour les opérations futures » et prennent « les mesures appropriées visant à détecter, prévenir et atténuer les éventuels biais repérés conformément au point f) ». À ces fins, les catégories sensibles de données à caractère personnel, dont le RGPD et le règlement européen sur l'IA interdisent généralement l'usage, peuvent être traitées au titre de l'exception prévue à l'article 10, paragraphe 5 :

Dans la mesure où cela est strictement nécessaire aux fins de la détection et de la correction des biais en ce qui concerne... des systèmes d'IA à haut risque [et sous certaines conditions¹¹],

-
11. Conditions applicables : « a) la détection et la correction des biais ne peuvent pas être satisfaites de manière efficace en traitant d'autres données, y compris des données synthétiques ou anonymisées; b) les catégories particulières de données à caractère personnel sont soumises à des limitations techniques relatives à la réutilisation des données à caractère personnel, ainsi qu'aux mesures les plus avancées en matière de sécurité et de protection de la vie privée, y compris la pseudonymisation; c) les catégories particulières de données à caractère personnel font l'objet de mesures visant à garantir que les données à caractère personnel traitées sont sécurisées, protégées et soumises à des garanties appropriées, y compris des contrôles stricts et une documentation de l'accès, afin d'éviter toute mauvaise utilisation et de veiller à ce que seules les personnes autorisées ayant des obligations de confidentialité appropriées aient accès à ces données à caractère personnel; d) les catégories particulières de données à caractère personnel ne doivent pas être transmises, transférées ou consultées d'une autre manière par d'autres parties; e) les catégories particulières de données à caractère personnel sont supprimées une fois que le biais a été corrigé ou que la période de conservation des données à caractère personnel a expiré, selon celle de ces deux échéances qui arrive en premier; f) les registres des activités de traitement visés dans les règlements (UE) 2016/679 et (UE) 2018/1725 et dans la directive (UE) 2016/680 comprennent les raisons pour lesquelles le traitement des catégories particulières de données à caractère personnel était strictement nécessaire pour détecter et corriger les biais, ainsi que la raison pour laquelle cet objectif n'a pas pu être atteint par le traitement d'autres données. »

les fournisseurs de ces systèmes peuvent exceptionnellement traiter des catégories particulières de données à caractère personnel, sous réserve de garanties appropriées pour les libertés et droits fondamentaux des personnes physiques.

Cette exception peut être utile pour identifier les discriminations algorithmiques, mais elle ne s'étend pas aux systèmes qui ne sont pas à haut risque.

Système de gestion des risques

L'article 9 du règlement européen sur l'IA exige que les fournisseurs de systèmes d'IA à haut risque mettent en place un système de gestion des risques. Il stipule qu'« un système de gestion des risques [doit être] établi, mis en œuvre, documenté et tenu à jour en ce qui concerne les systèmes d'IA à haut risque » et que la gestion des risques doit être comprise comme « un processus itératif continu qui est planifié et se déroule sur l'ensemble du cycle de vie d'un système d'IA à haut risque ». En particulier, les fournisseurs sont invités à identifier et analyser les « risques connus et raisonnablement prévisibles que le système d'IA à haut risque peut poser pour la santé, la sécurité ou les droits fondamentaux lorsque le système d'IA à haut risque est utilisé conformément à sa destination ». Ils doivent estimer et évaluer « les risques susceptibles d'apparaître lorsque le système d'IA à haut risque est utilisé conformément à sa destination et dans des conditions de mauvaise utilisation raisonnablement prévisible », ainsi que les « autres risques susceptibles d'apparaître, sur la base de l'analyse des données recueillies au moyen du système de surveillance après commercialisation visé à l'article 72 ». Enfin, ils doivent adopter des « mesures appropriées et ciblées de gestion des risques pour répondre aux risques identifiés ». Ces obligations seront mises en œuvre au moyen de normes techniques élaborées par le CEN-CENELEC. Par conséquent, la norme de gestion des risques qui sera publiée par le CEN-CENELEC constituera un dispositif essentiel pour gérer les risques pour les droits fondamentaux, et notamment pour la prévention et la protection contre la discrimination algorithmique. Même s'il n'est pas obligatoire, le respect des normes déclenchera une présomption de conformité des systèmes à haut risque conformément à l'article 40 du règlement, et le secteur sera ainsi incité à suivre ces normes. Néanmoins, il est important de mentionner que cette présomption de conformité se limite exclusivement aux exigences du règlement européen sur l'IA pour les systèmes d'IA à haut risque, et ne s'étend pas au droit européen des droits fondamentaux ou au droit anti-discrimination.

Autres exigences

Ces dispositions sont complétées par plusieurs obligations complémentaires pour les fournisseurs de systèmes d'IA à haut risque, telles que l'article 11 sur la documentation technique, l'article 12 sur la tenue de registres, l'article 13 sur la transparence, l'article 14 sur la surveillance humaine, l'article 15 sur l'exactitude, la robustesse et la cybersécurité, et l'article 17 sur les systèmes de gestion de la qualité. Par exemple, l'article 15, paragraphe 4, exige que les fournisseurs de systèmes à haut risque « qui continuent leur apprentissage après leur mise sur le marché ou leur mise en service sont développés de manière à éliminer ou à réduire dans la mesure du possible le risque que des sorties éventuellement biaisées n'influencent les entrées pour les opérations futures (boucles de rétroaction) et à veiller à ce que ces boucles de

rétroaction fassent l'objet d'un traitement adéquat au moyen de mesures d'atténuation appropriées ». Ces obligations peuvent faciliter l'application des dispositions anti-discriminatoires existantes.

Base de données de l'UE

L'article 49, paragraphes 1 et 2, oblige les fournisseurs à enregistrer les systèmes d'IA dans la base de données de l'UE visée à l'article 71 du règlement. En outre, l'article 49, paragraphe 3, exige que « les déployeurs qui sont des autorités publiques, des institutions, organes ou organismes de l'Union ou des personnes agissant en leur nom s'enregistrent, sélectionnent le système et enregistrent son utilisation ». Lorsque des systèmes d'IA sont utilisés dans les domaines des services répressifs, de la migration, de l'asile et de la gestion des contrôles aux frontières, cet enregistrement doit être effectué dans une section non publique de la base de données de l'UE à laquelle seules la Commission européenne et les autorités nationales peuvent accéder. Selon l'article 71, le reste de la base de données doit être « accessible et mise à la disposition du public d'une manière conviviale ». Ce registre public facilitera les enquêtes sur les impacts discriminatoires des systèmes d'IA et de PDA, y compris ceux déployés par les administrations publiques, ainsi que la tâche des victimes potentielles consistant à établir des preuves *prima facie* attestant les discriminations subies. Pourtant, bien qu'il s'agisse d'un bon point de départ, la publicité par le biais d'une base de données à l'échelle de l'UE ne suffira pas à parer aux utilisations discriminatoires des systèmes d'IA et de PDA. En outre, des organisations de la société civile ont exprimé des préoccupations quant au manque d'accès public aux informations relatives aux domaines des activités répressives, de la migration, de l'asile et de la gestion du contrôle aux frontières contenues dans le registre public (Access Now 2024).

Analyses d'impact sur les droits fondamentaux

L'article 27 exige que certains déployeurs, principalement ceux qui sont des organismes de droit public ou des entités privées fournissant des services publics, ainsi que les déployeurs de systèmes d'IA destinés à être utilisés pour évaluer la solvabilité de personnes physiques ou établir leur score de crédit, ou de systèmes d'IA destinés à être utilisés à des fins d'évaluation des risques et de tarification en fonction des personnes physiques dans les domaines de l'assurance-vie et de l'assurance-maladie, réalisent une analyse d'impact sur les droits fondamentaux (AIDF) avant toute utilisation de systèmes d'IA à haut risque. Ce type d'analyse consiste en une description des processus du déployeur; de la période et de la fréquence d'utilisation du système d'IA; des catégories de personnes physiques et de groupes susceptibles d'être impactés par l'utilisation du système d'IA dans le contexte concerné; des risques spécifiques de préjudice susceptibles de concerner ces groupes; des mesures de surveillance humaine prises; et de la manière dont ces risques sont traités, y compris par des mécanismes de gouvernance interne et de plainte. L'AIDF sera transmise aux autorités de surveillance du marché et un résumé devra être enregistré dans la base de données de l'UE (annexe IV, section C, point 4). L'accès aux AIDF facilitera la remise en question des systèmes discriminatoires d'IA et de PDA pour les individus et les organismes de promotion de l'égalité. En outre, conformément au considérant 96 :

Les déployeurs de systèmes d'IA à haut risque, en particulier lorsque des systèmes d'IA sont utilisés dans le secteur public, pourraient associer les parties prenantes concernées, y compris les représentants de groupes de personnes susceptibles d'être concernés par le système d'IA, les experts indépendants et les organisations de la société civile, à la réalisation de cette analyse d'impact et à la conception des mesures à prendre en cas de matérialisation des risques.

Les AIDF sont complétées par l'article 60, qui établit un cadre dans lequel les fournisseurs (avec ou sans déployeurs) peuvent tester des systèmes d'IA à haut risque en conditions réelles. Si les incidents graves doivent être signalés à l'autorité nationale de marché (article 60, paragraphe 7, et article 73) et immédiatement traités par des mesures d'atténuation, la mesure dans laquelle les parties prenantes concernées, telles que les organismes de promotion de l'égalité, peuvent accéder aux résultats des tests et aux mesures d'atténuation n'est pas claire et peut dépendre du contexte national. D'une part, l'article 71, paragraphe 4, dispose que les « informations [liées aux essais] enregistrées conformément à l'article 60 ne sont accessibles qu'aux autorités de surveillance du marché et à la Commission, sauf si le fournisseur ou fournisseur potentiel a donné son consentement pour que ces informations soient également accessibles au public ». En fait, il exclut les informations relatives aux essais enregistrés en vertu de l'article 60, paragraphe 4, point c), des exigences d'accessibilité et de publicité. D'autre part, l'article 77, paragraphe 1, stipule ce qui suit :

Les autorités ou organismes publics nationaux qui supervisent ou font respecter les obligations au titre du droit de l'Union visant à protéger les droits fondamentaux, y compris le droit à la non-discrimination, en ce qui concerne l'utilisation des systèmes d'IA à haut risque visés à l'annexe III sont habilités à demander toute documentation créée ou conservée en vertu du présent règlement et à y avoir accès dans une langue et un format accessibles lorsque l'accès à cette documentation est nécessaire à l'accomplissement effectif de leur mandat dans les limites de leurs compétences. (italiques ajoutés).

Autorités de protection des droits fondamentaux (organismes relevant de l'article 77)

En fait, l'article 77 du règlement européen sur l'IA sur les « pouvoirs des autorités de protection des droits fondamentaux » est particulièrement important parce qu'il définit le cadre d'application. Comme expliqué ci-dessus, l'article 77(1) accorde aux autorités de surveillance des droits fondamentaux désignées par les États membres le pouvoir de demander et d'accéder à toute documentation créée ou conservée en vertu du règlement européen sur l'IA, si nécessaire. Ce faisant, il confère à un ensemble d'autorités de protection des droits fondamentaux le droit d'accéder à la documentation établie en vertu du règlement européen sur l'IA, comme les plans de gestion des risques, les analyses d'impact ou tout autre document nécessaire pour s'acquitter efficacement de leur mandat. La Commission européenne a publié une interprétation large de la liste des autorités compétentes. Elle ne se limite pas aux organismes généralement considérés comme des structures de défense des droits humains, tels que les organismes de promotion de l'égalité, les institutions nationales de défense des droits humains et les médiateurs, mais comprend également les autorités chargées de la protection des données, les autorités chargées de la protection des consommateurs, les autorités chargées de la protection de l'enfance, les autorités chargées du droit du travail et les autorités de contrôle des médias. Les

autorités auxquelles sont accordés des droits d'accès en vertu de l'article 77 doivent soit superviser, soit faire appliquer la législation pertinente de l'Union (ou les deux). La Commission estime que l'article 77 n'a pas pour objet de couvrir les juridictions nationales ou autres autorités judiciaires, agissant au titre de leurs fonctions judiciaires, qui jouissent déjà de droits d'accès en vertu de l'article 47 de la Charte des droits fondamentaux de l'Union européenne. Les autorités ou organismes publics nationaux visés à l'article 77, paragraphe 1, devraient être annoncés publiquement et notifiés à la Commission européenne par les États membres. Compte tenu des nouveaux pouvoirs confiés aux organismes de promotion de l'égalité (voir ci-dessous la partie sur le rôle des organismes de promotion de l'égalité), ces institutions pourraient être appelées à jouer un rôle majeur dans l'application du règlement européen sur l'IA en matière de discrimination algorithmique.

En outre, l'article 77, paragraphe 3, dispose que :

Lorsque la documentation visée au paragraphe 1 ne suffit pas pour déterminer s'il y a eu violation des obligations au titre du droit de l'Union protégeant les droits fondamentaux, l'autorité ou l'organisme public visé au paragraphe 1 peut présenter à l'autorité de surveillance du marché une demande motivée visant à organiser les tests du système d'IA à haut risque par des moyens techniques. L'autorité de surveillance du marché organise les tests avec la participation étroite de l'autorité ou organisme public ayant présenté la demande dans un délai raisonnable après celle-ci.

En outre, l'article 73, paragraphe 7, du règlement européen sur l'IA dispose que les autorités nationales désignées qui protègent les droits fondamentaux doivent être informées par une autorité de surveillance du marché qui a reçu une notification relative à un incident grave, y compris un cas de discrimination, en rapport avec un système d'IA à haut risque. Conformément à l'article 79, paragraphe 2, ces autorités doivent être informées par une autorité de surveillance du marché, qui doit coopérer pleinement avec elles, lorsque cette dernière identifie un risque pour les droits fondamentaux, y compris un cas de discrimination, en rapport avec un système d'IA présentant un risque. Les opérateurs concernés doivent également coopérer avec eux, si nécessaire. Conformément à l'article 82, paragraphe 1, les autorités nationales chargées de la protection des droits fondamentaux doivent être consultées par toute autorité de surveillance du marché qui constate qu'un système d'IA conforme à haut risque présente un risque pour les droits fondamentaux, tel qu'une discrimination, lorsque l'autorité ou l'organisme en question est compétent vis-à-vis du droit fondamental concerné.

Responsabilité de l'IA

En 2022, la Commission européenne a proposé une directive spécifique sur la responsabilité des IA qui visait à éliminer les obstacles importants à l'accès à la justice dans le contexte de l'IA. Elle proposait d'harmoniser les règles de responsabilité dans l'ensemble de l'UE en ce qui concerne les actions en droit civil extracontractuel fondées sur la faute au titre de préjudices liés à l'IA. Toutefois, la proposition de directive a été retirée en 2025, ce qui renforce l'urgence de traiter les questions de responsabilité en matière d'IA aux niveaux national et européen, y compris par d'autres moyens. Le projet de directive sur la responsabilité de l'IA reconnaissait que « les caractéristiques spécifiques de l'IA [...] sont problématiques pour les règles

existant en matière de responsabilité » et proposait d'« all[éger] la charge de la preuve de manière très ciblée et proportionnée en recourant à la divulgation et aux présomptions réfragables » afin de faciliter l'accès à la justice en cas de préjudices causés par l'IA, y compris en cas de discrimination algorithmique (Commission européenne 2022a). En particulier, l'article 3 de la proposition retirée indiquait qu'une juridiction pouvait ordonner aux fournisseurs ou aux déployeurs de systèmes d'IA à haut risque soupçonnés d'avoir causé un préjudice de divulguer des éléments de preuve lorsque cela était nécessaire dans le cadre d'une procédure judiciaire. Le non-respect d'une demande de divulgation ou de conservation d'éléments de preuve aurait déclenché une présomption réfragable de non-respect de la part du ou de la défendeur.se. L'exposé des motifs de l'article 4 de la proposition retirée reconnaissait la possibilité que le ou la demandeur.se puisse « éprouver des difficultés à établir un lien de causalité entre ce non-respect et le résultat du système d'IA [...] à l'origine du dommage en question », comme la discrimination (ibid.). Par conséquent, le non-respect d'une obligation de diligence prévue par le règlement européen sur l'IA ou d'autres règles de l'UE devait être compris comme une faute et aurait déclenché une présomption réfragable de lien de causalité entre le préjudice et la faute. Par exemple, si le fournisseur d'un système d'IA à haut risque n'avait pas mis en place de façon adéquate un système de gestion des risques en vertu du règlement européen sur l'IA, un tribunal aurait pu présumer qu'il y avait un lien de causalité entre ce manquement à l'obligation de diligence et le préjudice causé par ce système, par exemple sous forme de discrimination. Ces règles auraient pu faciliter l'établissement d'une preuve *prima facie* de discrimination algorithmique devant un tribunal lorsque le défendeur n'avait pas respecté ses obligations en vertu du règlement européen sur l'IA. Toutefois, trois conditions devaient être remplies : le ou la demandeur.se aurait dû démontrer un manquement à l'obligation de diligence, il aurait dû être raisonnablement probable qu'une telle faute ait eu une influence sur le résultat nocif du système d'IA, et le ou la demandeur.se aurait dû démontrer que le résultat produit par le système d'IA a causé le dommage. Pour les particuliers victimes de discrimination, l'établissement de ces éléments aurait déjà représenté des obstacles importants. Par conséquent, des réflexions sur l'ajustement des règles existantes sur les présomptions de discrimination et le transfert de la charge de la preuve doivent avoir lieu dans le contexte des systèmes d'IA et de PDA.

Autres textes législatifs et lignes directrices de l'UE

D'autres actes législatifs peuvent être utilisés pour lutter contre la discrimination algorithmique au niveau de l'UE. Par exemple, l'article 34 du règlement sur les services numériques, relatif à l'évaluation des risques par les fournisseurs de très grandes plateformes en ligne et de très grands moteurs de recherche en ligne, inclut les risques de discrimination (Union européenne, 2022). Il en va de même de

l'article 35 sur l'atténuation de ces risques.¹² En outre, la Commission européenne a récemment lancé une consultation publique sur le règlement sur l'équité numérique, qui visera à renforcer la protection et l'équité numérique pour les consommateurs et consommatrices.

L'UE a également adopté d'importantes politiques sur la gouvernance des systèmes d'IA et de PDA qui peuvent être utilisées par les organismes de promotion de l'égalité, par exemple pour formuler des recommandations politiques aux décideurs ou pour les sensibiliser. Par exemple, la Déclaration européenne sur les droits et principes numériques pour la Décennie numérique adoptée en 2022 propose une déclaration des droits numériques (Commission européenne, 2022b). Elle exprime un engagement à « veiller à ce que les systèmes algorithmiques reposent sur des ensembles de données appropriés, afin d'éviter toute discrimination et de permettre une surveillance humaine de tous les résultats qui affectent la sécurité et les droits fondamentaux des citoyens ». Le Livre blanc de 2020 de la Commission européenne (2020) traite également de la discrimination algorithmique.

1.4.2. Instruments de lutte contre la discrimination

Bien que les règles spécifiques à l'IA soient de nouveaux outils importants pour les organismes de promotion de l'égalité et les autres organisations luttant contre la discrimination, l'activation de la législation anti-discrimination et de la législation sur la protection des données ainsi que ces nouvelles dispositions seront essentielles pour lutter contre la discrimination algorithmique. Les cadres juridiques spécifiques à l'IA complètent, mais ne remplacent en aucun cas, la législation sur l'égalité. Les exigences du règlement européen sur l'IA pour les systèmes d'IA à haut risque, en particulier, offrent un ensemble de garanties techniques qui favorisent, sans pouvoir garantir, le respect de la législation anti-discrimination. Dans ce contexte, les organismes de promotion de l'égalité joueront un rôle important dans l'application des cadres juridiques existants en matière d'égalité à la discrimination algorithmique. Par exemple, ils peuvent donner des conseils éclairés sur l'application concrète des règles relatives aux présomptions de discrimination et au transfert de la charge de la preuve en cas d'asymétrie d'information. Ils peuvent également jouer un rôle important en ce qui concerne les qualifications juridiques des préjudices algorithmiques à l'égard des cas de discrimination directe, indirecte et intersectionnelle, ou donner des conseils sur des questions telles que la manière d'appliquer les exceptions et fournir des justifications objectives dans ce contexte.

Conseil de l'Europe

Au niveau du Conseil de l'Europe, plusieurs instruments interdisent la discrimination et sont applicables à la discrimination algorithmique.

12. Sur la façon dont les organismes de promotion de l'égalité peuvent lutter contre la discrimination en ligne induite par l'IA, y compris les discours de haine, voir par exemple : Equinet (2018), « Extending the agenda. Equality Bodies addressing hate speech », disponible en anglais à l'adresse suivante : <https://equineteurope.org/extending-the-agenda-equality-bodies-addressing-hate-speech>, consulté le 7 novembre 2025; Facing Facts Network/CEJI (2022), « Current activities and gaps in hate speech responses : a mapping report for the Facing Facts Network », disponible en anglais à l'adresse www.facingfacts.eu/Hate-Speech-report/, consulté le 7 novembre 2025.

Avant tout, l'article 14 de la Convention européenne des droits de l'homme, qui établit l'interdiction de la discrimination, dispose que « [l]a jouissance des droits et libertés reconnus dans la [...] Convention doit être assurée, sans distinction aucune, fondée notamment sur le sexe, la race, la couleur, la langue, la religion, les opinions politiques ou toutes autres opinions, l'origine nationale ou sociale, l'appartenance à une minorité nationale, la fortune, la naissance ou toute autre situation ». Cette liste non exhaustive de formes de discrimination interdites peut être utilisée pour faire valoir les effets discriminatoires des systèmes d'IA et de PDA chaque fois qu'un des droits fondamentaux énumérés dans la Convention a été violé.

Cette disposition est complétée par l'article 1, paragraphe 1, du Protocole 12 à la Convention, qui énonce une interdiction indépendante de la discrimination : « La jouissance de tout droit prévu par la loi doit être assurée, sans discrimination aucune, fondée notamment sur le sexe, la race, la couleur, la langue, la religion, les opinions politiques ou toutes autres opinions, l'origine nationale ou sociale, l'appartenance à une minorité nationale, la fortune, la naissance ou toute autre situation ». En outre, le paragraphe 2 de l'article 1 dispose que « [n]ul ne peut faire l'objet de discrimination de la part d'une autorité publique quelle qu'elle soit fondée sur les motifs mentionnés au paragraphe 1 ». Cette interdiction de discrimination s'applique – indépendamment du fait qu'un autre droit fondamental soit violé ou non – aux administrations publiques qui déploient des systèmes discriminatoires d'IA et de PDA dans les pays qui ont ratifié le Protocole 12. Ce cadre anti-discrimination est complété, en ce qui concerne le travail, par l'article 20 de la Charte sociale européenne sur le droit à l'égalité des chances et à l'égalité de traitement en matière d'emploi et de travail sans discrimination fondée sur le sexe.

Parmi les autres instruments pertinents, figurent la Convention sur la prévention et la lutte contre la violence à l'égard des femmes et la violence domestique (Convention d'Istanbul), la Convention-cadre pour la protection des minorités nationales ainsi que la Charte européenne des langues régionales ou minoritaires. La Convention d'Istanbul, dans son article 1, paragraphe 1, point b), vise à « contribuer à éliminer toutes les formes de discrimination à l'égard des femmes et [...] promouvoir l'égalité réelle entre les femmes et les hommes, y compris par l'autonomisation des femmes » et interprète la violence à l'égard des femmes comme « une violation des droits de l'homme et une forme de discrimination à l'égard des femmes ». De son côté, l'article 4 de la Convention-cadre pour la protection des minorités nationales stipule que « toute discrimination fondée sur l'appartenance à une minorité nationale est interdite », y compris « les actes de discrimination, d'hostilité ou de violence en raison de [l']identité ethnique, culturelle, linguistique ou religieuse » (article 6). Parallèlement, l'article 7, paragraphe 2, de la Charte européenne des langues régionales ou minoritaires protège les personnes contre la discrimination fondée sur « l'utilisation d'une langue régionale ou minoritaire ».

Au-delà de ces instruments fondamentaux, d'autres documents importants peuvent être utilisés pour lutter contre la discrimination algorithmique. Par exemple, la Stratégie du Conseil de l'Europe pour l'égalité entre les hommes et les femmes 2024-2029 reconnaît « l'incidence spécifique de l'IA sur l'égalité des genres et les droits des femmes » et exhorte les États membres à s'attaquer « à la discrimination algorithmique fondée sur le genre et à la discrimination intersectionnelle par le biais de stratégies

multidimensionnelles d'égalité de genre et de non-discrimination fondées sur les droits humains [et à mettre] en œuvre des normes récemment élaborées dans le domaine de l'intelligence artificielle et de l'égalité de genre » (Conseil de l'Europe 2024c). Cela donne mandat aux États membres de revoir leur législation afin de remédier aux lacunes en matière de discrimination algorithmique.

La Recommandation CM/Rec(2019)¹ du Comité des Ministres aux États membres sur la prévention et la lutte contre le sexisme reconnaît que « Internet a donné une nouvelle dimension à l'expression et à la diffusion du sexisme, en particulier du discours de haine sexiste, auprès d'un large public, même si les origines du sexisme ne sont pas à chercher du côté des technologies mais dans la persistance des inégalités entre les femmes et les hommes » (Conseil de l'Europe 2019b). Il y est précisé que « l'intelligence artificielle pose des défis spécifiques en matière d'égalité entre les femmes et les hommes, et de stéréotypes de genre » et que « l'utilisation d'algorithmes risque de diffuser et de renforcer les stéréotypes existants, contribuant ainsi à perpétuer le sexisme ». Le rôle des systèmes d'IA et de PDA dans « la surveillance à laquelle le corps, l'expression et l'engagement des femmes sont soumis », en particulier en ligne, y est également établi. La Recommandation appelle les États membres à « intégrer une perspective d'égalité entre les femmes et les hommes dans toutes les politiques, programmes et recherches en matière d'intelligence artificielle afin d'éviter les risques potentiels de perpétuation du sexisme et des stéréotypes de genre, et d'examiner comment l'intelligence artificielle pourrait aider à combler les écarts entre les femmes et les hommes et à éliminer le sexisme ». Elle recommande de « combler le fossé dans la participation des femmes et des filles dans le domaine des technologies de l'information » et exige que « la conception d'instruments et d'algorithmes fondés sur les données [prenne en] compte les dynamiques liées au genre ». En outre, « [l]a transparence autour de ces questions devrait être améliorée, de même que la sensibilisation aux potentiels biais sexistes dans les mégadonnées; des solutions pour améliorer la responsabilisation devraient être proposées ». Dans l'ensemble, cette Recommandation représente un instrument juridique important pour lutter contre la discrimination algorithmique.

La Recommandation CM/Rec(2022)¹⁷ du Comité des Ministres aux États membres sur la protection des droits des femmes et des filles migrantes, réfugiées et demandeuses d'asile comporte une partie sur l'IA, la PDA et la protection des données. La Recommandation prévoit que « [t]oute conception, développement et application de l'intelligence artificielle et de systèmes automatisés de prise de décision par le secteur public ou privé ou par des prestataires de services et entreprises sous contrat devraient être non discriminatoires, conformes aux principes de protection de la vie privée, transparents et s'accompagner de mécanismes de gouvernance clairs » dans le contexte de la gestion des frontières et des migrations (Conseil de l'Europe 2022c). Elle encourage les États à veiller à ce que des évaluations d'impact sur les droits humains soient menées avant l'introduction de systèmes d'IA et de PDA dans le domaine de la migration. La Recommandation exige également que des OSC soient impliquées dans les discussions sur le développement et le déploiement de nouvelles technologies impactant des femmes et des filles migrantes, réfugiées et demandeuses d'asile (ibid.).

Les lignes directrices du Comité des Ministres du Conseil de l'Europe sur la défense de l'égalité et la protection contre la discrimination et la haine pendant la pandémie de COVID-19 et les crises similaires à venir contiennent plusieurs dispositions traitant du potentiel discriminatoire de la numérisation, de l'IA et des technologies de traçage des contacts à l'encontre des « groupes vulnérables », bien que l'accent soit mis sur l'exclusion numérique et le manque d'accès (Conseil de l'Europe 2021c). La Recommandation CM/Rec(2022)16 sur la lutte contre les discours de haine et la Recommandation CM/Rec(2024)4 sur la lutte contre les crimes de haine appellent les États membres à « protéger les droits de l'homme et les libertés fondamentales dans l'environnement numérique », soulignent le rôle et les responsabilités des intermédiaires d'Internet dans la diffusion des discours de haine, rappellent l'importance des preuves et des données numériques dans les enquêtes sur les crimes de haine et reconnaissent le rôle des « mouvements extrémistes opérant principalement par le biais de canaux numériques et d'une communication en ligne » dans les crimes de haine (Conseil de l'Europe 2022d). Pourtant, ces recommandations n'abordent pas entièrement le rôle des systèmes de recommandation algorithmique sur les réseaux sociaux dans la promotion d'une polarisation extrême en ligne et donc dans l'alimentation des stéréotypes discriminatoires et des préjugés parmi les utilisateurs et utilisatrices.

La Convention du Conseil de l'Europe sur la protection des enfants contre l'exploitation et les abus sexuels (Convention de Lanzarote) reconnaît « l'utilisation accrue des technologies de communication et d'information par les enfants et les auteurs d'infractions » et le rôle des technologies numériques dans la facilitation de la maltraitance des enfants (Conseil de l'Europe 2007). La Recommandation CM/Rec(2010)5 du Comité des Ministres sur les mesures de lutte contre la discrimination fondée sur l'orientation sexuelle ou l'identité de genre, adoptée en 2010, peut être utilisée transversalement pour lutter contre la discrimination algorithmique, mais ne contient aucune disposition traitant spécifiquement de l'impact discriminatoire des systèmes d'IA et de PDA, bien que de plus en plus de preuves montrent que ces technologies alimentent les préjugés et l'exclusion en ligne et hors ligne (Conseil de l'Europe 2010). Il en va de même pour les instruments existants interdisant la discrimination à l'égard des Roms et des Gens du voyage, tels que la Recommandation CM/Rec(2024)1 du Comité des Ministres aux États membres sur l'égalité pour les femmes et les filles roms et de la communauté des gens du voyage (Conseil de l'Europe 2024d).

Les recommandations de politique générale de la Commission européenne contre le racisme et l'intolérance fournissent également des ressources juridiques utiles pour lutter contre la discrimination algorithmique, en particulier par le prisme des utilisations spécifiques au terrain des systèmes d'IA et de PDA :

- ▶ La Recommandation n° 6 sur la lutte contre la diffusion de matériels racistes, xénophobes et antisémites par l'internet peut contribuer à lutter contre les contenus racistes, les discours de haine, le cyberharcèlement et l'extrémisme alimentés par des systèmes de recommandation algorithmique (Commission européenne contre le racisme et l'intolérance 2000);
- ▶ La Recommandation n° 7 sur la législation nationale pour lutter contre le racisme et la discrimination raciale offre des moyens procéduraux intéressants pour lutter contre la discrimination algorithmique au niveau structurel :

La loi doit disposer que les organisations telles que les associations, les syndicats et autres entités juridiques qui, selon leur législation nationale, ont un intérêt légitime à lutter contre le racisme et la discrimination raciale peuvent tenter des actions civiles, intervenir sur le plan administratif ou déposer des plaintes pénales, même si elles n'invoquent pas le sort d'une victime déterminée (Commission européenne contre le racisme et l'intolérance 2017 : 8)

...

[Ceci est fondamental] pour régler les cas de discrimination où il est difficile d'identifier une telle victime ou ceux qui frappent un nombre indéterminé de victimes (ibid. : 22);

- La Recommandation n° 8 sur la lutte contre le racisme tout en combattant le terrorisme invite les États membres à :

faire en sorte qu'aucune discrimination ne résulte de la législation et des réglementations – ou de leur application – régissant notamment les domaines suivants :

- contrôles auxquels procèdent les forces de l'ordre sur le territoire des États et le personnel de contrôle aux frontières
- détention administrative et provisoire
- ...
- procès équitable, procédure pénale
- protection des données personnelles
- protection de la vie privée et familiale
- expulsion, extradition, reconduite à la frontière et principe de non-refoulement
- délivrance de visas
- permis de séjour et de travail et regroupement familial
- acquisition et retrait de la nationalité

(Commission européenne contre le racisme et l'intolérance 2004).

Étant donné que les systèmes d'IA et de PDA sont de plus en plus utilisés dans ces domaines, en particulier pour faciliter les activités répressives, la recommandation fournit un cadre utile pour lutter contre leur potentiel discriminatoire, même si elle ne traite pas directement de ces technologies;

- La Recommandation n° 10 pour lutter contre le racisme et la discrimination raciale dans et à travers l'éducation scolaire s'applique aux cas où des systèmes d'IA et de PDA sont également déployés, par exemple dans des processus de sélection (Commission européenne contre le racisme et l'intolérance, 2006);
- La Recommandation n° 11 sur la lutte contre le racisme et la discrimination raciale dans les activités de la police, secteur dans lequel les systèmes d'IA et de PDA sont déployés, par exemple, sous forme de technologies de reconnaissance faciale et de comparaison faciale, mais aussi de systèmes de prévision et de surveillance de la criminalité, est pertinente (Commission européenne contre le racisme et l'intolérance, 2007);
- La Recommandation n° 14 sur la lutte contre le racisme et la discrimination raciale dans l'emploi fournit des lignes directrices utiles dans un domaine où les outils de lecture des CV et d'aide à l'embauche sont de plus en plus utilisés (Commission européenne contre le racisme et l'intolérance, 2012);

- La Recommandation n° 15 sur la lutte contre le discours de haine appelle les États membres à « veiller à ce que le cadre de ces infractions [liées au discours de haine] soit défini de manière à pouvoir s'adapter aux évolutions technologiques » (Commission européenne contre le racisme et l'intolérance, 2015).

Ce cadre est complété par d'autres recommandations axées sur des formes spécifiques de discrimination, telles que le racisme, la xénophobie, l'antitsiganisme, l'antisémitisme, la haine antimusulmane, et la discrimination contre les migrant.es en situation irrégulière et les personnes lesbiennes, gays, bisexuelles, transgenres et intersexuées (LGBTI).¹³

Union européenne

Au niveau de l'Union européenne, les principes de non-discrimination et d'égalité entre les femmes et les hommes sont garantis par l'article 2 du traité sur l'UE relatif aux valeurs de l'Union européenne et par l'article 3 du traité sur l'UE portant sur le marché intérieur, ainsi que par les articles 8 et 10 du traité sur le fonctionnement de l'Union européenne (TFUE) sur l'intégration des questions d'égalité entre les femmes et les hommes et de non-discrimination, l'article 19 du TFUE appelant l'UE à adopter une législation anti-discrimination et l'article 157 du TFUE sur l'égalité salariale. L'article 21, paragraphe 1), de la Charte des droits fondamentaux de l'Union européenne interdit « toute discrimination fondée notamment sur le sexe, la race, la couleur, les origines ethniques ou sociales, les caractéristiques génétiques, la langue, la religion ou les convictions, les opinions politiques ou toute autre opinion, l'appartenance à une minorité nationale, la fortune, la naissance, un handicap, l'âge ou l'orientation sexuelle ». L'article 23 de la Charte dispose que « l'égalité entre les hommes et les femmes doit être assurée dans tous les domaines, y compris en matière d'emploi, de travail et de rémunération ».

Plusieurs directives établissent un réseau dense de règles contre la discrimination, notamment :

- la directive 2000/43/CE mettant en œuvre le principe de l'égalité de traitement entre les personnes sans distinction de race ou d'origine ethnique;
- la directive 2000/78/CE portant création d'un cadre général en faveur de l'égalité de traitement en matière d'emploi et de travail;
- la directive 2004/113/CE mettant en œuvre le principe de l'égalité de traitement entre les femmes et les hommes dans l'accès à des biens et services et la fourniture de biens et services;
- la directive 2006/54/CE relative à la mise en œuvre du principe de l'égalité des chances et de l'égalité de traitement entre hommes et femmes en matière d'emploi et de travail (refonte);
- la directive 2010/41/CE relative à l'application du principe de l'égalité de traitement entre hommes et femmes exerçant une activité indépendante;

13. Pour accéder à toutes les recommandations, rendez-vous à l'adresse : www.coe.int/fr/web/european-commission-against-racism-and-intolerance/ecri-standards, consulté le 9 novembre 2025.

- la directive 79/7/CEE relative à la mise en œuvre progressive du principe de l'égalité de traitement entre hommes et femmes en matière de sécurité sociale.

Ces textes s'appliquent à la discrimination algorithmique à condition que les systèmes d'IA et de PDA discriminatoires soient déployés dans leur champ d'application matériel. Ce champ d'application présente toutefois des lacunes, par exemple en ce qui concerne la discrimination fondée sur l'âge, l'orientation sexuelle, le handicap et la religion ou les convictions dans le cadre de l'achat de biens et de services.

Directives de l'Union européenne sur les organismes de promotion de l'égalité

Deux directives ont récemment réformé le mandat des organismes de promotion de l'égalité, acteurs clés de la lutte contre la discrimination algorithmique :

- la directive 2024/1499 sur les normes applicables aux organismes pour l'égalité dans les domaines de l'égalité de traitement entre les personnes sans distinction de race ou d'origine ethnique, de l'égalité de traitement entre les personnes en matière d'emploi et de travail sans distinction de religion ou de convictions, de handicap, d'âge ou d'orientation sexuelle et de l'égalité de traitement entre les femmes et les hommes en matière de sécurité sociale ainsi que dans l'accès à des biens et services et la fourniture de biens et services;
- la directive 2024/1500 sur les normes applicables aux organismes pour l'égalité de traitement dans le domaine de l'égalité de traitement et de l'égalité des chances entre les femmes et les hommes en matière d'emploi et de travail.

D'ici juin 2026, les États membres de l'UE doivent adapter leur législation nationale aux dispositions des deux directives de l'UE sur les organismes de promotion de l'égalité. En ce qui concerne les systèmes automatisés et l'IA, ces directives prévoient que « les organismes pour l'égalité de traitement devraient disposer de ressources humaines et techniques appropriées. Ces ressources devraient notamment permettre aux organismes pour l'égalité de traitement d'utiliser des systèmes automatisés dans le cadre de leurs travaux, d'une part, et d'évaluer ces systèmes du point de vue de leur conformité avec les règles de non-discrimination, d'autre part ».

1.4.3. Instruments de protection des données

Les lois sur la protection des données sont une autre pièce du puzzle législatif que les organismes de promotion de l'égalité et les autres organisations qui luttent contre les inégalités en Europe doivent activer pour lutter contre la discrimination algorithmique.

Conseil de l'Europe

Au niveau du Conseil de l'Europe, le cadre législatif sur la protection des données peut être utilisé pour lutter contre la discrimination algorithmique. La Convention modernisée pour la protection des personnes à l'égard du traitement des données à caractère personnel (Convention 108+), et en particulier son article 6 limitant l'utilisation de catégories particulières de données, invite les États membres à mettre en place des « garanties [...] de nature à prévenir les risques que le traitement de

données sensibles peut présenter pour les intérêts, droits et libertés fondamentales de la personne concernée, notamment un risque de discrimination » (Conseil de l'Europe 2018). L'article 11 de la Convention-cadre sur l'IA sur le respect de la vie privée et la protection des données à caractère personnel étend ces restrictions aux « activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle », demandant en particulier aux États membres de garantir que « les droits à la vie privée des personnes et leurs données à caractère personnel [soient] protégés, notamment par les lois, normes et cadres nationaux et internationaux applicables » et que « des garanties efficaces [aient] été mises en place pour les personnes, conformément aux obligations juridiques nationales et internationales applicables ». Dès son entrée en vigueur, la Convention 108+ s'appliquera également aux domaines de la sécurité et de la défense nationales, ce qui signifie que son champ d'application est plus large que celui du RGPD.

Ce cadre est complété par les lignes directrices sur l'intelligence artificielle et la protection des données publiées en 2019 par le Comité consultatif de la Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel (Convention 108) (Conseil de l'Europe 2019c). Il y est souligné que « [l]es développeurs, fabricants et prestataires de service en IA devraient, à tous les stades du traitement des données, y compris lors de la collecte, adopter une approche des droits de l'homme dès la conception (by-design) et éviter tout biais potentiel, y compris les biais non intentionnels ou cachés, ainsi que les risques de discrimination ou d'autres effets négatifs sur les droits de l'homme et libertés fondamentales des personnes concernées ». On y encourage également « [c]oopération [...] entre autorités de contrôle de la protection des données et d'autres instances ayant des compétences liées à l'IA, telles que les autorités de régulation en matière de protection des consommateurs, concurrence, anti-discrimination, médias et autorités sectorielles ».

En outre, les lignes directrices sur la reconnaissance faciale adoptées en 2021 prévoient que « [l]'utilisation de la reconnaissance faciale dans le seul but de déterminer la couleur de la peau, les convictions religieuses ou autres convictions, le sexe, l'origine raciale ou ethnique, l'âge, l'état de santé ou la condition sociale d'une personne devrait être interdite à moins que des garanties appropriées soient prévues par la loi afin de prévenir tout risque de discrimination » (Conseil de l'Europe 2021d). Elles proposent également des orientations pour les développeurs, les fabricants et les prestataires de services concernant la représentativité des ensembles de données. Fondant leurs recommandations sur l'article 5 de la Convention 108+ relatif à l'exactitude des données, les lignes directrices stipulent que les développeurs, les fabricants et les utilisateurs doivent « éviter les erreurs d'étiquetage en testant suffisamment leurs systèmes et en identifiant et éliminant les disparités dans la précision, notamment en ce qui concerne les variations démographiques de la couleur de la peau, l'âge et le sexe, et éviter ainsi toute discrimination involontaire ». Les lignes directrices demandent également aux autorités répressives qui « envisagent le déploiement de technologies de reconnaissance faciale dans des environnements non contrôlés [...] [d'] aborder le risque pour différents droits fondamentaux, notamment pour le droit à la protection des données, à la liberté d'expression, à la liberté de réunion et à la liberté de circulation ou à la lutte contre la discrimination, en fonction des

utilisations potentielles en différents endroits », notamment par le biais d'analyses d'impact sur la protection des données.

Union européenne

Au niveau de l'Union européenne, plusieurs dispositions du RGPD peuvent jouer un rôle important en ce qui concerne la lutte contre la discrimination dans les systèmes d'IA et de PDA (Union européenne, 2016a). Tout d'abord, l'article 9 sur les catégories sensibles de données à caractère personnel empêche, en principe, les fournisseurs et les déployeurs d'utiliser ces données comme variables (étiquettes, variables d'entrée, facteurs de risque, par exemple) dans les systèmes d'IA et de PDA. Cela dit, comme énoncé plus haut, l'article 10, paragraphe 5, du règlement européen sur l'IA autorise ces utilisations, en interaction avec l'article 9, paragraphe 1, point g), du RGPD, lorsque l'objectif est de détecter et d'atténuer la discrimination algorithmique. En effet, le considérant 70 du règlement européen sur l'IA précise :

Afin de protéger le droit d'autrui contre la discrimination qui pourrait résulter des biais dans les systèmes d'IA, les fournisseurs devraient, à titre exceptionnel, et dans la mesure où cela est strictement nécessaire aux fins de la détection et de la correction des biais en ce qui concerne les systèmes d'IA à haut risque, sous réserve de garanties appropriées pour les libertés et droits fondamentaux des personnes physiques [...] être en mesure de traiter également des catégories particulières de données à caractère personnel, pour des raisons d'intérêt public substantiel au sens de l'article 9, paragraphe 2, point g), du [RGPD].

Par ailleurs, l'article 22 du RGPD confère le droit de ne pas faire l'objet de décisions entièrement automatisées. En particulier, « [l]a personne concernée a le droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire ». Cette disposition a été utilisée pour contester la notation de crédit discriminatoire dans les affaires SCHUFA Holding et Dun & Bradstreet Austria (voir section 1.4.4.). Bien que le considérant 71 du RGPD accorde aux personnes concernées le droit « d'obtenir une explication quant à la décision à l'issue [d'une] évaluation » par un système reposant exclusivement sur la PDA, pendant longtemps, la question était de savoir si le RGPD offrait un droit à l'explication et la nature et les limites potentielles d'un tel droit étaient caractérisés par une incertitude juridique (Wachter, Mittelstadt et Floridi 2017; Malgieri et Comandé 2017; Selbst et Powles 2017). Le RGPD autorise les personnes concernées à recevoir les informations « nécessaires pour garantir un traitement équitable et transparent » (articles 13, paragraphe 2, et 14, paragraphe 2), y compris en ce qui concerne « l'existence d'une prise de décision automatisée, y compris un profilage, visée à l'article 22, paragraphes 1 et 4, et, au moins en pareils cas, des informations utiles concernant la logique sous-jacente, ainsi que l'importance et les conséquences prévues de ce traitement pour la personne concernée » (article 13, paragraphe 2, point f; article 14, paragraphe 2, point g; et article 15, paragraphe 1, point h). Dans l'arrêt Dun & Bradstreet Austria (voir section 1.4.4.), la Cour de justice de l'Union européenne a confirmé l'existence et l'étendue du droit à explication prévu à l'article 15, paragraphe 1, point h) du RGPD, qui conditionne l'accès à des informations utiles sur la manière dont une décision a été prise et offre une base pour contester les décisions automatisées en vertu de l'article 22, paragraphe 3, du RGPD. L'article 22, paragraphe 1, du RGPD stipule que les personnes concernées ont le droit « de ne

pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques [les] concernant ou [les] affectant de manière significative de façon similaire ». Toutefois, ce droit ne s'applique pas lorsque le traitement automatisé ou le profilage est « nécessaire à la conclusion ou à l'exécution d'un contrat entre la personne concernée et un responsable du traitement », ou lorsque la personne concernée a donné son « consentement explicite » conformément à l'article 22, paragraphe 2, points a) et c) du RGPD. Dans ce cas de figure, l'article 22, paragraphe 3, du RGPD exige que le responsable du traitement mette en œuvre « des mesures appropriées pour la sauvegarde des droits et libertés et des intérêts légitimes de la personne concernée, au moins du droit de la personne concernée d'obtenir une intervention humaine de la part du responsable du traitement, d'exprimer son point de vue et de contester la décision ». Le droit à l'explication récemment créé par la Cour de justice peut faciliter l'accès à la justice, du moins dans les cas de discrimination algorithmique fondée sur le profilage ou de processus de PDA exclusifs.

Dans le domaine du droit pénal, la directive (UE) 2016/680 du Parlement européen et du Conseil du 27 avril 2016 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales, et à la libre circulation de ces données, s'applique (Union européenne 2016b). Elle fournit des garanties plus limitées qui peuvent ne pas suffire à contester la discrimination algorithmique dans ces domaines de déploiement.

1.4.4. Jurisprudence de la Cour européenne des droits de l'homme et de la Cour de justice de l'Union européenne

Cour européenne des droits de l'homme

Dans son arrêt *Glukhin c. Russie* de 2023, la Cour européenne des droits de l'homme a évalué la légalité du recours, par la police, à des technologies de reconnaissance faciale en direct (Cour européenne des droits de l'homme 2023). Dans cette affaire, un individu a été arrêté à la suite de la publication, sur ses réseaux sociaux, de photos et d'une vidéo d'une manifestation pacifique individuelle qu'il avait effectuée dans le métro de Moscou, que la police a découvert lors de ses activités de surveillance de routine d'Internet. Pour l'identifier, la police a utilisé des images de vidéosurveillance provenant de caméras de vidéosurveillance qui filment les espaces publics combinées à des données biométriques, en l'occurrence une technologie de reconnaissance faciale. La Cour a conclu à une violation de la liberté d'expression du requérant en vertu de l'article 10 et de son droit au respect de la vie privée en vertu de l'article 8 de la Convention européenne des droits de l'homme. En particulier, la Cour a déclaré que « [l']utilisation d'une technologie de reconnaissance faciale très intrusive dans une situation où le requérant exerçait son droit à la liberté d'expression tel que garanti par la Convention est incompatible avec les idéaux et valeurs d'une société démocratique régie par la prééminence du droit, que la Convention est destinée à sauvegarder et à promouvoir ». En tenant compte de « la difficulté pour le requérant de prouver ses allégations dès lors que le droit interne n'imposait pas que le recours

à une technologie de reconnaissance faciale fit l'objet d'un procès-verbal ou d'une notification officiels », la Cour a également reconnu que le manque de transparence sur l'utilisation de ces technologies rendait difficile la production de preuves et l'atténuation de ces obstacles (ibid., point 72).

Cour de justice de l'Union européenne

Une demande de décision préjudicielle présentée par la Belgique concernant sa large mise en œuvre de la directive de l'UE sur les dossiers passagers, y compris sur la manière dont l'automatisation du système de traitement des données devrait être mise en œuvre par les autorités nationales, a donné lieu à une décision dans l'affaire C-817/19 Ligue des droits humains ASBL c. Conseil des ministres (Cour de justice de l'Union européenne 2022). L'affaire portait sur la tension entre le traitement automatisé des données par le biais d'algorithmes d'apprentissage automatique et la non-discrimination. La Cour de justice a rappelé les garanties en matière de traitement automatisé des données, excluant l'utilisation d'algorithmes d'IA auto-apprentissage pour prévenir les résultats discriminatoires. De plus, les critères prédéterminés utilisés pour évaluer les menaces terroristes ne peuvent être fondés sur « l'origine raciale ou ethnique d'une personne, ses opinions politiques, sa religion ou ses convictions philosophiques, son appartenance à un syndicat, son état de santé, sa vie sexuelle ou son orientation sexuelle » (ibid.).

Dans son arrêt C-634/21 SCHUFA Holding, la Cour de justice a classé le profilage sous forme de notation des risques liés à des crédits comme une forme de décision individuelle automatisée relevant de l'article 22 du RGPD (Cour de justice de l'Union européenne 2023). Dans le cadre du litige ayant donné lieu à cette décision, le requérant s'était vu refuser un prêt par une institution financière sur la base d'une note de risque fournie par SCHUFA Holding, une société allemande de notation de crédit. La question était de savoir si l'établissement d'une [telle] valeur de probabilité [constituait] une décision individuelle automatisée au sens de l'article 22, paragraphe 1, du RGPD », en d'autres termes « une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques concernant cette personne ou l'affectant de manière significative de façon similaire » (ibid. : paragraphes 20, 27). Bien que la note de risque fournie par SCHUFA soit transmise à un tiers, qui est responsable en dernier ressort de la décision contractuelle, la Cour a soutenu que la décision de ce tiers était « guidée de manière déterminante par cette valeur » (ibid. : point 48). Il a également été suggéré que la notion de « décision » visée à l'article 22 du RGPD était « suffisamment large pour englober le résultat du calcul de la solvabilité d'une personne sous la forme d'une valeur de probabilité concernant la capacité de cette personne à honorer ses engagements de paiement à l'avenir » (ibid. : paragraphe 46). Ainsi, lorsque « la valeur de probabilité établie par une société fournissant des informations commerciales et communiquée à une banque joue un rôle déterminant dans l'octroi d'un crédit, l'établissement de cette valeur doit être qualifié en soi de décision produisant à l'égard d'une personne concernée 'des effets juridiques la concernant ou l'affectant de manière significative de façon similaire', au sens de l'article 22, paragraphe 1, du RGPD » (ibid. : paragraphe 50).

Il s'agit d'une décision historique qui constitue une base juridique pour contester la discrimination fondée sur le profilage et d'autres décisions fondées majoritairement sur des scores de risque automatisés.

Dans son arrêt *Dun & Bradstreet Austria* (Cour de justice de l'Union européenne 2025), la Cour de justice a enrichi les voies de recours ouvertes aux victimes de discrimination algorithmique. En l'espèce, le requérant s'était vu refuser un contrat de téléphonie mobile par une société autrichienne de télécommunications sur la base d'une évaluation automatisée de son crédit. Le score de risque algorithmique fourni par Dun & Bradstreet, entreprise spécialisée dans la fourniture d'évaluations de crédit, indiquait qu'elle ne présentait pas une solvabilité financière suffisante pour obtenir ce contrat, qui aurait nécessité un paiement mensuel de 10 euros (ibid. : points 2, 16). Le requérant a saisi l'autorité autrichienne de protection des données, qui a ordonné à la société de notation de crédit de lui communiquer « des informations utiles concernant la logique sous-jacente à la prise de décision automatisée fondée sur des données à caractère personnel » (ibid. : point 17). Invoquant des secrets commerciaux, la société d'évaluation du crédit a d'abord refusé de fournir au requérant d'autres informations que son indice de risque (ibid.). Elle avait seulement indiqué que « certaines données sociodémographiques [du requérant] avaient été « agrégées de manière équivalente » » pour établir ce score de risque, ce que la juridiction autrichienne a jugé insuffisant. Pour tenter de faire exécuter cette décision, un tribunal national a nommé un expert, lequel a estimé que le respect de l'obligation de fournir des « informations utiles » nécessiterait que Dun & Bradstreet Austria communique au ou à la demandeur-se quelles données personnelles (par exemple : date de naissance, adresse, sexe) ont été traitées pour générer les « facteurs » et « la valeur spécifique » attribuée au requérant pour chaque facteur, « la formule mathématique » sur laquelle le score reposait, « la précision des intervalles à l'intérieur desquels la même valeur est attribuée à différentes données pour le même facteur » ainsi qu'une « liste établissant les « scores » de personnes (« scoring ») pour la période couvrant les six mois qui ont précédé et les six mois qui ont suivi l'établissement du « score » [du requérant], et qui ont été obtenus sur le fondement de la même règle de calcul » afin de vérifier leur exactitude (ibid. : points 23 à 25). Selon le tribunal national, la possibilité pour les demandeur.ses de vérifier l'exactitude du profilage est cruciale aux fins de l'application de la loi.¹⁴ Dans le litige en cause, par exemple, « les informations fournies [au requérant], dont notamment le « score » obtenu, attestaient d'une très bonne solvabilité de [celui-ci], [alors que] le profilage réel aurait conduit à considérer qu'elle n'était pas solvable » (ibid. : point 27). L'accès à des informations utiles conditionne donc la capacité des requérant.es

14. Le tribunal national a fait valoir que « [d]ans l'hypothèse où l'article 15, paragraphe 1, sous h), du RGPD ne garantirait pas cette vérification, le droit d'accès aux données à caractère personnel de la personne concernée et à d'autres informations qui y est prévu serait vidé de toute portée et rendu inutile, d'autant que chaque responsable pourrait dans ce cas être en mesure de fournir des informations erronées » (ibid., point 29).

à contester le caractère discriminatoire du profilage ou des décisions automatisées dans le contexte de l'article 22(3) du RGPD (voir section 1.4.3.) (ibid.: point 55).¹⁵

La Cour de justice a expliqué que « le droit d'obtenir des « informations utiles concernant la logique sous-jacente » à une prise de décision automatisée, au sens de [l'article 15, paragraphe 1, point h], doit être compris comme un droit à l'explication de la procédure et des principes concrètement appliqués pour exploiter, par la voie automatisée, les données à caractère personnel de la personne concernée aux fins d'en obtenir un résultat déterminé, tel un profil de solvabilité », qui doivent être fournies « au moyen d'informations pertinentes et d'une façon concise, transparente, compréhensible et aisément accessible » (ibid. : points 58 à 59). Bien que la Cour n'ait pas énuméré les éléments pertinents, elle a indiqué que « ni la simple communication d'une formule mathématique complexe, telle qu'un algorithme, ni la description détaillée de toutes les étapes d'une prise de décision automatisée » ne satisferait les exigences de l'article 15, paragraphe 1, point h car « aucune de ces modalités ne constituerait une explication suffisamment concise et compréhensible. » (ibid. : point 59). En d'autres termes :

Les « informations utiles concernant la logique sous-jacente » à la prise de décision automatisée, au sens de l'article 15, paragraphe 1, point h) du RGPD, doivent décrire la procédure et les principes effectivement appliqués de manière à ce que la personne concernée puisse comprendre quelles données à caractère personnel la concernant ont été utilisées et de quelle manière dans la prise de décision automatisée en cause, étant donné que la complexité des opérations à effectuer dans le cadre de la prise de décision automatisée n'est pas de nature à dégager le responsable du traitement de l'obligation de fournir une explication (ibid. : point 61).

Pour la Cour, il peut être « suffisamment transparent et intelligible [...] d'informer la personne concernée de la mesure dans laquelle une variation au niveau des données à caractère personnel prises en compte aurait conduit à un résultat différent » (ibid.). Dans le contexte de la législation anti-discrimination, ce droit à l'explication est essentiel car il pourrait permettre aux requérant.es de se comparer avec des comparateurs réels ou hypothétiques, élément essentiel du test de discrimination, qui est cependant devenu très inaccessible dans le contexte de l'opacité algorithmique et de la fragmentation épistémique (voir section 1.4.3.). Enfin, lorsque des secrets commerciaux pourraient être compromis, le responsable du traitement « est tenu de communiquer ces informations prétendument protégées à l'autorité de contrôle ou à la juridiction compétentes, auxquelles il incombe de pondérer les droits et les intérêts en cause aux fins de déterminer l'étendue du droit d'accès de la personne concernée prévu à l'article 15 du RGPD » (ibid. : point 76).

En France, 15 organisations ont récemment contesté une décision de la Caisse nationale des allocations familiales d'utiliser un algorithme de notation des risques pour décider sur quels bénéficiaires enquêter en matière de fraude auprès du Conseil d'État. Il a été démontré que le système de PDA avait des effets discriminatoires,

15. La Cour a déclaré que « dans le contexte spécifique de l'adoption d'une décision fondée exclusivement sur un traitement automatisé, la finalité principale du droit de la personne concernée à obtenir les informations prévues à l'article 15, paragraphe 1, sous h), du RGPD consiste à lui permettre d'exercer de manière efficace les droits qui lui sont reconnus par l'article 22, paragraphe 3, de ce règlement, à savoir celui d'exprimer son point de vue sur cette décision et celui de contester celle-ci » (ibid. paragraphe 55).

notamment pour les femmes et les personnes vivant dans la pauvreté. Les requérants et requérantes ont demandé au Conseil d'État de soumettre les questions à la Cour de justice, notamment s'agissant de la discrimination algorithmique. Cette affaire devrait faire l'objet d'un suivi car elle donnera à la Cour de justice la possibilité d'étudier la discrimination algorithmique sous l'angle de la discrimination directe et indirecte.

1.5. Principaux acteurs concernés

Bureau européen de l'IA

Plusieurs institutions joueront un rôle dans la lutte contre la discrimination algorithmique. Le Bureau européen de l'IA récemment créé¹⁶ (établi au sein de la Commission européenne) soutiendra la mise en œuvre du règlement européen sur l'IA et enquêtera sur les infractions éventuelles. En vertu de l'article 90, le Bureau européen de l'IA sera notamment chargé de traiter les alertes déclenchées par le groupe scientifique d'experts indépendants (article 68 du règlement européen sur l'IA) sur les modèles d'IA à usage général présentant des risques systémiques. Semblable à une autorité de surveillance du marché, le Bureau de l'IA aura également le pouvoir de surveiller et de superviser la conformité des systèmes d'IA qui intègrent un modèle d'IA à usage général (article 75(1) du règlement européen sur l'IA). Il coopérera également avec les autorités de surveillance du marché pour surveiller, enquêter et évaluer la conformité des modèles d'IA à usage général utilisés dans les domaines à haut risque (article 75, paragraphe 2, du règlement européen sur l'IA). Conformément à l'article 75, paragraphe 3, du règlement européen sur l'IA, si « une autorité de surveillance du marché n'est pas en mesure de conclure son enquête sur le système d'IA à haut risque en raison de son incapacité à accéder à certaines informations relatives au modèle d'IA à usage général malgré tous les efforts appropriés pour obtenir ces informations, elle peut présenter une demande motivée au Bureau de l'IA, par laquelle l'accès à ces informations est mis en œuvre ». À la lumière de ces dispositions, il est probable que les organismes de promotion de l'égalité désignés en vertu de l'article 77 du règlement européen sur l'IA coopéreront directement ou indirectement avec le Bureau de l'IA dans le cadre de leur coopération avec les autorités de surveillance du marché lorsqu'ils aborderont les risques ou les atteintes aux droits anti-discrimination. En effet, comme expliqué à la section 1.4.1., l'article 73, paragraphe 7, du règlement européen sur l'IA impose aux autorités de surveillance du marché concernées d'informer les autorités ou organismes publics nationaux visés à l'article 77, paragraphe 1, du règlement européen sur l'IA en cas d'incident grave compris comme une « violation des obligations en vertu du droit de l'Union visant à protéger les droits fondamentaux » au sens de l'article 3, paragraphe 49, point c), du règlement européen sur l'IA. L'article 79, paragraphe 2, du règlement européen sur l'IA exige également que « [l]orsque sont identifiés des risques pour les droits fondamentaux, l'autorité de surveillance du marché informe également les autorités ou organismes publics nationaux concernés visés à l'article 77, paragraphe 1, et coopère pleinement avec eux ». La consultation des autorités visées à l'article 77(1) est également prévue par l'article 82(1) du règlement européen sur l'IA.

16. Voir <https://digital-strategy.ec.europa.eu/fr/policies/ai-office>, consulté le 10 novembre 2025.

Comité européen de l'intelligence artificielle

L'article 65 du règlement européen sur l'IA institue également un Comité européen de l'intelligence artificielle¹⁷. Le Comité de l'IA est composé de représentants des États membres et accorde le statut d'observateur au contrôleur européen de la protection des données. Le Comité de l'IA joue un rôle clé dans l'application du règlement européen sur l'IA, notamment en aidant à « coordonner et à assurer la coopération entre les États membres de l'UE, en vue d'une mise en œuvre et d'une application cohérentes du règlement européen sur l'IA dans toute l'Union » (Commission européenne, 2025c). Il fournit une plate-forme de coopération aux autorités nationales compétentes chargées de faire appliquer le règlement européen sur l'IA, et facilite la mise en commun d'expertises techniques et réglementaires ainsi que de bonnes pratiques. Il est également chargé de créer des sous-groupes pour faciliter la coopération sur les questions d'application, en particulier deux sous-groupes permanents pour offrir une plate-forme de coopération et d'échange entre les autorités de surveillance du marché et les autorités notifiantes, mais aussi pour fournir des conseils sur des questions de politique en matière d'IA.

Forum consultatif sur l'IA

L'article 67 du règlement européen sur l'IA institue un forum consultatif sur l'IA¹⁸ chargé d'apporter une expertise technique et de conseiller et de soutenir les travaux du Comité de l'IA et de la Commission européenne. Le forum consultatif a une composition « équilibrée en ce qui concerne la représentation des parties prenantes, y compris l'industrie, les jeunes pousses, les PME [petites et moyennes entreprises], la société civile et le monde universitaire », avec une représentation équilibrée des intérêts commerciaux et non commerciaux ainsi que des PME et des autres entreprises.

Groupe scientifique d'experts indépendants

L'article 68 du règlement européen sur l'IA institue un groupe scientifique d'expertes et experts indépendant.es destiné.es à appuyer les activités de contrôle de l'application du règlement. Ce groupe est composé de personnes sélectionnées par la Commission sur la base d'une expertise scientifique ou technique actualisée dans le domaine de l'IA, nécessaire à l'accomplissement des missions suivantes :

- ▶ Soutenir la mise en œuvre et le contrôle de l'application du règlement européen sur l'IA, en particulier en ce qui concerne les modèles d'IA à usage général (par exemple en alertant le Bureau de l'IA sur les éventuels risques systémiques, en contribuant à la mise au point d'outils et de méthodologies d'évaluation des capacités de l'IA, en fournissant des conseils quant à la classification des modèles d'IA à usage général présentant ou non un risque systémique, et en contribuant à la mise au point d'outils et de modèles);
- ▶ soutenir, à leur demande, les autorités de surveillance du marché dans leur travail;
- ▶ soutenir les activités transfrontalières de surveillance du marché;

17. Voir <https://artificialintelligenceact.eu/fr/article/65/>, consulté le 10 novembre 2025.

18. Voir <https://artificialintelligenceact.eu/fr/article/67/>, consulté le 10 novembre 2025.

- Soutenir le Bureau de l'IA dans l'exercice de ses fonctions dans le cadre de la procédure de sauvegarde de l'Union.

CEN-CENELEC

Le Comité technique mixte CEN-CENELEC 21 (JTC 21) mène des travaux sur des normes pour l'IA pour répondre à une demande de normalisation de la Commission européenne.¹⁹ Actuellement, le JTC 21 travaille à l'élaboration d'une norme harmonisée traduisant l'exigence, établie par l'article 9 du règlement européen sur l'IA, selon laquelle les fournisseurs de systèmes d'IA à haut risque doivent disposer d'un système de gestion des risques.

Rôle des organismes de promotion de l'égalité

Les organismes de promotion de l'égalité ont un rôle clé à jouer. La discrimination algorithmique prend souvent des formes qui la rendent difficile à détecter et à prouver. Cela constitue un obstacle important à l'accès à la justice pour les personnes impactées. Les personnes qui subissent une discrimination algorithmique peuvent avoir des difficultés à comprendre comment ou pourquoi elles ont été traitées injustement, ce qui rend difficile la contestation de ces décisions. Les organismes de promotion de l'égalité jouent donc un rôle crucial dans le soutien aux victimes et la promotion de la responsabilisation. Dans ce contexte, les directives récemment adoptées sur les organismes de promotion de l'égalité (directive 2024/1500 et directive 2024/1499) renforcent le rôle et les pouvoirs des organismes de promotion de l'égalité en ce qui concerne leur indépendance, leurs ressources et leur rôle dans la facilitation de l'accès à la justice. Elles confirment également le rôle des organismes de promotion de l'égalité dans la lutte contre la discrimination induite par l'IA :

Il est essentiel d'accorder une attention particulière aux possibilités et aux risques que présente l'utilisation de systèmes automatisés, y compris l'intelligence artificielle. En particulier, les organismes pour l'égalité de traitement devraient disposer de ressources humaines et techniques appropriées. Ces ressources devraient notamment permettre aux organismes pour l'égalité de traitement d'utiliser des systèmes automatisés dans le cadre de leurs travaux, d'une part, et d'évaluer ces systèmes du point de vue de leur conformité avec les règles de non-discrimination, d'autre part. Lorsque l'organisme pour l'égalité de traitement fait partie d'un organisme à plusieurs mandats, il convient de garantir les ressources nécessaires à l'accomplissement de son mandat ayant trait à l'égalité (Union européenne 2024b : Considérant 21; Union européenne 2024a : considérant 22).

Dans ce contexte, les organismes de promotion de l'égalité peuvent jouer un rôle essentiel en soutenant les victimes, en les sensibilisant, en menant des recherches et en plaidant pour une meilleure réglementation des systèmes d'IA et de PDA :

- agir d'office et traiter les plaintes des victimes de discrimination algorithmique : les organismes de promotion de l'égalité peuvent disposer d'office de pouvoirs d'enquête et servir de point de contact accessible pour les victimes potentielles de discrimination algorithmique. Ils peuvent traiter les plaintes des victimes,

19. Décision d'exécution de la Commission européenne du 22 mai 2023 exposant la demande de normalisation C(2023)3215.

enquêter sur les cas présumés de partialité algorithmique et fournir des conseils et un soutien sur la manière d'intenter une action en justice;

- ▶ apporter une assistance juridique et saisir les tribunaux : cela pourrait inclure la prise de décisions concernant des cas de discrimination algorithmique, conseiller les particuliers sur la marche à suivre pour contester des décisions discriminatoires induites par des systèmes d'IA ou de PDA ou, dans certains cas, représenter des victimes potentielles ou agir en leur nom devant les tribunaux;
- ▶ sensibiliser et informer les victimes et le grand public : les organismes de promotion de l'égalité peuvent jouer un rôle clé dans l'éducation du public sur les dangers potentiels de la discrimination algorithmique, y compris en fournissant des informations sur la procédure de contestation des décisions automatisées injustes et les droits des personnes en ce qui concerne les processus basés sur l'IA ou la PDA;
- ▶ collecter des données et mener des recherches sur les discriminations liées à des systèmes d'IA et de PDA : un autre rôle essentiel des organismes d'égalité est de collecter des données sur survenue de cas de discrimination algorithmique. En effectuant un suivi systématique des plaintes liées aux systèmes d'IA et de PDA et en procédant à des recherches, à une collecte de données et à un suivi, les organismes chargés de l'égalité peuvent identifier les tendances de discrimination et mettre en évidence les problèmes systémiques qui peuvent nécessiter une intervention réglementaire;
- ▶ formuler des recommandations à l'intention des décideurs et des législateurs : ces données peuvent également fournir des preuves concrètes de la nécessité de réformes juridiques (par exemple en ce qui concerne les obligations d'enregistrement et d'information). Les organismes de promotion de l'égalité peuvent proposer aux décideurs et aux législateurs des recommandations fondées sur des données factuelles en ce qui concerne la façon de réglementer les systèmes d'IA et de PDA en fonction de leurs activités de surveillance, en particulier du point de vue des activités répressives et des difficultés connexes;
- ▶ Collaborer avec les parties prenantes pour promouvoir l'égalité dans les systèmes d'IA et de PDA : les organismes chargés de l'égalité peuvent collaborer avec un large éventail de parties prenantes – y compris des employeurs, des fournisseurs de services, des institutions publiques et des OSC – afin d'encourager l'adoption de bonnes pratiques d'égalité en ce qui concerne les systèmes d'IA et de PDA. Cela pourrait consister à aider les organisations à élaborer et à mettre en œuvre des plans d'égalité qui minimisent le risque de biais algorithmique, en fournissant des conseils sur la façon de tester les systèmes d'IA et de PDA pour détecter toute discrimination.



2. Lacunes juridiques

Cette partie identifie les lacunes et les besoins dans les cadres juridiques, procéduraux et de gouvernance actuels à la lumière des risques discriminatoires découlant de l'utilisation de systèmes d'IA et de PDA dans les administrations publiques et d'autres secteurs relevant du champ d'étude de la présente étude. Elle résume également les principaux obstacles, limites et besoins rencontrés par les organismes de promotion de l'égalité et les autres parties prenantes concernées dans leurs fonctions de suivi et de supervision.

2.1. Justification de l'adoption de systèmes d'IA et de PDA

La partie I a montré que l'adoption de technologies d'IA et de PDA est de plus en plus répandue. Pourtant, plusieurs problèmes se posent en ce qui concerne leur adoption. D'une part, les discours promouvant l'adoption de ces systèmes sont problématiques et augmentent les risques juridiques. Les systèmes d'IA et de PDA sont souvent présentés comme « neutres » ou du moins « moins biaisés » que les décideurs humains. Leur déploiement pourrait être justifié par une volonté de rationaliser les processus décisionnels. Puisque ces systèmes sont connus pour être largement biaisés, ce type de discours peut créer l'illusion que le déploiement de l'IA et de la PDA réduit, voire empêche la discrimination. Cela crée le risque que les violations de la législation anti-discrimination soient négligées.

En outre, les systèmes d'IA et de PDA sont souvent promus comme des moyens d'accroître la rentabilité des processus décisionnels, notamment dans les administrations publiques qui sont soumises à une surveillance renforcée dans le contexte des réductions des dépenses et des politiques d'austérité. Cette justification mène à l'adoption de l'IA et de la PDA pour des usages spécifiques, tels que la détection de fraude ou l'évaluation des risques, qui sont présentés comme des moteurs pour l'efficacité des dépenses. Ce discours obscurcit au moins deux questions importantes :

- ▶ ces technologies devraient-elles être adoptées pour d'autres usages, par exemple pour améliorer l'accès aux droits et aux services publics, dont les études montrent régulièrement qu'ils sont sous-estimés ? (Défenseur des droits 2024);
- ▶ ces systèmes peuvent-ils réellement conduire à une baisse des dépenses publiques sans compromettre la qualité des services publics, une fois que tous les coûts liés à la sauvegarde des droits fondamentaux ont été internalisés ?

Par exemple, empêcher les risques de discrimination de se matérialiser grâce à des procédures adéquates de test, de suivi et d'audit, former les agent.es chargé.es de traiter les recommandations algorithmiques afin d'assurer une surveillance humaine efficace, et mettre en place des mécanismes efficaces d'information et de recours pour les utilisateurs des services publics est coûteux. Ces coûts devraient sans doute être pleinement internalisés par les administrations publiques lors de la prise de décisions concernant la mise en œuvre de ce type de systèmes. Ces délibérations devraient également inciter les décideurs publics à poser des « questions zéro » : lorsque le déploiement d'un système d'IA ou de PDA apparaît trop risqué ou coûteux une fois que les mesures de prévention et les garanties des droits fondamentaux ont été correctement mises en œuvre, cela devrait conduire à une décision de ne pas déployer ce système.

2.2. Manque de transparence

Comme détaillé dans les sections correspondantes à venir, il y a un manque généralisé d'informations systématiques, claires et accessibles sur le développement, l'expérimentation et le déploiement des systèmes d'IA et de PDA dans les administrations publiques et ailleurs. L'utilisation de systèmes d'IA et de PDA dans les administrations publiques n'est pas systématiquement cartographiée malgré les projets ou les efforts actuels visant à créer des bases de données recensant ces utilisations (par exemple, le Registre néerlandais des algorithmes)²⁰. Les cas de discrimination algorithmique sont encore très peu nombreux et reposent souvent sur des questions de protection des données plutôt que de discrimination et d'égalité. Les rares affaires existantes qui sont apparues devant les tribunaux l'ont souvent été grâce au soutien d'organismes de promotion de l'égalité et d'organisations de la société civile. Cependant, l'absence d'informations sur le moment, les raisons et les contextes dans lesquels les administrations publiques recourent à des systèmes d'IA et de prise de décision automatisée (ADM) – et, par conséquent, l'absence d'obligations juridiques de reporting applicables à ce jour – rend difficile pour ces acteurs de contrôler, d'évaluer et de contester de manière effective et systématique

20. Voir <https://algoritmes.overheid.nl/en> (en 5en anglais), consulté le 10 novembre 2025.

les cas de discrimination algorithmique, malgré leurs pouvoirs d'enquête. En effet, pour faire valoir efficacement ces pouvoirs d'enquête, ils doivent d'abord et avant tout être conscients de l'existence des utilisations de l'IA ou de la PDA. Au lieu de cela, les organismes de promotion de l'égalité doivent souvent s'appuyer sur des enquêtes médiatiques, des OSC spécialisées ou des réseaux informels pour identifier des cas potentiels de discrimination algorithmique. Les experts nationaux indiquent que des demandes d'information (par exemple au titre de la législation sur la liberté d'information) sont régulièrement adressées aux autorités publiques afin de constituer des archives publiques des usages de l'IA et des systèmes ADM dans les administrations. Toutefois, leur légitimité est parfois remise en question, ou elles peuvent rester sans réponse lorsqu'elles n'ont pas de caractère contraignant, en particulier lorsque ces demandes portent sur des applications sensibles, telles que les technologies de surveillance numérique. En outre, souvent, les demandeur.ses ne peuvent pas vérifier les renseignements reçus. D'autres canaux d'information sont parfois utilisés, comme les questions adressées au gouvernement en collaboration avec les membres des parlements nationaux. Cependant, l'efficacité des demandes et des questions relatives à la liberté d'information adressées au parlement est compromise par le manque même de transparence sur l'utilisation de systèmes d'IA et de PDA. Certains organismes de promotion de l'égalité offrent des exemples positifs à cet égard : le Médiateur finlandais pour la non-discrimination, par exemple, peut adresser des demandes d'information et de clarification contraignantes aux administrations publiques. Ces pouvoirs peuvent faciliter l'accès à l'information et la transparence en ce qui concerne l'utilisation des systèmes d'IA et de PDA et leurs effets discriminatoires potentiels.

Dans une certaine mesure, le manque d'information sera corrigé par les nouvelles réglementations sur l'IA. Par exemple, la base de données de l'UE sur l'IA prévue par l'article 71 du règlement européen sur l'IA contiendra des informations publiques sur les applications d'IA à haut risque.²¹ En outre, l'article 14(a)(b) de la Convention-cadre sur l'IA, qui impose que « des informations pertinentes concernant les systèmes d'intelligence artificielle susceptibles d'avoir une incidence significative sur les droits humains et leur utilisation pertinente sont documentées, fournies aux organismes autorisés à avoir accès à ces informations et, si nécessaire et applicable, mises à la disposition des personnes concernées ou communiquées à ces dernières » permet à ces personnes « de contester la ou les décisions prises par le biais de l'utilisation du système ou fondées en grande partie sur celle-ci ». Ces dispositions sont également renforcées par les articles 8 et 9 de la Convention sur la transparence et le contrôle et sur la responsabilité. Ces exigences d'information et de transparence peuvent permettre un contrôle public. Dans le même ordre d'idées, l'article 16 de la Convention exige l'adoption et le maintien d'un cadre de gestion des risques et des impacts (étayé par l'analyse d'impact normalisée HUDERIA), et l'article 16, paragraphe 2, point f de la Convention exige « la documentation des risques, des impacts réels

21. Toutefois, les systèmes d'IA à haut risque, y compris les systèmes biométriques, qui sont utilisés dans les domaines des activités répressives, de la migration, de l'asile et de la gestion des contrôles aux frontières sont enregistrés dans une partie non publique de la base de données, à laquelle ne peuvent accéder que la Commission européenne et certaines autorités nationales de surveillance du marché (par exemple, les autorités chargées de la protection des données).

et potentiels, et de l'approche de la gestion des risques ». L'accès à ces informations pourrait aider les organismes chargés de la protection des droits fondamentaux à prévenir et à combattre la discrimination algorithmique. Le règlement européen sur l'IA, quant à lui, prévoit que les autorités désignées en vertu de son article 77 aient accès à la documentation technique et aux systèmes d'évaluation des risques qui sont basés sur le règlement européen sur l'IA et les analyses d'impact sur les droits fondamentaux, lorsque l'article 27 du règlement européen sur l'IA l'exige (limité à certains déployeurs qui proposent des services publics ou privés essentiels).

Cela dit, ces obligations en matière d'information et de transparence auraient pu être formulées en termes plus larges. À l'heure actuelle, cela ne concerne que les systèmes d'IA à haut risque, ce qui épargne les systèmes décrits comme n'étant pas à haut risque et les systèmes de PDA qui ne sont pas considérés comme relevant de la définition de l'IA. En 2020, l'Assemblée parlementaire du Conseil de l'Europe a exigé que les gouvernements soient priés « d'informer à l'avance le parlement de l'utilisation [de technologies d'IA et de PDA] » et de « rendre obligatoire l'inscription systématique dans un registre public de chaque utilisation de ces technologies par les autorités » (Conseil de l'Europe, Assemblée parlementaire 2020b). Ces principes ne sont pas directement reflétés dans la Convention-cadre sur l'IA récemment adoptée, qui n'exige pas la création d'un registre public. De son côté, le règlement européen sur l'IA ne prévoit pas de déclaration aux parlements nationaux ou aux institutions de l'UE lors du déploiement de systèmes d'IA ou de PDA par les autorités publiques.

2.3. Problématiques d'accès à la justice

L'accès à la justice est rendu difficile en raison des multiples asymétries en matière d'information et de pouvoir. Comme expliqué plus haut, une des principales conclusions des travaux de recherche menés au niveau national montre que l'absence de cartographie claire et cohérente des utilisations des systèmes d'IA et de PDA par les administrations publiques compromet le droit à la non-discrimination et son application. Le fait de ne pas savoir qu'une décision administrative donnée a été prise avec (l'aide d') un système d'IA ou de PDA empêche le sujet de cette décision de soupçonner une discrimination algorithmique. Même en cas de suspicion, l'absence d'obligations généralisées imposant aux administrations publiques de fournir des informations sur le type de système utilisé, les facteurs de décision mis en œuvre et la manière dont le système est utilisé dans le processus décisionnel empêche les victimes de présenter des preuves suffisantes pour établir des présomptions de discrimination devant les tribunaux.²²

Il est vrai que les lois européennes sur la non-discrimination prévoient un partage de la charge de la preuve entre les demandeur.ses et les défendeur.ses afin d'alléger la charge des demandeur.ses lorsqu'ils invoquent une discrimination. La législation anti-discrimination de l'UE prévoit que « dès lors qu'une personne s'estime lésée par le non-respect à son égard du principe de l'égalité de traitement et établit,

22. Dans une certaine mesure, cette problématique peut être atténuée en faisant valoir le droit à l'explication prévu dans le RGPD, bien que les obstacles procéduraux et les coûts puissent être importants pour les particuliers qui souhaitent se porter victimes.

devant une juridiction ou une autre instance compétente, des faits qui permettent de présumer l'existence d'une discrimination directe ou indirecte, il incombe à la partie défenderesse de prouver qu'il n'y a pas eu violation du principe de l'égalité de traitement ».²³ De même, la Cour européenne des droits de l'homme transfère aux défendeurs la charge de prouver qu'il n'y a pas eu discrimination lorsqu'un.e demandeur.se présente des faits crédibles et cohérents à partir desquels on peut présumer qu'il y a eu discrimination : « La preuve peut découler de la coexistence d'inférences suffisamment fortes, claires et concordantes ou de présomptions de fait similaires non réfutées » (Cour européenne des droits de l'homme, 2007). En principe, ces règles facilitent la tâche des requérant-es en leur imposant uniquement d'établir une présomption de discrimination avant que la charge de la preuve ne puisse être transférée au ou à la défendeur-se. Il appartient alors au ou à la défendeur-se de réfuter la présomption de discrimination.

Toutefois, ces règles ne traitent pas suffisamment des asymétries de pouvoir et d'information qui se manifestent lorsqu'une décision est prise avec (l'aide d') un système d'IA ou de PDA, et peuvent ne pas être à la hauteur face à des cas de discrimination algorithmique. D'une part, les experts et expertes nationaux.ales signalent que les tribunaux fixent parfois un seuil de preuve élevé pour déclencher un transfert de la charge de la preuve sur la partie défenderesse. D'autre part, l'opacité algorithmique complique la tâche des particuliers qui tentent d'établir une présomption de discrimination algorithmique (voir partie 2.2. ci-dessus). Les asymétries en matière d'information et de pouvoir prennent également la forme de ce que Milano et Prunkl (2024) appellent la « fragmentation épistémique », qui est « une caractéristique structurelle des environnements à médiation algorithmique qui isolent les personnes », empêchant ainsi « les personnes et les communautés de partager et de comparer de manière significative leurs expériences ». Ainsi, la fragmentation épistémique « complique l'identification et la conceptualisation des préjudices émergents dans ces environnements », notamment parce que le droit européen en matière de non-discrimination repose sur la capacité des victimes à démontrer qu'elles ont fait l'objet d'un traitement différencié ou d'un désavantage particulier par rapport à d'autres personnes dans une situation similaire (ibid.). Ensemble, l'opacité algorithmique et la fragmentation épistémique rendent très difficiles les recours légaux contre l'IA et la PDA.

Ainsi, l'accès à la justice est compromis dans la pratique. Pour donner un exemple concret, les systèmes de PDA utilisés par plusieurs administrations de sécurité sociale pour détecter les fraudes en Europe créent une catégorie de « personnes suspectes » par nature, qui est tout à fait discriminatoire mais très difficile à contester pour ces personnes, avec des conséquences désastreuses sur leur vie (suspension des prestations sociales, demandes de remboursements indus, etc.). Bien que le droit à l'explication récemment confirmé par la Cour de justice dans le contexte de l'article 15, paragraphe 1, point h, et de l'article 22 du RGPD dans l'affaire Dun & Bradstreet Austria remédie en partie au problème (voir partie 1.4.4.), il subsiste des asymétries de pouvoir et d'information auxquelles il convient de remédier en

23. Voir article 8, paragraphe 1, de la directive 2000/43/CE. Cette disposition est présente dans toutes les directives anti-discrimination : 2000/78/CE, 2004/113/CE et 2006/54/CE.

allégeant encore davantage la charge de la preuve pour les demandeur.ses, en habilitant les organismes de promotion de l'égalité et les parties prenantes concernées à soutenir les victimes potentielles et en renforçant la prévention.²⁴ Cette question a été abordée dès 2020 par l'Assemblée parlementaire du Conseil de l'Europe, qui soulignait la nécessité de « veiller tout particulièrement à garantir la présomption d'innocence et à éviter d'imposer aux victimes de discrimination une charge de la preuve disproportionnée » (Conseil de l'Europe, Assemblée parlementaire 2020b). Les États membres y étaient appelés « à veiller à ce que les organismes de promotion de l'égalité soient pleinement habilités à traiter les problèmes concernant les principes d'égalité et de non-discrimination résultant d'une utilisation de l'IA ».

À cet égard, les pouvoirs conférés à certains organismes européens de promotion de l'égalité en matière de protection contre la discrimination au niveau collectif et structurel peuvent contribuer à combler ces lacunes. En effet, certains organismes de promotion de l'égalité peuvent contribuer à prévenir la discrimination algorithmique par le biais de pouvoirs de supervision, d'enquête et de sanction. Par ailleurs, certains organismes de promotion de l'égalité peuvent surveiller l'utilisation des systèmes d'IA et de PDA et contester les résultats discriminatoires devant les tribunaux au nom de l'intérêt public, sans qu'aucune victime particulière n'ait à être identifiée au préalable (voir Cour de justice de l'Union européenne, 2008). Bien que ce ne soit pas le cas pour tous les organismes européens de promotion de l'égalité, de tels pouvoirs peuvent atténuer les problématiques associées au fait de devoir s'appuyer sur des victimes particulières qui signalent des préjudices discriminatoires pour pouvoir ouvrir des enquêtes. Toutefois, pour être efficaces, ces pouvoirs doivent être étayés par des droits à l'information solides : les organismes de promotion de l'égalité doivent être habilités à demander et à obtenir des informations pertinentes et utiles sur la manière dont les systèmes d'IA et de PDA sont utilisés. Ces pouvoirs devraient être étendus aux situations où des systèmes d'IA et de PDA sont utilisés pour appuyer et aider la prise de décision, plutôt que d'être limités aux situations d'automatisation complète uniquement, afin de refléter le fait que la prise de décision hybride est la forme d'utilisation la plus répandue. La transposition des nouvelles directives sur les normes applicables aux organismes de promotion de l'égalité d'ici juin 2026 sera une bonne occasion de revoir les pouvoirs des organismes de promotion de l'égalité à la lumière de ces nouveaux défis. En outre, d'autres voies procédurales telles que les mécanismes de plainte collective (par exemple *actio popularis*, recours collectifs) pourraient atténuer les problématiques d'accès à la justice.

2.4. Application des dispositions existantes et futures

2.4.1. Coopération institutionnelle

Le système d'application établi en vertu du règlement européen sur l'IA est complexe et ne laisse que peu de place aux plaintes individuelles. Bien que les autorités de surveillance du marché puissent recevoir des plaintes, elles n'ont aucune obligation

24. Ceci est d'autant plus important que la directive sur la responsabilité en matière d'IA a été retirée. Cette directive proposait de faciliter l'établissement des présomptions et le transfert de la charge de la preuve, quoique dans une mesure limitée.

d'enquêter à cet égard. Dans cette optique, le rôle des organismes de promotion de l'égalité devient encore plus important. Le système d'application établi par le règlement européen sur l'IA prévoit une coopération multipartite entre les organismes de promotion de l'égalité, les autres autorités de surveillance des droits fondamentaux désignées en vertu de l'article 77, les autorités de surveillance du marché et les autres institutions concernées. En outre, les organismes chargés de l'égalité coopéreront avec d'autres institutions de défense des droits fondamentaux désignées en vertu de l'article 77 du règlement européen sur l'IA. Toutefois, selon certaines des institutions interrogées aux fins du présent rapport, il pourrait y avoir des difficultés importantes à cet égard. En particulier, des méthodologies de coopération doivent être élaborées et mises en œuvre, et des canaux de communication institutionnels doivent être assurés afin de promouvoir une approche fondée sur la complémentarité institutionnelle et l'intersectionnalité. En outre, les autorités compétentes, et notamment les guichets uniques et les autorités de surveillance du marché, doivent avoir une compréhension adéquate des risques de discrimination présentés par les technologies algorithmiques, ainsi que des moyens juridiques et institutionnels de lutte contre la discrimination.

2.4.2. Lacunes conceptuelles

D'autres problèmes d'application se posent en ce qui concerne les concepts et les procédures consacrés par la législation européenne anti-discrimination. Concrètement, le fonctionnement de l'IA, et en particulier de l'apprentissage automatique, repose sur des corrélations, et non sur la causalité. Cela accroît le risque de discrimination par variables de substitution, que les juges pourraient être réticents à traiter comme une discrimination directe même lorsque ces variables sont connues pour être en corrélation avec des motifs de discrimination protégés.²⁵ Par exemple, dans certains contextes, le lieu de résidence des individus est connu pour être fortement corrélé à l'origine raciale et ethnique. Au niveau européen, le lieu de résidence ne figure pas dans la liste des motifs protégés inscrits dans la législation sur la non-discrimination. Par conséquent, ce type de discrimination algorithmique par variables de substitution est susceptible d'être traitée comme une discrimination indirecte, sauf dans les cas où le lieu de résidence est considéré comme étant « inextricablement » ou « intrinsèquement lié » à un motif protégé tel que la race ou l'origine ethnique.

À cet égard, le droit dérivé de l'UE interdisant la discrimination pourrait s'avérer trop limité, car il ne le fait que par rapport à une liste « fermée » de motifs protégés : sexe, race ou origine ethnique, handicap, orientation sexuelle, religion ou croyances et âge. Les préjugés algorithmiques liés à d'autres caractéristiques sociales ne seraient jugés directement discriminatoires que s'ils sont considérés comme « intrinsèquement liés » à l'un de ces motifs protégés. Ces préjugés pourraient

25. Pour de plus amples informations sur la discrimination par variables de substitution, voir Bartoletti I. et Xenidis R. (2023), « Étude sur l'impact des systèmes d'intelligence artificielle, leur potentiel de promotion de l'égalité, y compris l'égalité de genre, et les risques qu'ils peuvent entraîner en matière de non-discrimination », disponible à l'adresse suivante : <https://edoc.coe.int/en/artificial-intelligence/11649-study-on-the-impact-of-artificial-intelligence-systems-their-potential-for-promoting-equality-including-gender-equality-and-the-risks-they-may-cause-in-relation-to-non-discrimination.html>, consulté le 10 novembre 2025, Éditions du Conseil de l'Europe, Strasbourg.

toutefois être traités comme une discrimination indirecte s'ils touchent un groupe lié à l'un de ces motifs protégés. D'autres préjudices algorithmiques sortiraient du champ d'application juridiquement protégé, à condition qu'ils ne soient pas pris en compte par des cadres législatifs nationaux spécifiques, qui peuvent aller au-delà du droit de l'Union.

En revanche, des listes dites « ouvertes » telles que les critères protégés consacrés à l'article 21, paragraphe 1, de la Charte européenne des droits fondamentaux et à l'article 14 de la Convention européenne des droits de l'homme pourraient offrir plus de souplesse dans la lutte contre la discrimination algorithmique par variables de substitution. En effet, elles permettent aux juges de conclure à une discrimination « fondée notamment sur », respectivement, « le sexe, la race, la couleur, les origines ethniques ou sociales, les caractéristiques génétiques, la langue, la religion ou les convictions, les opinions politiques ou toute autre opinion, l'appartenance à une minorité nationale, la fortune, la naissance, le handicap, l'âge ou l'orientation sexuelle » et « le sexe, la race, la couleur, la langue, la religion, les opinions politiques ou toutes autres opinions, l'origine nationale ou sociale, l'appartenance à une minorité nationale, la fortune, la naissance ou toute autre situation ».

La question de savoir comment qualifier la discrimination algorithmique par variables de substitution est importante. Si cette qualification ne change rien au regard de la législation anti-discrimination de la Convention en raison de l'existence d'un régime unitaire de justifications ouvertes pour les formes de discrimination directe et indirecte, le problème est différent au regard de la législation anti-discrimination de l'UE. En UE, la discrimination directe bénéficie d'un régime de justification plus strict fondé sur des exceptions fermées. Compte tenu du manque de transparence, d'opacité et de complexité des systèmes d'IA et de PDA, un régime de justification ouvert qui exige une analyse contextuelle n'est peut-être pas le plus adapté. Pourtant, justifier une discrimination directe est souvent difficile dans le contexte algorithmique. Comme cela a été dit ailleurs, au-delà des problèmes de justification, les concepts mêmes de discrimination directe et indirecte ne sont pas tout à fait adaptés pour refléter la discrimination algorithmique, qui mélange des caractéristiques des deux notions (Xenidis 2021).

Le niveau élevé de différenciation effectué par les systèmes d'IA et de PDA accroît également les risques de discrimination intersectionnelle, en particulier lorsque les utilisateurs combinent plusieurs catégories de risque qui font grimper leur score de risque (par exemple, les antécédents en matière de genre et de migration dans des exemples de prédictions de chômage biaisées). Pourtant, la discrimination intersectionnelle n'est toujours pas pleinement reconnue dans les directives anti-discrimination de l'UE (exception faite de la récente directive sur la transparence en matière d'égalité salariale 2023/970) et n'est pas uniformément interdite dans la législation nationale malgré les évolutions récentes. Par conséquent, les États membres devront revoir leur législation anti-discrimination et la modifier si nécessaire, de manière à ce qu'elle couvre tous les cas où une discrimination directe ou indirecte, y compris une discrimination par association et une discrimination intersectionnelle, peut être causée par l'utilisation de systèmes d'IA et de PDA.

2.4.3. Restrictions pour la collecte et le traitement des données

Enfin, les restrictions juridiques existantes encadrant la collecte de données sur l'égalité limitent les possibilités d'enquêter efficacement sur la discrimination algorithmique et de promouvoir l'égalité dans le cadre du recours à des systèmes d'IA et de PDA. Pour lutter efficacement contre la discrimination dans l'IA, il pourrait être nécessaire de plaider en faveur de procédures de collecte de données plus adéquates en matière d'égalité. D'autres problématiques liées à l'application de la législation concernent l'utilisation de catégories sensibles de données à caractère personnel. L'article 10(5) du règlement européen sur l'IA autorise exceptionnellement le traitement de catégories sensibles de données personnelles (définies à l'article 9 du RGPD), telles que les origines ethniques ou l'orientation sexuelle, afin d'assurer la détection des biais, mais cette possibilité ne s'étend pas aux systèmes d'IA qui ne relèvent pas de la catégorie des systèmes à haut risque. Cela pourrait présenter d'importantes limites lorsqu'il s'agit de détecter, de prévenir et d'atténuer les biais discriminatoires de systèmes considérés comme à faible risque, auxquels les obligations de non-discrimination s'appliquent transversalement. Par ailleurs, il subsiste des incertitudes quant à la possibilité que les organismes de promotion de l'égalité et les autres autorités de surveillance des droits fondamentaux se prévalent des exceptions prévues à l'article 9 du RGPD et à l'article 10(5) du règlement européen sur l'IA pour évaluer les impacts discriminatoires de certains systèmes d'IA. La discrimination algorithmique peut être plus difficile à détecter et à surveiller lorsque des catégories sensibles de données à caractère personnel – qui se chevauchent avec de nombreuses catégories (mais pas toutes) protégées par le droit communautaire et national en matière de non-discrimination – ne peuvent pas être utilisées à des fins de test, d'audit et d'évaluation.

2.5. “Dériskuer” les systèmes d'IA

La classification fondée sur le risque du règlement européen sur l'IA a fait l'objet de critiques. La liste des applications à haut risque figurant à l'annexe III du règlement a été décrite comme trop rigide face au contexte de l'évolution rapide des technologies. En outre, le règlement européen sur l'IA ne fournit pas d'orientations claires pour décider si une utilisation relevant des secteurs énumérés à l'annexe III présente ou non un risque élevé. Cela est particulièrement problématique à la lumière de l'article 6, paragraphe 3, du règlement européen sur l'IA, qui prévoit un certain nombre d'exceptions. Par exemple, les fournisseurs peuvent déclarer que les systèmes qui relèveraient normalement de la liste des secteurs à haut risque ne présentent en fait aucun risque significatif de préjudice pour la santé, la sécurité ou les droits fondamentaux des personnes physiques. Les systèmes d'IA ne sont pas considérés comme à haut risque lorsqu'ils sont utilisés pour « effectuer une tâche procédurale étroite » ou une « tâche préparatoire à une évaluation²⁶ », pour « améliorer le résultat d'une activité humaine préalablement réalisée » ou pour « détecter les constantes en matière de prise de décision ou les écarts par rapport aux constantes habituelles

26. Cependant, en vertu de l'article 6, paragraphe 3, point d, un système d'IA utilisé pour établir le profil de personnes physiques dans les zones énumérées à l'annexe III est systématiquement considéré comme à haut risque.

antérieures ou des écarts par rapport à des schémas décisionnels antérieurs et [qui ne sont] pas destiné[s] à remplacer ou influencer l'évaluation humaine préalablement réalisée, sans examen humain approprié ». Pour les fournisseurs qui considèrent que leurs systèmes ne sont pas à haut risque, il leur suffit de « documenter leur évaluation avant que ce système ne soit mis sur le marché ou mis en service » (article 6, paragraphe 4) et enregistrer le système conformément à l'article 49, paragraphe 2.

Il n'y a pas de contrôle systématique à ce stade, car les autorités nationales compétentes peuvent demander les évaluations effectuées, sans toutefois en avoir l'obligation systématique. Ainsi, cela crée une faille potentielle qui pourrait être utilisée à volonté pour « dérisquer » des systèmes d'IA afin d'éviter de se conformer aux exigences et aux garanties applicables aux systèmes à haut risque en vertu du règlement européen sur l'IA²⁷. Ici, l'approche de la Convention-cadre sur l'IA, qui est fondée sur les droits plutôt que sur les risques, pourrait s'avérer particulièrement utile pour étendre les garanties des droits fondamentaux aux applications d'IA et de PDA « dérisquées ». Une autre lacune potentielle qui pourrait être exploitée dans les deux ensembles de règlements concerne la définition de l'IA. Comme les deux textes s'appliquent à l'IA, par opposition, par exemple, aux systèmes simples fondés sur des règles, certains systèmes de PDA pourraient être exclus s'ils sont considérés comme n'entrant pas dans le champ d'application de la définition de l'IA (Commission européenne, 2025d). Cette autre faille pourrait également être exploitée par les fournisseurs de systèmes de PDA plus simples pour se soustraire aux obligations du règlement européen sur l'IA.

2.6. Les normes techniques, les organismes de promotion de l'égalité et la question de l'harmonisation

Les normes harmonisées en matière d'IA qui seront publiées par le CEN-CENELEC constitueront un dispositif essentiel pour appuyer le respect des droits fondamentaux, y compris via la prévention et la protection contre la discrimination algorithmique. La question clé reste toutefois de savoir quels risques de discrimination les normes obligeront-elles les fournisseurs à identifier, évaluer et traiter ? En particulier, quelles formes de discrimination les fournisseurs devront-ils prendre en compte et comment devront-ils les rectifier pour être considérés comme conformes au règlement européen sur l'IA ? Plusieurs problèmes se posent ici. Le premier est celui du pouvoir confié aux organisations privées impliquées dans le processus de corégulation. La plupart des organisations impliquées dans le JTC 21 du CEN-CENELEC, en charge de l'élaboration de ces normes, sont des acteurs de l'industrie qui disposent des ressources nécessaires pour financer une telle participation. Les OSC et les organisations ayant un intérêt légitime peuvent participer dans une certaine mesure (Equinet participe par exemple actuellement au processus de normalisation), mais ont souvent des moyens financiers limités pour le faire. Cela signifie également que les compétences sur le terrain peuvent être inégalement réparties, avec un niveau de compétences techniques adéquat et des compétences juridiques plus limitées, en particulier en matière de droits humains. Les projets de normes ne sont pas publics, ce qui rend difficiles le suivi et l'évaluation de l'efficacité et des lacunes des stratégies

27. Ces exigences et garanties sont énumérées dans la section sur les systèmes d'IA à haut risque.

d'application qui sont actuellement à l'étude. Cela préempte les tentatives visant à influencer ces discussions pour appréhender les limites existantes.

Le deuxième problème est celui de la compatibilité et de la complémentarité de cette approche avec celle des droits fondamentaux garantis par la législation de l'UE, notamment par droit à la non-discrimination. D'une part, le niveau d'atténuation des risques que les normes prescriront à des fins de conformité reste encore à clarifier. Il s'agit d'une question clé, car l'article 9, paragraphe 5, du règlement européen sur l'IA prévoit que « [l]es mesures de gestion des risques [soient] telles que le risque résiduel pertinent associé à chaque danger, ainsi que le risque résiduel global lié aux systèmes d'IA à haut risque sont jugés acceptables ». En ce qui concerne les biais et la discrimination, les risques résiduels acceptables au titre du règlement européen sur l'IA ne garantiront pas le respect de la législation européenne en matière de droits fondamentaux, qui adopte une approche différente centrée sur les préjudices réels. Toutefois, l'article 52, paragraphe 1, de la Charte des droits fondamentaux de l'UE stipule que « [t]oute limitation de l'exercice des droits et libertés reconnus par la présente Charte doit être prévue par la loi et respecter le contenu essentiel desdits droits et libertés » et que « [d]ans le respect du principe de proportionnalité, des limitations ne peuvent être apportées que si elles sont nécessaires et répondent effectivement à des objectifs d'intérêt général reconnus par l'Union ou au besoin de protection des droits et libertés d'autrui ». La portée des ingérences permises – limitées par l'essence des droits fondamentaux et conditionnées par la nécessité et l'intérêt général – semble être en conflit avec l'approche du risque résiduel adoptée dans le règlement européen sur l'IA. À tout le moins, la relation entre ces concepts est ambiguë et reste à clarifier, éventuellement par la Cour de justice.

D'autre part, l'approche de normalisation adoptée par l'UE, combinée au cadre d'harmonisation maximale (au moins partielle) inscrit dans le nouveau règlement de l'UE sur l'IA, entre en conflit avec l'approche d'harmonisation minimale des directives de l'UE sur la non-discrimination. Alors que le droit dérivé de l'Union interdit six motifs de discrimination (race et origine ethnique, sexe et genre, handicap, orientation sexuelle, religion ou convictions, et âge), il permet également aux États membres d'aller au-delà de ces exigences minimales et d'interdire plus largement la discrimination par des « dispositions plus favorables », pour autant qu'elles respectent les traités de l'UE. Pourtant, il reste là aussi à clarifier quelles formes de biais les fournisseurs de systèmes à haut risque devront tester. Par exemple, si des tests et la prise de mesures d'atténuation des biais ne sont requis que pour les six motifs protégés au niveau de l'UE, et que les fournisseurs qui respectent ces normes sont alors présumés respecter le règlement européen sur l'IA, cela leur permettra de vendre leurs produits sur l'ensemble du marché intérieur. Même si les dépoyeurs peuvent encore être tenus responsables de discriminations en vertu des lois anti-discrimination de l'UE et des États membres qui imposent une protection plus large, des questions se posent à au moins deux égards.

Premièrement, des dispositions nationales plus favorables interdisant la discrimination sont explicitement autorisées par le droit communautaire en matière de discrimination et ne devraient pas être considérées comme des obstacles à la libre circulation. Cette approche serait également contraire à l'article 21, paragraphe 1 de la Charte des droits fondamentaux de l'Union européenne, qui interdit « toute discrimination

fondée notamment sur le sexe, la race, la couleur, les origines ethniques ou sociales, les caractéristiques génétiques, la langue, la religion ou les convictions, les opinions politiques ou toute autre opinion, l'appartenance à une minorité nationale, la fortune, la naissance, un handicap, l'âge ou l'orientation sexuelle ». Un autre scénario consisterait à inclure la définition plus large de la discrimination consacrée dans la Charte. Or, il subsisterait des tensions avec une législation nationale plus protectrice, expressément autorisée par la législation de l'UE en matière de discrimination. Dans le même temps, si les normes ne mentionnent pas les motifs de discrimination interdits dont les fournisseurs de systèmes d'IA à haut risque doivent tenir compte lorsqu'ils procèdent à des tests de partialité ou évaluent l'impact de leur système sur les droits fondamentaux, il existe un risque important que ces procédures finissent par être largement inefficaces.

La deuxième question a trait à l'influence de ces mécanismes sur le transfert de la charge de la preuve dans les affaires de discrimination. Il est important de rappeler ici que la présomption de conformité prévue à l'article 40 du règlement européen sur l'IA ne s'applique qu'aux exigences du règlement pour les systèmes à haut risque (voir la partie sur le système de gestion des risques), et non à la législation anti-discrimination ou relative aux droits fondamentaux. À cet égard, l'article 82, paragraphe 1, du règlement européen sur l'IA reconnaît que « bien qu'un système d'IA à haut risque soit conforme au présent règlement, il [peut néanmoins] présente[r] un risque pour la santé ou la sécurité des personnes, pour les droits fondamentaux ».

Par conséquent, il convient de souligner que l'approche d'harmonisation maximale (partielle) du règlement européen sur l'IA ne devrait pas être interprétée comme annulant ou altérant les obligations existantes en matière d'anti-discrimination établies en vertu du droit primaire et secondaire de l'UE.

Qu'est-ce que cela signifie pour les organismes chargés de l'égalité et les autorités compétentes qui protègent les droits fondamentaux, comme ceux désignés en vertu de l'article 77 du règlement européen sur l'IA ? Premièrement, il peut être difficile d'évaluer la conformité des systèmes de gestion des risques et des évaluations d'impact sur les droits fondamentaux vis-à-vis de la législation anti-discrimination si ces systèmes ne fournissent pas d'informations sur les motifs de discrimination qui ont été testés. La question qui se pose est de savoir si les organismes de promotion de l'égalité et les parties prenantes concernées chargées de faire respecter les droits fondamentaux disposeront d'informations suffisamment précises pour évaluer la manière dont la discrimination a été empêchée et/ou dont les risques de discrimination ont été atténués en ce qui concerne les groupes légalement protégés dans le contexte des systèmes d'IA à haut risque, et comment ces groupes ont été définis au cours des procédures d'évaluation des risques. Deuxièmement, la dissonance juridique soulignée ci-dessus crée un décalage entre, d'une part, les obligations des fournisseurs (et, le cas échéant, des déployeurs) de prévenir et de traiter les risques de biais discriminatoires dans le contexte des systèmes à haut risque, et d'autre part, le mandat des organismes de promotion de l'égalité et des autres autorités compétentes. La plupart des organismes de promotion de l'égalité sont mandatés pour lutter contre les discriminations fondées sur tous les motifs protégés par le droit national, qui vont souvent au-delà des motifs protégés par le droit anti-discrimination dérivé de l'UE. En fonction de l'approche privilégiée dans

les nouvelles normes ou dans leur application (par exemple, le plus petit dénominateur commun ou des types de biais non spécifiés), les organismes chargés de l'égalité devront être particulièrement vigilants en ce qui concerne les motifs de discrimination qui ne sont interdits que par le droit national. En effet, les risques de discrimination algorithmique peuvent ne pas avoir été testés par les fournisseurs vis-à-vis de ces motifs protégés au niveau national.

En outre, la normalisation présente plusieurs défis pour les organismes de promotion de l'égalité. Il peut être important que ces organismes contrôlent et influencent la manière dont les nouvelles normes harmonisées en matière d'IA seront définies par le CEN-CENELEC, car le respect de ces normes déclenchera une présomption de conformité aux exigences substantielles du règlement européen sur l'IA pour les fournisseurs de systèmes d'IA à haut risque. Plusieurs difficultés se posent cependant. Premièrement, même si Equinet est impliqué dans certains sous-groupes du JTC 21 chargés de rédiger ces normes sur l'IA, le degré de publicité et d'accessibilité des normes convenues reste flou. Dans l'affaire *Public.Resource.Org et Right to Know c. Commission* et autres, la Cour de justice a décidé que les normes devaient être rendues publiquement accessibles aux citoyens de l'Union européenne (Cour de justice de l'Union européenne 2024). Étant donné que le litige portait sur l'accès à des normes relatives à la sécurité des jouets, il reste à déterminer si cette exigence de transparence s'appliquera de manière similaire aux normes relatives à l'IA et si elles seront accessibles gratuitement.

Deuxièmement, une expertise technique peut être nécessaire pour comprendre et évaluer de manière critique si les normes d'IA garantissent suffisamment le respect des droits fondamentaux et de la législation en matière de non-discrimination. Selon un rapport d'Equinet, les normes « détermineront quelles informations et quel niveau de détail sont inclus dans [la] documentation » produite en vertu des exigences substantielles pour les systèmes d'IA à haut risque en vertu du règlement européen sur l'IA et, par conséquent, « le type de gestion des risques et de surveillance après commercialisation que les fournisseurs doivent effectuer pour anticiper, identifier et atténuer les risques pour les droits fondamentaux » (Mittelstadt 2025).

Ainsi, les organismes de promotion de l'égalité devraient être particulièrement attentifs aux choix techniques mis en œuvre par les nouvelles normes d'IA :

- ▶ Quels indicateurs d'équité sont utilisés pour mesurer les performances ? Quels sont les seuils fixés pour quantifier les risques de biais discriminatoires ? Quelles sont les mesures de débiaisage et les méthodes d'équité utilisées ? Sont-elles alignées sur la législation et la jurisprudence européennes en matière de lutte contre la discrimination ?
- ▶ Comment sont définis les niveaux acceptables de « risques résiduels » en matière de discrimination ?
- ▶ Comment la finalité prévue des systèmes, le niveau d'exactitude attendu et les utilisations abusives prévisibles sont-ils définis ?
- ▶ Comment les risques et impacts discriminatoires sont-ils mesurés et atténués vis-à-vis des groupes protégés par la législation européenne anti-discrimination ?
- ▶ Comment les groupes particuliers sont-ils définis dans la documentation technique, la gestion des risques et les analyses d'impact ?

- La discrimination intersectionnelle est-elle adéquatement prise en compte ?
- Les définitions techniques des groupes couvrent-elles l'ensemble de la population protégée par des motifs de discrimination interdits ?

Être attentif à ces questions est important pour éviter l'« éthique-washing » et le respect de façade des droits fondamentaux.

Ceci n'est pas facile pour les organismes de promotion de l'égalité, qui pourraient être amenés à chercher à acquérir une nouvelle expertise, de nouveaux savoir-faire et de nouvelles ressources pour pouvoir traiter à ces questions. Créer des modèles de coopération solides avec les autorités compétentes (autorités de surveillance du marché, points de contact uniques, Bureau de l'IA, etc.) et les organismes possédant une expertise complémentaire (agences de protection des données, médiateurs, offices de protection des consommateurs, etc.) ainsi qu'avec des tiers concernés (chercheurs, OSC, etc.) peut aider les organismes de promotion de l'égalité à remédier aux écueils concernant l'expertise et les ressources nécessaires.

Résumé – Utilisation de l’IA dans l’administration publique belge et recommandations politiques sur la discrimination algorithmique



Le résumé est basé sur le rapport d'analyse des lacunes et de l'état actuel (« As-is and gap analysis »), dans lequel la partie belge a été rédigée par Ine van Zeeland.²⁸

Utilisation de l’IA dans l’administration publique

L’administration publique belge utilise plusieurs applications d’intelligence artificielle (IA) et mène actuellement divers projets d’IA, notamment dans les secteurs de la police, de la finance, de l’emploi et de l’éducation. Face à l’absence d’un registre public complet recensant ces systèmes, il n’est pas possible de cartographier de manière exhaustive et cohérente les utilisations de systèmes d’IA et de prise de décision automatisée (PDA) dans les services publics belges. Le Service public fédéral (SPF) Stratégie et Appui (BOSA) héberge un Observatoire de l’IA, qui a créé des initiatives visant à collecter des informations sur l’utilisation de systèmes d’IA ou de PDA dans les services publics en Belgique. Des démarches similaires de collecte d’informations sont menées ou prévues par d’autres organismes tels que l’Institut FARI de Bruxelles et le SPF Sécurité sociale.

28. Ceci est un rapport confidentiel qui examine l’utilisation des systèmes d’IA et de PDA dans le secteur public en Belgique, évalue les cadres juridiques et politiques liés à l’IA du point de vue de l’égalité et de la non-discrimination, et formule des recommandations à l’intention des autorités et des décideurs-euses politiques.

Législation et jurisprudence en la matière

Le règlement européen de l'UE sur l'IA et la législation européenne et belge en matière de protection des données et de lutte contre la discrimination constituent un cadre juridique pour lutter contre la discrimination algorithmique. La jurisprudence belge concernant l'IA et la PDA et la discrimination est relativement balbutiante. Le rapport d'analyse des lacunes et de l'état actuel (« As-is and gap analysis ») a identifié quelques affaires concernant ces problématiques, comme la décision de la Cour de justice dans l'affaire C-817/19 Ligue des droits humains ASBL c. Conseil des ministres,²⁹ qui a statué sur la tension entre le traitement automatisé de données par des algorithmes d'apprentissage automatique et la non-discrimination.

En ce qui concerne la protection des données, la Chambre contentieuse du GBA-APD a statué à deux reprises sur des affaires liées à des systèmes de PDA potentiellement pertinentes pour la lutte contre la discrimination : une affaire concernait une notation de crédit par une société de covoiturage³⁰ et l'autre concernait le développement d'un modèle algorithmique basé sur les données des cartes de paiement.³¹ Cependant, dans les deux affaires, les décisions de la Chambre contentieuse concernaient des exigences du RGPD qui n'avaient aucune incidence directe sur la discrimination. Dans l'ensemble, il peut y avoir des cas d'inégalité de traitement par des systèmes d'IA et de PDA non détectés par les tribunaux. En effet, en 2022, le Tribunal de première instance de Bruxelles a réglé une affaire de discrimination indirecte par un système de PDA dans le contexte de « scan-cars » (voitures de contrôle automatisé du stationnement) déployées par Parking.³²

Outre l'acquis de l'UE, d'autres possibilités de prévention et de lutte contre la discrimination algorithmique sont prévues dans la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits humains, la démocratie et l'État de droit, ouverte à la signature en septembre 2024, dans laquelle l'article 10 sur l'égalité et la non-discrimination interdit la discrimination et appelle les États à « maintenir des mesures visant à supprimer les inégalités [...] afin d'obtenir des résultats impartiaux, justes et équitables ».

Lacunes dans les cadres juridiques et politiques

Manque de transparence

En Belgique, l'un des problèmes critiques concerne le manque de transparence et de reporting sur l'utilisation des systèmes d'IA et de PDA, car il n'existe actuellement aucune obligation formelle pour les organisations du secteur public de déclarer ou d'enregistrer leur utilisation de systèmes d'IA et de PDA. L'absence d'une base de

29. C-817/19 Ligue des droits humains ASBL c. Conseil des ministres EU:C:2022:491.

30. Décision 168/2023 du 19 décembre 2023, Chambre contentieuse GBA-APD

31. Décision quant au fond 46/2024 du 15 mars 2024, Chambre contentieuse GBA-APD.

32. Affaire 21/6495/C, disponible à l'adresse : www.unia.be/fr/legislation-et-jurisprudence/jurisprudence/tribunal-de-premiere-instance-de-bruxelles-le-2-mai-20221, consulté le 10 novembre 2025. Dans cette affaire, les plaignants étaient une personne handicapée et le Collectif accessibilité Wallonie Bruxelles (CAWaB), soutenu par l'organisme interfédéral de promotion de l'égalité Unia.

données complète et accessible au public présente des risques tant pour le contrôle que pour la responsabilisation, et complique également la tâche des organismes chargés de l'égalité consistant à surveiller les risques de discrimination ou à aider les personnes qui risquent d'être victimes de discrimination de la part de ces systèmes. Les recherches menées dans ce domaine révèlent systématiquement que les cas de discrimination algorithmique sont sous-déclarés et sous-examinés.

Manque de sensibilisation et absence d'approche harmonisée des droits fondamentaux

Malgré les diverses formations et directives éthiques, la connaissance systématique des risques liés à la discrimination est limitée dans de nombreux services publics. Dans l'ensemble, l'administration publique ne dispose d'aucune approche standardisée pour évaluer les risques liés aux droits fondamentaux et les discriminations liés à l'utilisation de l'IA dans les services publics. Il manque une approche harmonisée pour intégrer les points de vue des groupes ou des personnes touchés, et ceux des experts en droits fondamentaux dans la conception, le développement et le déploiement des systèmes d'IA et de PDA.

Un paysage fragmenté en matière de surveillance

Le paysage institutionnel de surveillance belge est complexe, avec des mandats répartis entre les niveaux fédéral, régional et communautaire. Cela crée des obstacles à l'harmonisation de la surveillance et des interprétations juridiques, et introduit le risque que le règlement européen sur l'IA soit appliqué de manière hétérogène dans les différents services publics belges.

Lacunes en matière de ressources et d'expertise

Les ressources financières, humaines et techniques publiques limitées des organismes de promotion de l'égalité et les autorités de surveillance belges représentent des difficultés importantes pour lutter contre la discrimination induite par l'IA et la PDA.

Recommandations clés

- **Mettre en place un registre public des systèmes d'IA et de PDA :** Introduire une obligation légale de signaler toute utilisation des systèmes d'IA et de PDA dans le secteur public dans un registre ou un référentiel national disposant d'un certain niveau d'accès public, en plus du registre européen, afin de garantir la transparence et l'efficacité de la surveillance.
- **Mettre en place des mécanismes avancés de coordination entre les autorités nationales compétentes à différents niveaux :** Faciliter et mettre au point des accords de coopération formels pour garantir une harmonisation de la supervision et des décisions au niveau des autorités compétentes, et créer des lignes directrices en faveur d'une implication mutuelle afin de garantir que les autorités du règlement européen sur l'IA et les structures nationales de protection des droits humains visées à l'article 77 soient incluses de manière proactive dans les enquêtes.
- **Créer un réseau pour partager les pratiques prometteuses et renforcer les synergies :** Aider les pouvoirs publics dans la mise en place d'un réseau structurel

de coopération interfédérale autour de conseils pratiques et d'orientations et pour tirer des enseignements des meilleures pratiques. Cela pourrait permettre de remédier à la fragmentation des connaissances et des capacités pertinentes nécessaires pour développer ou acquérir des systèmes d'IA et de PDA qui intègrent des mesures de protection des droits fondamentaux dès leur conception, ou permettent tout au moins un suivi de leur impact tout au long du cycle de vie du système. Ce réseau devrait réunir des expert.es de la lutte contre la discrimination, des droits fondamentaux et de la protection des données à caractère personnel.

- ▶ **Assurer des garanties procédurales pour rendre les recours plus efficaces dans les affaires impliquant des systèmes d'IA ou de PDA, y compris en modifiant la législation sur la non-discrimination au besoin :** Rendre possible un transfert de la charge de la preuve en droit en introduisant une présomption de biais algorithmique qui aiderait les personnes confrontées à des problématiques quasi insolubles en matière de preuves. Si cela est nécessaire, la législation en matière de non-discrimination devrait être modifiée pour étayer cette présomption. Dans le même temps, il convient de faciliter la mise en place de procédures de recours collectif par les organisations d'intérêt public dans des contextes tels que l'emploi, la santé et la sécurité sociale, comme c'est le cas dans les procédures de règlement et de plainte axées sur les consommateurs.
- ▶ **Élaborer et publier des synthèses des analyses d'impact sur les droits fondamentaux :** Introduire une approche harmonisée des analyses d'impact sur les droits fondamentaux (AIDF). Cette approche devrait être appliquée à toutes les organisations du secteur public et inclure les points de vue des groupes touchés et des organismes de promotion de l'égalité, avec publication des synthèses des AIDF.
- ▶ **Inclure la lutte contre la discrimination dans les marchés publics de systèmes d'IA et de PDA :** Veiller à ce que les exigences anti-discrimination soient intégrées aux marchés publics d'IA et de PDA afin d'améliorer les vérifications préalables.
- ▶ **Sensibiliser le public à l'utilisation de l'IA et améliorer leur maîtrise de l'IA :** La sensibilisation du public à l'utilisation de l'IA dans les services publics devrait être renforcée par des campagnes de sensibilisation et des démarches de maîtrise de l'IA, y compris dans le domaine de l'éducation.
- ▶ **Garantir la disponibilité de ressources publiques et d'expertise suffisantes pour les organismes de promotion de l'égalité et les autorités de surveillance en matière d'utilisation de l'IA :** Celles-ci sont indispensables pour permettre aux organismes de promotion de l'égalité et aux autorités de contrôle de s'acquitter de leur mandat en vertu du règlement européen sur l'IA et de la législation anti-discrimination.
- ▶ **Former le personnel du secteur public sur les impacts de l'IA et de la PDA sur les droits fondamentaux :** La formation doit aller au-delà des questions d'éthique pour aborder les impacts de l'IA sur les droits fondamentaux, en accordant une attention particulière aux risques de discrimination.
- ▶ **Renforcer la collaboration entre les organismes de promotion de l'égalité, les autres institutions publiques et la société civile :** Renforcer et institutionnaliser la collaboration, par exemple par le biais de « communautés de pratique » qui allègent les obstacles pour les personnes lésées qui recherchent de l'aide.

Résumé – Utilisation de l’IA dans l’administration publique finlandaise et recommandations politiques sur la discrimination algorithmique



Ce résumé est basé sur le rapport d'analyse des lacunes et de l'état actuel (« As-is and gap analysis »), dans lequel la partie finlandaise a été rédigée par Emeline Banzuzi.³³

Recours établi à la prise de décision automatisée, utilisation émergente de l'intelligence artificielle dans l'administration publique

En Finlande, l'utilisation de systèmes de prise de décision automatisée par les administrations publiques est déjà installée depuis longtemps et sert à traiter efficacement un grand nombre de décisions administratives, en particulier dans des domaines tels que la migration, les activités répressives, l'emploi et le recrutement. En revanche, le recours à l'intelligence artificielle (IA) demeure limité, et se concentre principalement à l'utilisation d'IA génératives à des fins d'efficacité, avec notamment l'utilisation de Microsoft CoPilot pour faciliter la rédaction de documents ou la prise de notes. Plusieurs autorités publiques explorent néanmoins ou expérimentent des applications d'IA : c'est le cas du service finlandais de l'immigration (Migri), qui envisage de recourir à l'IA en particulier pour la transcription de texte verbal en texte écrit.

33. Ceci est un rapport confidentiel qui examine l'utilisation des systèmes d'IA et de PDA dans le secteur public en Finlande, évalue les cadres juridiques et politiques liés à l'IA du point de vue de l'égalité et de la non-discrimination, et formule des recommandations à l'intention des autorités et des décideurs-euses politiques.

Par ailleurs, la police finlandaise utilise l'IA pour lire les plaques d'immatriculation, et la ville d'Helsinki expérimente actuellement « Helsinki GPT », un outil d'IA conçu pour classer et prioriser automatiquement les retours clients et générer des réponses.

Cadre juridique et institutionnel

Le règlement européen de l'UE sur l'IA et la législation européenne finlandaise en matière de protection des données et de lutte contre la discrimination constituent un cadre juridique pour lutter contre la discrimination algorithmique. Les autorités finlandaises se préparent à adapter leurs cadres juridiques et institutionnels aux obligations introduites par le règlement sur l'IA. Deux propositions législatives sont en cours pour mettre en œuvre le règlement européen sur l'IA en Finlande. La première vise, par exemple, à réglementer la surveillance de certains systèmes d'IA et l'attribution d'amendes, mais également à désigner dans la législation nationale les autorités compétentes et un point de contact unique, comme l'exige le règlement européen sur l'IA. La seconde vise à réglementer, par exemple, la mise en place et le fonctionnement des « bacs à sable sur l'IA ». Les autres travaux préparatoires institutionnels en cours comprennent la mise en place de mécanismes de coordination et la clarification des rôles de supervision.

En Finlande, le recours à la PDA est réglementé séparément dans la loi administrative du pays, qui ne permet l'utilisation de la PDA que dans les cas où aucun jugement spécifique au cas en question n'est nécessaire pour prendre la décision ou lorsqu'un ou une agent.e a résolu tout aspect nécessitant une décision spécifique au cas en question. La PDA doit être basée sur des règles prédéterminées par une personne physique.

La législation finlandaise en matière de non-discrimination et d'égalité est relativement complète, car elle renforce la promotion de l'égalité plutôt que de se concentrer uniquement sur la prévention de la discrimination. L'obligation de promouvoir l'égalité impose aux pouvoirs publics de prendre des mesures actives et anticipatives pour assurer l'égalité de traitement, et cette obligation s'applique également à la conception, au développement et à l'utilisation des systèmes d'IA et de PDA. Outre l'acquis de l'UE, la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits humains, la démocratie et l'État de droit, ouverte à la signature en septembre 2024 et signée par l'UE, apporte d'autres possibilités dans la prévention et la lutte contre la discrimination algorithmique, avec son article 10 sur l'égalité et la non-discrimination, qui interdit la discrimination et appelle les États à « maintenir des mesures visant à supprimer les inégalités [...] afin d'obtenir des résultats impartiaux, justes et équitables ».

Notamment, le seul arrêt bien connu concernant la discrimination et la PDA en Finlande est une affaire qui concernait une société de services financiers dans laquelle le tribunal national de non-discrimination et d'égalité de Finlande a ordonné à la

société de cesser d'utiliser une méthode statistique de notation de crédit fondée sur des motifs de discrimination interdits, à savoir le lieu de résidence, l'âge et le sexe.³⁴

Lacunes dans les cadres juridiques et politiques

Nécessité d'une approche normalisée pour évaluer les impacts sur l'égalité et la non-discrimination

Le rapport constate que les analyses d'impact sur l'égalité et la non-discrimination ne sont pas systématiquement menées sur les systèmes d'IA ou de PDA, bien que l'obligation légale de promouvoir l'égalité et d'évaluer les impacts sur l'égalité s'applique aux autorités publiques. Cependant, un cadre national d'évaluation élaboré dans le cadre des activités d'analyse, d'évaluation et de recherche du gouvernement finlandais³⁵ fournit un outil existant pour évaluer les systèmes d'IA. Une méthodologie commune pour réaliser les analyses d'impact sur l'égalité et les droits fondamentaux (AIDF) est nécessaire pour garantir l'intégration systématique des considérations d'égalité tout au long du cycle de vie des systèmes d'IA et de PDA. Le Médiateur pour la non-discrimination a exhorté le gouvernement à veiller à ce que les autorités publiques disposent d'une méthode efficace et systématiquement appliquée pour évaluer les impacts sur l'égalité, y compris dans les marchés publics.³⁶

Connaissance limitée des risques de discrimination

Bien que les pouvoirs publics aient reçu une formation et élaboré des lignes directrices éthiques, ceux-ci mériteraient de disposer de ressources supplémentaires, d'une formation ciblée complémentaire et d'une meilleure compréhension des lois sur la discrimination algorithmique et la non-discrimination dans le cadre de leurs obligations en matière de droits fondamentaux.

Ressources limitées et manque de compétences techniques fondées sur les droits humains pour les organismes chargés de l'égalité et les autorités de contrôle

Les organismes chargés de l'égalité et les autres autorités chargées de la surveillance des droits fondamentaux devraient jouer un rôle plus important en vertu de l'article 77 du règlement européen sur l'IA, qui leur octroie des pouvoirs d'accès à

34. Décision 216/2017 du tribunal national de non-discrimination et d'égalité de Finlande. Disponible à l'adresse : www.yvtltk.fi/material/sites/ytltk/tapausselosteet/zedi26zya/Decision_YVTltk_xx16_2017_Multiple_discrimination_in_the_assessment_of_creditworthiness.pdf (en anglais), consulté le 10 novembre 2025.

35. Ojanen A. et al. (2022), "Promoting equality in the use of Artificial Intelligence – An assessment framework for non-discriminatory AI", Policy Brief 2022:25 (en anglais).

36. « Recommandations du Médiateur pour la non-discrimination pour le programme gouvernemental 2023-2027 » (en finnois : Yhdenvertaisuusvaltuutetun suosituksset hallitusohjelmaan 2023-2027), disponible en anglais à l'adresse suivante : [https://yhdenvertaisuusvaltuutettu.fi/documents/25249352/143315235/Yhdenvertaisuusvaltuutetun+suositukset+hallitusohjelmaan+\(PDF\).pdf/dd8d4ddf-2ab7-9c25-437d-c6619da62712/Yhdenvertaisuusvaltuutetun+suositukset+hallitusohjelmaan+\(PDF\).pdf?t=1681804769918](https://yhdenvertaisuusvaltuutettu.fi/documents/25249352/143315235/Yhdenvertaisuusvaltuutetun+suositukset+hallitusohjelmaan+(PDF).pdf/dd8d4ddf-2ab7-9c25-437d-c6619da62712/Yhdenvertaisuusvaltuutetun+suositukset+hallitusohjelmaan+(PDF).pdf?t=1681804769918), consulté le 10 novembre 2025.

la documentation et de demande de test des systèmes à haut risque. Cependant, leur capacité à remplir ce rôle élargi peut être limitée par le manque de ressources humaines, techniques et financières, alors même que le recours aux systèmes d'IA augmente, dans le contexte d'une sensibilisation et d'une maîtrise de l'IA encore limitée au niveau du grand public. En outre, une connaissance approfondie de la protection des droits fondamentaux, notamment avec l'identification et l'évaluation des différentes formes de discrimination, est requise pour tous les acteurs chargés de tâches de supervision en vertu du règlement européen sur l'IA. Les ressources devraient également être augmentées pour assurer un mécanisme de prévention efficace afin de garantir l'égalité avant que les risques de discrimination ne se matérialisent.

Absence de registre public et harmonisé des systèmes d'IA et de PDA

Il n'existe pas de base de données nationale publique unique compilant les informations sur l'utilisation des systèmes d'IA ou de PDA dans les administrations publiques finlandaises. Cependant, l'utilisation de la PDA est façonnée par les obligations de transparence de la Loi administrative finlandaise : les autorités doivent publier des informations générales sur l'utilisation de PDA sur leurs sites web et informer les personnes chaque fois que des décisions les concernant ont été prises par l'intermédiaire de la PDA. Bien que les agences doivent publier des informations concernant ces utilisations sur leurs propres sites Web, ces transmissions d'informations restent fragmentées.

Surveillance fragmentée et coordination institutionnelle

Les responsabilités en matière de surveillance sont actuellement réparties entre plusieurs autorités agissant en tant qu'autorités nationales compétentes, telles que les autorités de surveillance du marché et les autorités chargées des droits fondamentaux. Bien que cette répartition garantisse une couverture intersectorielle, elle risque en revanche de donner lieu à des interprétations incohérentes et des lacunes en matière de mise en œuvre pour garantir la protection des droits fondamentaux, en particulier si les mécanismes de coopération ne sont pas formalisés et si les autorités ne disposent pas d'une expertise suffisante en matière de non-discrimination.

Recommandations clés

- **Veiller à ce que la législation sur la non-discrimination soit dûment prise en compte et mise en œuvre, y compris les obligations de promotion de l'égalité :** Pour ce faire, il convient de veiller à ce que les autorités compétentes disposent de compétences suffisantes en matière de non-discrimination, de dispenser des formations et d'améliorer leur connaissance et leur compréhension de ces obligations. Une expertise et des conseils suffisants en matière de non-discrimination devraient également être apportés aux personnes impliquées ou à celles qui supervisent les achats et les subventions liés à l'achat et à l'utilisation de systèmes d'IA.

- ▶ **Introduire un cadre normalisé pour les analyses d'impact sur l'égalité et les droits fondamentaux dans le cycle de vie des systèmes d'IA :** Ces analyses devraient être appliquées à toutes les autorités publiques, avec la participation des organismes de promotion de l'égalité, de la société civile et des groupes impactés. Des synthèses des analyses devraient être publiées pour renforcer la responsabilisation.
- ▶ **Augmenter les allocations de ressources aux organismes de promotion de l'égalité et aux autorités de contrôle :** Les autorités nationales compétentes chargées de faire appliquer le règlement européen sur l'IA devraient disposer de ressources suffisantes pour s'acquitter efficacement de leurs fonctions de surveillance et participer de manière significative au cadre national de gouvernance de l'IA.
- ▶ **Mettre en place un registre public national des systèmes d'IA et de PDA utilisés par les autorités publiques, en complément de la base de données à l'échelle de l'UE prévue par le règlement européen sur l'IA :** Ce registre devrait permettre d'améliorer la transparence, de favoriser la surveillance et aider les organismes de promotion de l'égalité et le public à surveiller les risques de discrimination.
- ▶ **Normaliser les mécanismes de coordination et de coopération entre autorités compétentes, notamment en matière de protection et de promotion de l'égalité :** Il faudrait notamment doter les autorités compétentes de compétences suffisantes en matière d'égalité et de non-discrimination pour assurer une application cohérente du règlement européen sur l'IA. Il est important de renforcer la coopération entre le Médiateur chargé de la non-discrimination et le Médiateur chargé de la protection des données en matière de discrimination algorithmique, car les cas qui surviennent dans ce domaine peuvent chevaucher le droit à la non-discrimination et les droits à la protection des données.

Résumé – Utilisation de l’IA dans l’administration publique portugaise et recommanda- tions politiques sur la discri- mination algorithmique



Le résumé est basé sur le rapport d'analyse des lacunes et des données en l'état (« As-is and gap analysis »), dont la partie portugaise a été rédigée par Eduardo dos Santos.³⁷

L'utilisation de l'intelligence artificielle dans l'administration publique

Au Portugal, l'utilisation de l'intelligence artificielle (IA) en est encore à ses débuts, mais se développe progressivement sous l'effet d'initiatives pilotes et des démarches de numérisation. Par exemple, le portail national des marchés publics répertorie 119 contrats liés à l'IA depuis 2009, dont près de la moitié ont été attribués en 2023 et 2024 seulement³⁸. L'IA est utilisée dans des secteurs tels que la justice et la santé, et plusieurs services publics ont lancé des chatbots IA pour assurer l'assistance virtuelle des services.

Cadre juridique et institutionnel

Le règlement européen sur l'IA, ainsi que la législation européenne et portugaise en matière de lutte contre la discrimination et de protection des données,

37. Ceci est un rapport confidentiel qui examine l'utilisation des systèmes d'IA et de PDA dans le secteur public au Portugal, évalue les cadres juridiques et politiques liés à l'IA du point de vue de l'égalité et de la non-discrimination, et formule des recommandations à l'intention des autorités et des décideurs-euses politiques.

38. Portal Base, disponible à l'adresse : www.base.gov.pt, consulté le 10 novembre 2025.

forment un cadre juridique pour lutter contre la discrimination algorithmique. Outre l'acquis de l'UE, d'autres possibilités de prévention et de lutte contre la discrimination algorithmique sont prévues dans la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit, ouverte à la signature en septembre 2024, dans laquelle l'article 10 sur l'égalité et la non-discrimination interdit la discrimination et appelle les États à « maintenir des mesures visant à supprimer les inégalités [...] afin d'obtenir des résultats impartiaux, justes et équitables ».

Au Portugal, une affaire judiciaire notable impliquant une prise de décision algorithmique a eu lieu en 2021, lorsque la compagnie aérienne public-privé TAP a utilisé un algorithme pour sélectionner jusqu'à 500 employés à licencier en fonction de paramètres tels que la productivité, l'absentéisme, l'expérience, la contribution, le coût et les qualifications. Les travailleurs ont critiqué le manque de transparence, car la pondération de chaque critère n'a pas été divulguée. En 2023, le tribunal du travail de Lisbonne a jugé illégal le licenciement collectif concernant 23 salariés et salariées.

Le gouvernement portugais a adopté plusieurs instruments stratégiques pour encourager la numérisation et l'utilisation responsable de l'IA, tels que la Stratégie nationale sur l'intelligence artificielle³⁹ publiée en 2019 dans le cadre de l'Initiative nationale sur les compétences numériques (INCoDe.2030) et la Stratégie nationale sur le numérique⁴⁰ à l'horizon 2030.

Toutefois, le rapport d'analyse des lacunes et de l'état actuel note que la prolifération des stratégies, des plans et des programmes au fil des ans peut être contre-productive si elle n'est pas correctement coordonnée et intégrée.

En 2024, le Portugal a créé le Conseil pour le numérique dans l'administration publique avec un groupe de travail technique dédié à l'intelligence artificielle chargé de surveiller la mise en œuvre du règlement européen sur l'IA. Un autre rôle central est joué par l'ancienne Agence pour la modernisation administrative, restructurée en Agence pour la réforme technologique de l'État en 2025, qui veille à l'utilisation responsable des systèmes numériques et d'IA dans l'administration publique. Les entités publiques qui ont l'intention d'acquérir des biens ou des services dans ce domaine, y compris des systèmes d'IA, doivent en informer l'agence et peuvent faire l'objet d'une évaluation. L'agence a élaboré un guide sur l'IA dans l'administration publique et un outil d'évaluation des risques, qui aborde certaines préoccupations éthiques, d'égalité et d'inclusivité.

39 Estratègia Nacional de Inteligência Artificial, disponible en anglais à l'adresse suivante : www.incode2030.gov.pt/en/estrategia-nacional-de-inteligencia-artificial-en, consulté le 10 novembre 2025.

40 Estratègia Digital Nacional, disponible en anglais à l'adresse suivante : <https://digital.gov.pt/documentos/portugal-digital-strategy>, consulté le 10 novembre 2025.

Lacunes dans les cadres juridiques et politiques

Transparence limitée et absence de registre national d'IA et de prise de décision automatisée

Il n'existe aucun registre public complet recensant les systèmes d'IA ou de prise de décision automatisée (PDA) utilisés dans l'administration publique portugaise. Actuellement, aucun mécanisme public ne permet aux citoyens ou aux organismes de surveillance d'identifier quels systèmes d'IA ou de PDA sont utilisés par les entités publiques. Ce manque de transparence peut rendre difficile l'évaluation des risques de discrimination et des préjudices potentiels de ces systèmes.

Exclusion des organismes chargés des droits fondamentaux et de l'égalité du cadre de gouvernance du règlement européen sur l'IA

La liste des autorités désignées au Portugal en vertu de l'article 77 du règlement européen sur l'IA comprend 14 entités au total : les inspections générales des finances, de la défense, de la justice, de l'administration interne et de l'éducation, ainsi que les régulateurs sectoriels tels que l'Autorité nationale des communications, l'Autorité de sécurité alimentaire et économique et l'Autorité des conditions de travail. Ni aucune institution de défense des droits fondamentaux, ni aucun organisme de promotion de l'égalité ni le Médiateur ne sont inclus en tant qu'entité « protégeant les droits fondamentaux, y compris le droit à la non-discrimination » au sens de l'article 77. L'absence de ces organismes soulève des préoccupations quant à la manière dont les cas potentiels de discrimination algorithmique seront traités au titre du de l'UE et du droit national, et à la manière dont les personnes victimes de discrimination pourront systématiquement demander réparation en vertu du cadre de gouvernance du règlement européen sur l'IA.

Mécanismes de surveillance et de coordination fragmentés

Les responsabilités en matière de gouvernance de l'IA sont dispersées entre plusieurs autorités et il n'existe aucun canal de communication structuré entre ces entités et les institutions dotées de mandats en matière de droits humains ou d'égalité, comme le Médiateur ou les organismes de promotion de l'égalité tels que la Commission pour la citoyenneté et l'égalité des genres (CIG). En l'absence de mécanisme de coopération, il est difficile d'identifier et de remédier aux chevauchements ou aux lacunes potentiels dans la surveillance, en particulier en ce qui concerne les droits fondamentaux et les obligations en matière de non-discrimination.

Sensibilisation, ressources et capacités techniques limitées

Au Portugal, les considérations relatives à l'égalité et aux droits humains ne sont pas encore systématiquement intégrées dans les processus de conception d'IA ou de passation de marchés et dans l'utilisation de l'IA dans l'administration publique. Les organismes de promotion de l'égalité et les autres institutions compétentes disposent de moyens limités pour mobiliser une expertise technique suffisante en matière d'IA et ne disposent pas des ressources nécessaires pour procéder à des évaluations techniques des systèmes d'IA afin d'aider les personnes susceptibles

d'être impactées par la discrimination. La Commission pour la citoyenneté et l'égalité des genres et d'autres organismes compétents sont confrontés à des contraintes de ressources qui limitent leur capacité à faire face aux écueils émergents, ainsi qu'à des pouvoirs limités (c'est-à-dire l'absence de mandats d'enquête et de plaider), qui limitent leur capacité à traiter les questions de discrimination algorithmique.

Recommandations clés

- ▶ **Mettre en place un registre public des systèmes d'IA et de PDA utilisés dans l'administration publique :** L'introduction d'une obligation pour les autorités publiques de déclarer et de documenter les systèmes d'IA dans un registre centralisé améliorerait la transparence, favoriserait le contrôle par les organismes de surveillance et d'égalité et renforcerait la confiance du public.
- ▶ **Consolider les pouvoirs d'enquête et d'exécution des organismes chargés de l'égalité et inclure les organismes chargés des droits fondamentaux et de l'égalité en tant qu'organismes nationaux relevant de l'article 77 :** Dans le contexte de la transposition des directives de l'UE sur les normes applicables aux organismes de promotion de l'égalité⁴¹, cela renforcerait le rôle de ces institutions pour permettre un suivi proactif, un engagement au niveau de la gouvernance de l'IA et une application effective des obligations anti-discrimination en vertu du droit européen et national.
- ▶ **Jusqu'à ce que les organismes chargés des droits fondamentaux et de l'égalité soient désignés en tant qu'organes relevant de l'article 77, clarifier les mandats institutionnels et renforcer la coordination inter-institutions entre les autorités :** La définition de responsabilités claires et de mécanismes formels de coopération entre les autorités chargées de l'égalité, de la protection des données et de la protection des consommateurs garantirait un échange cohérent d'informations et la mise en œuvre du règlement européen sur l'IA.
- ▶ **Créer un cadre harmonisé pour évaluer les incidences sur l'égalité, l'égalité des sexes et les droits fondamentaux, en s'appuyant sur les stratégies nationales liées à l'IA, le guide de l'Agence pour la modernisation administrative et le règlement européen sur l'IA :** Ce cadre élaborerait une approche nationale normalisée pour guider les autorités publiques dans l'évaluation des risques de discrimination algorithmique et assurer une application cohérente des garanties dans le secteur public.
- ▶ **Mettre en place des programmes de formation et de renforcement des capacités en matière de droits fondamentaux pour les fonctionnaires et les autorités compétentes en vertu du règlement européen sur l'IA :** Des mesures ciblées de formation et de renforcement des capacités devraient être mises en œuvre pour que les autorités puissent identifier et combattre la discrimination algorithmique, notamment la discrimination fondée sur le sexe.
- ▶ **Intégrer les principes de lutte contre la discrimination et d'égalité des sexes de manière plus visible dans les marchés publics liés à des systèmes d'IA et de PDA :** Cela comprend la poursuite de la surveillance des systèmes d'IA par le biais d'audits réguliers pendant leur déploiement et leur utilisation.

41. Directive UE 2024/1499 et directive UE 2024/1500.

- ▶ **Promouvoir des campagnes de sensibilisation et de maîtrise de l'IA pour le grand public :** Des campagnes de sensibilisation et du matériel éducatif pour informer le public sur la façon dont l'IA est utilisée dans l'administration publique, ses impacts potentiels et les mécanismes de recours disponibles doivent être mises au point. Renforcer la maîtrise de l'IA permettrait d'améliorer la transparence et la responsabilisation dans l'utilisation de ces systèmes.
- ▶ **Assurer des financements adéquats et une expertise technique en IA pour les organismes de promotion de l'égalité :** Les organismes de promotion de l'égalité doivent disposer de financements suffisants et d'une expertise technique spécialisée en IA pour lutter contre la discrimination algorithmique.
- ▶ **Réfléchir à la manière dont l'initiative amorcée par le gouvernement pour développer une nouvelle génération de services publics omnicanaux et centrés sur la personne dans le cadre de la plateforme gov.pt permettrait aux citoyens de déposer des plaintes et d'être orientés en ce qui concerne les problématiques de discrimination induite par l'IA.**

Références

Access Now (2024), Joint statement – A dangerous precedent: how the EU AI Act fails migrants and people on the move, disponible en anglais sur : www.accessnow.org/press-release/joint-statement-ai-act-fails-migrants-and-people-on-the-move, consulté le 10 novembre 2025.

Algorithm Audit (2024a), Preventing prejudice. Recommendations for risk profiling in the College Grant Control process: a quantitative and qualitative analysis, disponible en anglais sur : https://algorithmaudit.eu/algoprudence/cases/aa202401_preventing-prejudice/, consulté le 10 novembre 2025.

Algorithm audit (2024b), DUO control process bias Towards Students with an non European migration background, disponible à l'adresse : https://algorithmaudit.eu/events/press_room/#DUO_CBS, consulté le 15 novembre 2025.

Algorithm Watch (2018), Finnish credit score ruling raises questions about discrimination and how to avoid it, disponible en anglais à l'adresse : [Finnish Credit Score Ruling raises Questions about Discrimination and how to avoid it - AlgorithmWatch](#), consulté le 15 novembre 2025.

Allhutter D. et al. (2020), « Algorithmic profiling of job seekers in Austria: how austerity politics are made effective », *Frontiers in Big Data*, Vol. 3, No. 5.

Amnesty International (2021), Dutch Childcare Benefit Scandal an urgent Wake-up Call to ban racist Algorithms, disponible en anglais à l'adresse : www.amnesty.org/en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal, consulté le 10 novembre 2025.

Conseil d'État (2024), Requête introductive d'instance, 15 octobre, disponible à l'adresse : www.laquadrature.net/wp-content/uploads/sites/8/2024/10/LQDN_Cnaf_Memoire_Introductif_anon.pdf, consulté le 10 novembre 2025.

Conseil de l'Europe (2007), Convention 201 sur la protection des enfants contre l'exploitation et les abus sexuels (Lanzarote, 25 octobre 2007), disponible à l'adresse : <https://rm.coe.int/1680084833>, consulté le 10 novembre 2025.

Conseil de l'Europe (2010), Recommandation cm/Rec(2010)5 du Comité des Ministres sur des mesures visant à combattre la discrimination fondée sur l'orientation sexuelle ou l'identité de genre, disponible à l'adresse suivante : [https://search.coe.int/cm/#%22CoEIdentifier%22:\[%2209000016805b1652%22\],%22sort%22:\[%22CoEValidationDate%20Descending%22\]](https://search.coe.int/cm/#%22CoEIdentifier%22:[%2209000016805b1652%22],%22sort%22:[%22CoEValidationDate%20Descending%22]), consulté le 10 novembre 2025.

Conseil de l'Europe (2013), Déclaration du Comité des Ministres sur les risques présentés par le suivi numérique et les autres technologies de surveillance pour les droits fondamentaux, disponible à l'adresse suivante : [https://search.coe.int/cm/#%22CoEIdentifier%22:\[%2209125948801e0622%22\],%22sort%22:\[%22CoEValidationDate%20Descending%22\]](https://search.coe.int/cm/#%22CoEIdentifier%22:[%2209125948801e0622%22],%22sort%22:[%22CoEValidationDate%20Descending%22]), consulté le 10 novembre 2025.

Conseil de l'Europe (2018), Convention 108+ pour la protection des personnes à l'égard du traitement des données à caractère personnel, disponible à l'adresse suivante : <https://rm.coe.int/convention-108-convention-pour-la-protection-des-personnes-a-l-egard-d/16808b3726>, consulté le 10 novembre 2025.

Conseil de l'Europe (2019a), Déclaration du Comité des Ministres sur les capacités de manipulation des processus algorithmiques, Decl(13/02/2019)1, disponible à l'adresse suivante : [https://search.coe.int/cm/#{%22CoEIdentifier%22:\[%220912594880233b7e%22\],%22sort%22:\[%22CoEValidationDate%20Descending%22\]}](https://search.coe.int/cm/#{%22CoEIdentifier%22:[%220912594880233b7e%22],%22sort%22:[%22CoEValidationDate%20Descending%22]}), consulté le 10 novembre 2025.

Conseil de l'Europe (2019b), Recommandation cm/Rec(2019)1 du Comité des Ministres aux États membres sur la prévention et la lutte contre le sexisme, disponible à l'adresse suivante : <https://rm.coe.int/168093b269>, consulté le 10 novembre 2025.

Conseil de l'Europe (2019c), « Lignes directrices sur l'intelligence artificielle et la protection des données », Comité consultatif de la Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel (Convention 108), disponible à l'adresse suivante : <https://rm.coe.int/2018-lignes-directrices-sur-l-intelligence-artificielle-et-la-protecti/168098e1b8>, consulté le 10 novembre 2025.

Conseil de l'Europe (2020), Recommandation cm/Rec(2020)1 du Comité des Ministres aux États membres sur les impacts des systèmes algorithmiques sur les droits de l'homme, disponible à l'adresse suivante : [https://search.coe.int/cm/#{%22CoEIdentifier%22:\[%2209125948802662bf%22\],%22sort%22:\[%22CoEValidationDate%20Descending%22\]}](https://search.coe.int/cm/#{%22CoEIdentifier%22:[%2209125948802662bf%22],%22sort%22:[%22CoEValidationDate%20Descending%22]}), consulté le 10 novembre 2025.

Conseil de l'Europe (2021a), Recommandation générale n° 1 sur la dimension numérique de la violence à l'égard des femmes, Groupe d'experts sur la lutte contre la violence à l'égard des femmes et la violence domestique (GREVIO), disponible à l'adresse suivante : <https://edoc.coe.int/fr/violence-l-gard-des-femmes/10639-recommandation-generale-n-1-du-grevio-sur-la-dimension-numerique-de-la-violence-a-legard-des-femmes.html>, consulté le 10 novembre 2025.

Conseil de l'Europe (2021b), « Modération de contenu : Meilleures pratiques en vue de la mise en place de cadres juridiques et procéduraux efficaces pour les mécanismes d'autorégulation et de corégulation de la modération de contenu », note d'orientation, Comité directeur sur les médias et la société de l'information (CDMSI), disponible à l'adresse suivante : <https://rm.coe.int/content-moderation-fr/1680a40294>, consulté le 10 novembre 2025.

Conseil de l'Europe (2021c), « Lignes directrices sur la défense de l'égalité et la protection contre la discrimination et la haine pendant la pandémie de COVID-19 et les crises similaires à venir », Comité des Ministres, CM(2021)37-add1 rev, disponible à l'adresse suivante : <https://edoc.coe.int/fr/vivre-ensemble-diversite-et-liberte-en-europe/9733-lignes-directrices-du-comite-des-ministres-du-conseil-de-leurope-sur-la-defense-de-legalite-et-la-protection-contre-la-discrimination-et-la-haine-pendant-la-pandemie-de-covid-19-et-dautres-crisis-similaires-a-venir.html>, consulté le 10 novembre 2025.

Conseil de l'Europe (2021d), « Lignes directrices sur la reconnaissance faciale », Comité consultatif de la Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel (Convention 108), disponible à l'adresse suivante : <https://edoc.coe.int/fr/intelligence-artificielle/9749-lignes-directrices-sur-la-reconnaissance-faciale.html>, consulté le 10 novembre 2025.

Conseil de l'Europe (2022a), « Contours de la méthodologie pour l'évaluation des risques et de l'impact Huderia », Comité sur l'intelligence artificielle, disponible à l'adresse suivante : <https://rm.coe.int/cai-bu-2022-03-fr-contours-de-la-methodologie-huderia/1680a81e15>, consulté le 10 novembre 2025.

Conseil de l'Europe (2022b), Recommandation CM/Rec(2022)13 du Comité des Ministres aux États membres sur les effets des technologies numériques sur la liberté d'expression, disponible à l'adresse suivante : [https://search.coe.int/cm/#{%22CoEIdentifier%22:\[%2209125948802680f8%22\],%22sort%22:\[%22CoEValidationDate%20Descending%22\]}](https://search.coe.int/cm/#{%22CoEIdentifier%22:[%2209125948802680f8%22],%22sort%22:[%22CoEValidationDate%20Descending%22]}), consulté le 10 novembre 2025.

Conseil de l'Europe (2022c), Recommandation CM/Rec(2022)17 du Comité des Ministres aux États membres sur la protection des droits des femmes et des filles migrantes, réfugiées et demandeuses d'asile, disponible à l'adresse suivante : <https://edoc.coe.int/fr/droit-international/11093-la-protection-des-droits-des-femmes-et-des-filles-migrantes-refugiees-et-demandeuses-d-asile-recommandation-cmrec202217.html>, consulté le 10 novembre 2025.

Conseil de l'Europe (2022d), Recommandation CM/Rec(2022)16 du Comité des Ministres sur la lutte contre les discours de haine, disponible à l'adresse suivante : <https://search.coe.int/cm?i=0912594880268339>, consulté le 10 novembre 2025.

Conseil de l'Europe, Recommandation CM/Rec(2024)4 du Comité des Ministres sur la lutte contre les crimes de haine, disponible à l'adresse suivante : <https://search.coe.int/cm?i=091259488025f97d>, consulté le 10 novembre 2025.

Conseil de l'Europe (2024a), Convention-cadre sur l'intelligence artificielle et les droits de l'Homme, la démocratie et l'État de droit, Vilnius.

Conseil de l'Europe (2024b), « Méthodologie pour l'évaluation des risques et des impacts des systèmes d'intelligence artificielle du point de vue des droits humains, de la démocratie et de l'État de droit (méthodologie Huderia) », Comité sur l'intelligence artificielle, disponible à l'adresse suivante : <https://rm.coe.int/cai-2024-16rev2-fr-methodologie-pour-l-evaluation-des-risques-et-des-i/1680b2a09e> consulté le 10 novembre 2025.

Conseil de l'Europe (2024c), Stratégie pour l'égalité entre les hommes et les femmes 2024-2029, disponible à l'adresse suivante : <https://rm.coe.int/prems-080324-fra-2573-gender-equality-strategy-2024-29-txt-web-a5-1-27/1680afc66c>, consulté le 10 novembre 2025.

Conseil de l'Europe (2024d), Recommandation CM/Rec(2024)1 du Comité des Ministres aux États membres sur l'égalité pour les femmes et des filles roms et de la communauté des Gens du voyage, disponible à l'adresse suivante : <https://search.coe.int/cm?i=091259488025f13c>, consulté le 10 novembre 2025.

Conseil de l'Europe (à paraître), projet de Recommandation du Comité des Ministres aux États membres sur l'impact des systèmes d'intelligence artificielle, leur potentiel de promotion de l'égalité, y compris l'égalité de genre, et les risques qu'ils peuvent entraîner en matière de non-discrimination, Comité d'expert-es sur l'intelligence artificielle, l'égalité et la discrimination (GEC/ADI-AI).

Conseil de l'Europe, Commissaire aux droits de l'homme (2019), « Décoder l'intelligence artificielle : 10 mesures pour protéger les droits de l'homme », disponible à l'adresse suivante : <https://book.coe.int/fr/commissaire-aux-droits-de-l-homme/9668-decoder-lintelligence-artificielle-10-mesures-pour-protoger-les-droits-de-lhomme.html>, consulté le 10 novembre 2025.

Conseil de l'Europe, Commissaire aux droits de l'homme (2023), « Les droits de l'homme dès la conception de l'IA : une protection durable des droits humains à l'ère de l'intelligence artificielle », Recommandation de suivi, disponible à l'adresse suivante : <https://rm.coe.int/-les-droits-humains-des-la-conception-de-l-ia-une-protection-durable-d/1680ac78b3>, consulté le 10 novembre.

Conseil de l'Europe, Assemblée parlementaire (2017), Recommandation 2102 (2017) sur « La convergence technologique, l'intelligence artificielle et les droits de l'homme », disponible à l'adresse suivante : <https://pace.coe.int/en/files/23726>, consulté le 10 novembre 2025.

Conseil de l'Europe, Assemblée parlementaire (2020a), résolution 2343 sur la « prévention de la discrimination causée par l'utilisation de l'intelligence artificielle », disponible à l'adresse suivante : <https://pace.coe.int/fr/files/23726>, consulté le 10 novembre 2025.

Conseil de l'Europe, Assemblée parlementaire (2020b), Recommandation 2183 (2020) pour « Prévenir la discrimination résultant de l'utilisation de l'intelligence artificielle », disponible à l'adresse suivante : <https://pace.coe.int/fr/files/28809>, consulté le 10 novembre 2025.

Cour de justice de l'Union européenne (2008), C-54/07 Centrum voor gelijkheid van kansen en voor racismebestrijding contre Firma Feryn NV EU:C:2008:397.

Cour de justice de l'Union européenne (2022), C-817/19 Ligue des droits humains ASBL c. Conseil des ministres EU:C:2022:491.

Cour de justice de l'Union européenne (2023), C-634/21 SCHUFA Holding (Scoring) EU:C:2023:957.

Cour de justice de l'Union européenne (2024), C-588/21 P Public.Resource.Org and Right to Know c. Commission et autres EU:C:2024:201.

Cour de justice de l'Union européenne (2025), C 203/22 Dun & Bradstreet Austria EU:C:2025:117.

Défenseur des droits (2024), « Algorithmes, systèmes d'IA et services publics : quels droits pour les usagers ? », disponible à l'adresse suivante : https://www.defenseurdes-droits.fr/sites/default/files/2024-11/DDD_rapport_algorithmes-systemes-d-ia-et-services-publics_2024_20241025.pdf, consulté le 10 novembre 2025.

Dubois V. (2021), *Contrôler les assistés : Genèses et usages d'un mot d'ordre*, Raisons d'agir, Paris.

Commission européenne (2020), « Livre blanc : Sur l'intelligence artificielle – Une approche européenne axée sur l'excellence et la confiance », COM(2020)65 final.

Commission européenne (2022a), « Proposition de directive du Parlement européen et du Conseil relative à l'adaptation des règles en matière de responsabilité civile extracontractuelle au domaine de l'intelligence artificielle (directive sur la responsabilité en matière d'IA) », COM/2022/496 final.

Commission européenne (2022b), « Déclaration européenne sur les droits et principes numériques pour la décennie numérique », COM(2022) 28 final.

Commission européenne (2025a), « Lignes directrices de la Commission sur les pratiques interdites en matière d'intelligence artificielle au sens du règlement (UE) 2024/1689 (règlement sur l'IA) », disponible à l'adresse suivante : <https://digital-strategy.ec.europa.eu/fr/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act>, consulté le 10 novembre 2025.

Commission européenne (2025b), Code de pratique général de l'IA à usage général, disponible à l'adresse suivante : <https://digital-strategy.ec.europa.eu/fr/policies/contents-code-gpai>, consulté le 10 novembre 2025.

Commission européenne (2025c), Comité de l'IA, <https://digital-strategy.ec.europa.eu/fr/policies/ai-board>, consulté le 10 novembre 2025.

Commission européenne (2025d), « Lignes directrices de la Commission sur la définition d'un système d'intelligence artificielle au sens du règlement (UE) 2024/1689 (règlement sur l'IA) », C(2025) 924 final.

Commission européenne contre le racisme et l'intolérance (2000), « Recommandation n° 6 sur la lutte contre la diffusion de matériels racistes, xénophobes et antisémites par l'internet », disponible à l'adresse suivante : <https://rm.coe.int/recommandation-de-politique-generale-n-6-de-l-ecri-sur-la-lutte-contre/16808b5a8e>, consulté le 10 novembre 2025.

Commission européenne contre le racisme et l'intolérance (2004), « Recommandation de politique générale n° 8 de l'ECRI pour lutter contre le racisme tout en combattant le terrorisme », disponible à l'adresse suivante : <https://rm.coe.int/recommandation-de-politique-generale-n-8-de-l-ecri-pour-lutter-contre-/16808b5abe>, consulté le 10 novembre 2025.

Commission européenne contre le racisme et l'intolérance (2006), « Recommandation de politique générale n° 10 de l'ECRI pour lutter contre le racisme et la discrimination raciale dans et à travers l'éducation scolaire », disponible à l'adresse suivante : <https://rm.coe.int/recommandation-de-politique-generale-n-10-de-l-ecri-pour-lutter-contre/16808b5ad6>, consulté le 10 novembre 2025.

Commission européenne contre le racisme et l'intolérance (2007), « Recommandation de politique générale n° 11 de l'ECRI sur la lutte contre le racisme et la discrimination raciale dans les activités de la police », disponible à l'adresse suivante : <https://rm.coe.int/recommandation-de-politique-generale-n-11-de-l-ecri-sur-la-lutte-contr/16808b5ae0>, consulté le 10 novembre 2025.

Commission européenne contre le racisme et l'intolérance (2012), « Recommandation de politique générale n° 14 de l'ECRI sur la lutte contre le racisme et la discrimination raciale dans le monde du travail », disponible à l'adresse suivante : <https://rm.coe.int/recommandation-de-politique-generale-n-14-de-l-ecri-sur-la-lutte-contr/16808b5afd>, consulté le 10 novembre 2025.

Commission européenne contre le racisme et l'intolérance (2015), « Recommandation de politique générale n° 15 de l'ECRI sur la lutte contre le discours de haine », disponible à l'adresse suivante : <https://rm.coe.int/recommandation-de-politique-generale-n-15-de-l-ecri-sur-la-lutte-contr/16808b5b03>, consulté le 10 novembre 2025.

Commission européenne contre le racisme et l'intolérance (2017), « Recommandation de politique générale n° 7 (révisée) de l'ECRI sur la législation nationale pour lutter contre le racisme et la discrimination raciale », disponible à l'adresse suivante : <https://rm.coe.int/recommandation-de-politique-generale-n-7-revisee-de-l-ecri-sur-la-legi/16808b5ab0>, consulté le 10 novembre 2025.

Cour européenne des droits de l'homme (2007), requête n° 57325/00, D.H. et autres c. République tchèque (Grande Chambre, 13 novembre 2007), point 178.

Cour européenne des droits de l'homme (2023), requête 11519/20, Glukhin c. Russie (troisième section, 2023).

European Digital Rights (2024), How to Fight Biometric mass surveillance after the ai Act : a legal and Practical guide, disponible en anglais à l'adresse suivante : <https://edri.org/our-work/how-to-fight-biometric-mass-surveillance-after-the-ai-act-a-legal-and-practical-guide>, consulté le 10 novembre 2025.

Réseau européen des migrations et Organisation de coopération et de développement économiques (2022), « The use of digitalisation and artificial intelligence in migration management », disponible à l'adresse suivante : www.oecd.org/en/topics/sub-issues/migration-policies-returns-and-attracting-talent/migration-policy-debates-and-data-briefs.html, consulté le 10 novembre 2025.

Réseau européen des institutions nationales des droits de l'homme (ENNHRI) (2024), « Technologies, migration, and human rights: the role of European NHRIs », document de cadrage de l'ENNHRI, disponible en anglais à l'adresse : <https://ennhri.org/publications-statements/#artificial-intelligence-publications>, consulté le 10 novembre 2025.

Union européenne (2016a), règlement (UE) 2016/679 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, abrogeant la directive 95/46/CE;

Union européenne (2016b), directive 2016/680 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales, et à la libre circulation de ces données.

Union européenne (2022), règlement (UE) 2022/2065 du Parlement européen et du Conseil du 19 octobre 2022 relatif à un marché unique des services numériques.

Union européenne (2024a), la directive du Conseil 2024/1499 relative aux normes applicables aux organismes pour l'égalité dans les domaines de l'égalité de traitement entre les personnes sans distinction de race ou d'origine ethnique, de l'égalité de traitement entre les personnes en matière d'emploi et de travail sans distinction de religion ou de convictions, de handicap, d'âge ou d'orientation sexuelle et de l'égalité de traitement entre les femmes et les hommes en matière de sécurité sociale ainsi que dans l'accès à des biens et services et la fourniture de biens et services;

Union européenne (2024b), directive du Conseil 2024/1500 sur les normes applicables aux organismes pour l'égalité de traitement dans le domaine de l'égalité de traitement et de l'égalité des chances entre les femmes et les hommes en matière d'emploi et de travail.

Union européenne (2024c), 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 (le « règlement européen sur l'intelligence artificielle »), disponible à l'adresse : https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401689, consulté le 10 novembre 2025

Comité de Lanzarote (2024), Déclaration sur la protection des enfants contre l'exploitation sexuelle et les abus sexuels facilités par les technologies émergentes, disponible à l'adresse suivante : <https://rm.coe.int/declaration-sur-la-protection-des-enfants-contre-l-exploitation-et-les/1680b25aa7>, consulté le 10 novembre 2025.

La Quadrature du Net (2023), Notation des allocataires : l'indécence des pratiques de la CAF désormais indéniable, disponible à l'adresse : www.laquadrature.net/2023/11/27/notation-des-allocataires-lindecence-des-pratiques-de-la-caf-desormais-indeniable, consulté le 10 novembre 2025.

La Quadrature du Net (2024), À France Travail, l'essor du contrôle algorithmique, 25 juin, disponible à l'adresse : www.laquadrature.net/2024/06/25/a-france-travail-lessor-du-controle-algorithmique, consulté le 10 novembre 2025.

Malgieri G. et Comandé G. (2017), « Why a right to legibility of automated decision-making exists in the General Data Protection Regulation », *International Data Privacy Law*, Vol. 7, Numéro 4, p. 243-65.

McGregor L. et Molnar P. (2023), « Digital Border Governance : a Human Rights Based Approach », Université d'Essex et Haut Commissariat des Nations Unies aux droits de l'homme.

Milano S. et Prunkl C. (2024), « Algorithmic profiling as a source of hermeneutical injustice », *Philosophical Studies*, vol. 182, p. 185-203.

Mittelstadt B. (2025), « How to use the Artificial Intelligence Act to investigate AI bias and discrimination: a guide for Equality Bodies », *European Network of Equality Bodies (Equinet)*, Bruxelles.

National Non-Discrimination and Equality Tribunal (2017), Décision 216/2017, disponible en anglais à l'adresse : www.yvtltk.fi/material/sites/ytaltk/tapausselosteet/zedi26zya/Decision_YVTltk_xx16_2017_Multiple_discrimination_in_the_assessment_of_creditworthiness.pdf, consulté le 10 novembre 2025.

Selbts A. D. et Powles J. (2017) « Meaning information and the right to explication », *International Data Privacy Law*, vol. 7, n° 4, p. 233-42.

Rapporteuse spéciale des Nations Unies (2021), Discrimination raciale et xénophobe et utilisation des technologies numériques dans le contrôle des frontières et de l'immigration - Rapport de la Rapporteuse spéciale sur les formes contemporaines de racisme, de discrimination raciale, de xénophobie et de l'intolérance qui y est associée, Achiume E. Tendayi, Rapporteuse spéciale des Nations Unies sur les formes contemporaines de racisme, de discrimination raciale, de xénophobie et de l'intolérance qui y est associée, A/HRC/48/76.

Wachter S., Mittelstadt B. et Floridi L. (2017), « Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation », *International Data Privacy Law*, vol. 7, numéro 2, pp. 76-99.

Xenidis R. (2021), « Tuning eu Equality law to Algorithmic discrimination : Three Pathways to Resilience », *Maastricht Journal of European and comparative Law*, vol. 27, numéro 6, p. 736-58.

L'intégration croissante de l'intelligence artificielle (IA) dans diverses applications dans un large éventail de secteurs pose des défis importants pour la protection des droits fondamentaux. Parmi ceux-ci, la discrimination algorithmique apparaît comme une préoccupation particulièrement urgente. Des recherches empiriques ont démontré que les biais algorithmiques non seulement reflètent, mais exacerbent également les inégalités sociales existantes à grande échelle, en particulier dans les domaines où la prise de décision algorithmique est courante, tels que la police, l'emploi, l'éducation et les assurances. En 2024, l'Union européenne et le Conseil de l'Europe ont adopté des instruments juridiques historiques : le règlement européen sur l'IA (règlement 2024/1689) et la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit. Le présent rapport examine comment ces nouveaux instruments juridiques contribuent à renforcer la protection contre la discrimination algorithmique en Europe et évalue les lacunes qui continuent d'affecter ces cadres juridiques. Le rapport vise également à retracer et à synthétiser ces derniers développements afin de doter les acteurs concernés, tels que les organismes de promotion de l'égalité, d'outils leur permettant de s'adapter à un paysage juridique en mutation et, le cas échéant, d'intervenir dans les évolutions en cours.



YHDENVERTAISUUSVALTUUTETTU
NON-DISCRIMINATION OMBUDSMAN



CIG
COMISSÃO PARA A CIDADANIA
E A IGUALDADE DE GÉNERO

PREMS 163825

FRA

Les Etats membres de l'Union européenne ont décidé de mettre en commun leur savoir-faire, leurs ressources et leur destin. Ensemble, ils ont construit une zone de stabilité, de démocratie et de développement durable tout en maintenant leur diversité culturelle, la tolérance et les libertés individuelles. L'Union européenne s'engage à partager ses réalisations et ses valeurs avec les pays et les peuples au-delà de ses frontières.

www.europa.eu

Le Conseil de l'Europe est la principale organisation de défense des droits humains du continent. Il comprend 46 Etats membres, dont l'ensemble des membres de l'Union européenne. Tous les Etats membres du Conseil de l'Europe ont signé la Convention européenne des droits de l'homme, un traité visant à protéger les droits humains, la démocratie et l'Etat de droit. La Cour européenne des droits de l'homme contrôle la mise en œuvre de la Convention dans les Etats membres.

www.coe.int



UNION EUROPÉENNE

COUNCIL OF EUROPE



CONSEIL DE L'EUROPE