LIGNES DIRECTRICES POLITIQUES EUROPÉENNES RELATIVES AUX DISCRIMINATIONS INDUITES PAR L'IA ET LES ALGORITHMES

à l'intention des organismes de promotion de l'égalité et des autres structures nationales des droits humains

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

Kris Shrishak
Soizic Pénicaud

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

LIGNES DIRECTRICES POLITIQUES EUROPÉENNES RELATIVES AUX DISCRIMINATIONS INDUITES PAR L'IA ET LES ALGORITHMES

à l'intention des organismes de promotion de l'égalité et des autres structures nationales des droits humains

Kris Shrishak
Soizic Pénicaud

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 ses auteur.es. 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.

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

AFAR	Algorithmic fairness for asylum seekers and refugees (équité des algorithmes pour les demandeurs d'asile et les réfugiés)
AIDF	Analyse d'impact sur les droits fondamentaux
AIPD	Analyse d'impact sur la protection des données
ASM	Autorité de surveillance du marché
CJUE	Cour de justice de l'Union européenne
CNIL	Commission nationale de l'informatique et des libertés
Commission	Commission européenne
Convention 108+	Convention modernisée du Conseil de l'Europe n° 108 pour la protection des personnes à l'égard du traitement des données à caractère personnel
Convention-cadre	Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit (CETS N° 225)
Directive SMA	Directive sur les services de médias audiovisuels (2010/13/EU)
ECRI	Commission européenne contre le racisme et l'intolérance
EEE	Espace économique européen
Equinet	Réseau européen d'organismes de promotion de l'égalité
FARI	Institut consacré à l'IA et aux questions d'intérêt général
HUDERIA	Méthodologie d'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
IA	Intelligence artificielle
OCR	Reconnaissance optique de caractères
OPE	Organismes de promotion de l'égalité
OSC	Organisations de la société civile
PDA	Prise de décision automatisée

Abréviations

PEReN	Pôle d'expertise de la régulation numérique
Règlement européen sur l'IA	Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 fixant des règles harmonisées en matière d'intelligence artificielle
Règlement sur les dispositifs médicaux	Règlement (UE) 2017/745 du Parlement européen et du Conseil du 5 avril 2017 relatif aux dispositifs médicaux
RGPD	Règlement général sur la protection des données (règlement (UE) 2016/679)
SIBD	Systèmes d'identification biométrique à distance
SNDH	Structure nationale des droits humains
STIM	Sciences, technologie, ingénierie et mathématiques
UE	Union européenne
Unia	Centre interfédéral pour l'égalité des chances belge

Pour le glossaire, se reporter aux définitions établies à l'article 3 du règlement européen sur l'IA

Remerciements

Ce rapport a été élaboré 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. Ce projet est mis en œuvre par le Conseil de l'Europe en coopération avec le Commission européenne, le Centre pour l'égalité des chances et la lutte contre le racisme (Unia, Belgique), le Médiateur de la non-discrimination (Finlande) et la Commission pour la citoyenneté et l'égalité de genre (Portugal).

Ce rapport a été rédigé par Kris Shrishak (consultant international, Conseil de l'Europe) et Soizic Pénicaud (consultante internationale, Conseil de l'Europe).

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 et Nadine Brauns (Centre interfédéral pour l'égalité des chances, Unia, Belgique); Tiina Valonen et Ville Rantala (Médiateur de la non-discrimination, YVV, Finlande); Carla Peixe, Ana Martinho Fernandes, Alexandra Andrade et Susana Miguel (Commission pour la citoyenneté et l'égalité de genre, 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).

Nous tenons à remercier tout particulièrement Louise Hooper (consultante internationale, Conseil de l'Europe) pour ses contributions expertes et sa relecture de l'avant-projet de ce rapport, ainsi que Milla Vidina (Equinet, réseau européen d'organismes de promotion de l'égalité) pour ses retours constructifs.

Ce rapport a également bénéficié des apports d'experts nationaux issus d'organismes de promotion de l'égalité et d'organisations de la société civile qui ont participé à des entretiens au printemps 2025. Leurs perspectives ont permis d'enrichir les parties analytiques et opérationnelles de ces lignes directrices.

Note de synthèse

Les lignes directrices européennes sur l'IA et la discrimination fondée sur des algorithmes définissent la manière dont les organismes de promotion de l'égalité et, le cas échéant, d'autres structures nationales de défense des droits humains (SNDH) peuvent utiliser leurs mandats dans le cadre juridique européen, en particulier le règlement (UE) 2024/1689 établissant des règles harmonisées concernant l'intelligence artificielle (Règlement européen sur l'IA) – pour protéger les droits fondamentaux et lutter contre les risques de discrimination dans le déploiement de l'IA et des systèmes de prise de décision automatisée (PDA), en particulier dans le secteur public. À l'heure où ces systèmes sont de plus en plus déployés dans les services publics et privés, ces institutions jouent un rôle essentiel dans la protection des droits fondamentaux et la lutte contre les risques de discrimination.

Ces lignes directrices sont organisées en deux sections principales. La première section se concentre sur les dispositions clés du Règlement européen sur l'IA qui sont les plus pertinentes pour les organismes de promotion de l'égalité et les SNDH, et explique comment celles-ci peuvent être utilisées dans la pratique :

1. Systèmes d'IA interdits : les lignes directrices détaillent les interdictions explicites prévues à l'article 5 pour les systèmes d'IA considérés comme incompatibles avec les valeurs et les droits fondamentaux de l'Union, y compris le droit à la non-discrimination. Ces interdictions couvrent, entre autres et dans certaines conditions, les systèmes d'IA qui manipulent, trompent, exploitent les vulnérabilités, les systèmes utilisés pour l'évaluation sociale ou l'évaluation des risques criminels, le scraping d'images pour créer ou développer des bases de données de reconnaissance faciale, ainsi que les systèmes de reconnaissance des émotions, de catégorisation biométrique et d'identification biométrique à distance en temps réel.
2. Systèmes d'IA à haut risque : les lignes directrices précisent comment les systèmes d'IA utilisés dans des domaines critiques tels que la biométrie, l'application de la loi, la protection sociale et la sécurité sociale, l'emploi, l'éducation et l'accès aux services essentiels peuvent être classés comme « à haut risque » et donc soumis à des obligations strictes en matière de gestion des risques, de gouvernance des données, de documentation, de contrôle humain et d'évaluation de l'impact sur les droits fondamentaux. Elles décrivent les éléments dont les organismes de promotion de l'égalité et les SNDH doivent tenir compte lorsqu'ils participent à des décisions de classification et lorsqu'ils examinent le respect de ces obligations du point de vue de l'égalité et de la non-discrimination.
3. Transparence et bases de données : les lignes directrices décrivent comment les nouvelles exigences en matière de transparence et les bases de données au niveau de l'UE – en particulier la base de données de l'UE pour certains systèmes d'IA à haut risque – peuvent créer des possibilités de contrôle. Les obligations d'enregistrement et de connexion peuvent aider les organismes de promotion de l'égalité et les SNDH à identifier les domaines dans lesquels les systèmes d'IA

sont utilisés à des fins de surveillance et d'enquête, et à soutenir les personnes susceptibles d'être concernées.

4. Application : les lignes directrices examinent le rôle des organismes de protection des droits fondamentaux énumérés à l'article 77 du Règlement sur l'IA, les recours dont disposent les personnes, ainsi que les mécanismes et les possibilités de coopération entre les organismes de promotion de l'égalité, les autorités chargées de la protection des données, les autorités de surveillance du marché et d'autres entités régulatrices et actrices. Elles fournissent des recommandations politiques sur la manière dont les organismes de promotion de l'égalité et les SNDH peuvent utiliser leurs pouvoirs en matière de traitement des plaintes, d'enquête, de contentieux, de conseil et de sensibilisation pour prévenir, détecter et corriger la discrimination algorithmique dans ce cadre de coopération multipartite.

Tout au long de cette première section, les rôles et responsabilités des organismes de promotion de l'égalité et des SNDH sont explicitement énoncés, avec des recommandations politiques ciblées sur la manière dont ces institutions peuvent prévenir, détecter et répondre à la discrimination algorithmique, par exemple en promouvant l'égalité, en menant des actions de sensibilisation, en soutenant les plaintes, en engageant des poursuites judiciaires et en travaillant en étroite collaboration avec les autorités nationales compétentes dans le cadre plus large du système de surveillance de l'IA.

La deuxième partie des lignes directrices repose sur deux piliers. Premièrement, elle analyse les nouvelles directives européennes relatives aux normes applicables aux organismes de promotion de l'égalité (« directives sur les normes ») et explique comment leurs dispositions en matière de mandat, d'indépendance, de ressources et de pouvoirs – y compris la promotion, l'accès à la justice et la collecte de données – peuvent être mobilisées pour lutter contre les risques de discrimination liés aux systèmes d'IA et de PDA. Ensuite, elle propose des points d'entrée thématiques pour l'utilisation de l'IA dans des secteurs souvent couverts par les mandats des organismes de promotion de l'égalité : application de la loi, migration, asile et contrôle des frontières, protection sociale et sécurité sociale, emploi et éducation. Pour chaque secteur, elle établit un lien entre les utilisations concrètes de l'IA et les dispositions pertinentes du Règlement sur l'IA (interdictions, classifications à haut risque, exigences de transparence et d'enregistrement) et souligne les domaines dans lesquels des garanties spécifiques au secteur et la participation des organismes de promotion de l'égalité et des SNDH sont essentielles.

Les lignes directrices sont conçues pour s'adapter à différents contextes nationaux et pour aider les organismes de promotion de l'égalité et les SNDH à conseiller les décideurs.euses politiques et les régulateurs.rices, à se conformer aux normes du Conseil de l'Europe telles que la Convention 108+ et la Convention-cadre sur l'intelligence artificielle et les droits humains, la démocratie et l'État de droit, et à garantir que les systèmes d'IA et de PDA sont développés et utilisés conformément à la législation européenne en matière d'égalité et de non-discrimination.

Introduction

Contexte des lignes directrices

Les administrations publiques de toute l'Europe utilisent des systèmes d'intelligence artificielle (IA) et/ou de prise de décision automatisée (PDA) dans un grand nombre de domaines politiques, notamment en ce qui concerne la migration, l'aide sociale, la justice, l'éducation, l'emploi, la fiscalité, le maintien de l'ordre ou encore la santé. Ces systèmes sont également déployés dans des domaines sensibles du secteur privé, tels que la banque (ex. : applications d'évaluation des risques-clients) et l'assurance. Bien que les systèmes d'IA et de PDA présentent des risques importants de discrimination, l'identification et l'atténuation de ces risques restent encore compliquées. Le récent rapport « Legal protection against algorithmic discrimination in Europe: Current frameworks and remaining gaps » (Xenidis, 2025), rédigé dans le cadre du projet de l'Union européenne et du Conseil de l'Europe, met en évidence plusieurs problèmes majeurs : manque de sensibilisation aux risques de discrimination, manque de transparence et manque d'informations significatives concernant l'utilisation des systèmes d'IA/PDA par les autorités publiques, difficultés d'accès à la justice et absence de pratiques de gouvernance normalisées. Les organismes de promotion de l'égalité (OPE) et les autres structures nationales des droits humains (SNDH) ont

donc un rôle clé à jouer dans la promotion du déploiement de systèmes d'IA/PDA respectant les droits fondamentaux par les organisations du secteur public.

De nouveaux cadres juridiques, notamment la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits humains, la démocratie et l'État de droit1, ainsi que le règlement de l'Union européenne sur l'IA2 visent à protéger les droits fondamentaux et à prévenir la discrimination dans les systèmes d'IA. Dans le même temps, l'Union européenne a adopté deux nouvelles directives introduisant des normes minimales pour renforcer le rôle et les capacités des organismes de promotion de l'égalité en Europe (appelées ci-après « directives sur les normes3 »). Une mise en œuvre adéquate de ces nouveaux cadres est essentielle pour garantir leur efficacité.

Objectif des lignes directrices

Dans ce contexte, les présentes lignes directrices ont pour but d'équiper les organismes de promotion de l'égalité et les autres SNDH, en particulier dans l'Union européenne, pour lutter contre la discrimination dans les systèmes d'IA/PDA en :

- les tenant informés de leurs responsabilités face à l'évolution du paysage réglementaire encadrant l'intelligence artificielle, notamment concernant son impact direct ou indirect sur leur mandat;
- leur proposant des lignes directrices spécifiques, des recommandations et des exemples de bonnes pratiques pour superviser l'application et la mise en œuvre des nouvelles réglementations, tout en les reliant à la réglementation existante;
- assumant un rôle de ressource pour les organismes de promotion de l'égalité et les SNDH afin d'aider et de conseiller les parties prenantes nationales, telles que les décideurs politiques et les régulateurs, en ce qui concerne les droits humains, l'égalité et la non-discrimination.

Ces lignes directrices sont particulièrement axées sur les systèmes d'IA/PDA dans le secteur public.

---

1. Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit (Vilnius, 5 septembre 2024), ci-après dénommée la « Convention-cadre du Conseil de l'Europe ».
2. 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 (ci-après dénommé « règlement sur l'intelligence artificielle »).
3. Directive (UE) 2024/1499 du Conseil du 7 mai 2024 relative aux normes applicables aux organismes pour l'égalité de traitement 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, et modifiant les directives 2000/43/CE et 2004/113/CE; Directive (UE) 2024/1500 du Parlement européen et du Conseil du 14 mai 2024 relative aux 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, et modifiant les directives 2006/54/CE et 2010/41/UE, ci-après dénommées les « directives sur les normes ».

Méthodologie

Ces lignes directrices donnent un aperçu des nouvelles missions, mandats et opportunités apportés par les nouvelles réglementations aux organismes de promotion de l'égalité et les SNDH, ainsi que des lignes directrices générales adaptables à chaque contexte national.

Elles s'inspirent et s'appuient sur les éléments suivants :

- Recherche documentaire et analyse du cadre juridique;
- Résultats du rapport réalisé dans le cadre du projet de l'Union européenne et du Conseil de l'Europe (Xenidis, 2025);
- Cinq entretiens semi-structurés avec des représentants d'organisations de la société civile et d'organismes de promotion de l'égalité choisis pour leur expertise en la matière.

L'UE a récemment adopté d'autres règlements dans le domaine de la gouvernance numérique. Il s'agit notamment du paquet « Législation sur les services numériques[4] », composé de la législation sur les services numériques et de la législation sur les marchés numériques[5], qui est axé sur la réglementation des services en ligne. Ces réglementations, notamment la législation sur les services numériques, abordent également la nécessité de protéger les droits fondamentaux dans la sphère numérique en termes d'identification et d'atténuation des risques. Par exemple, l'article 34 de la législation sur les services numériques, qui porte sur l'évaluation des risques par les fournisseurs de très grandes plateformes en ligne et de très grands moteurs de recherche en ligne, et l'article 35 sur l'atténuation de ces risques mentionnent tous deux les risques de discrimination. En revanche, ni la législation sur les services numériques ni la législation sur les marchés numérique ne traite des utilisations de systèmes d'IA dans le secteur public. Ainsi, sachant que les présentes lignes directrices portent essentiellement sur les utilisations de l'IA/PDA dans le secteur public, elles ne concernent pas directement la législation sur les services numériques ni celle sur les marchés numériques.

Structure des lignes directrices

La première partie des lignes directrices est structurée autour des articles du règlement européen sur l'IA qui sont les plus critiques pour les organismes de promotion de l'égalité, et les autres SNDH, soit parce qu'ils concernent directement ou indirectement leur mandat, soit parce qu'ils impliquent des changements significatifs pour d'autres institutions, régulateurs ou cadres de gouvernance. Pour chacun de ces articles, des liens sont établis avec les réglementations existantes en matière de protection des données (le règlement général sur la protection des données de

4. Règlement (UE) 2022/2065 du Parlement européen et du Conseil du 19 octobre 2022 relatif à un marché unique des services numériques et modifiant la directive 2000/31/CE, ci-après dénommé « législation sur les services numériques ».
5. Règlement (UE) 2022/1925 du Parlement européen et du Conseil du 14 septembre 2022 relatif aux marchés contestables et équitables dans le secteur numérique et modifiant les directives (UE) 2019/1937 et (UE) 2020/1828, ci-après dénommé « législation sur les marchés numériques ».

l'Union européenne[6], la directive « Police-Justice[7] » et la convention 108+ du Conseil de l'Europe pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel) et la Convention-cadre du Conseil de l'Europe.

La seconde partie dresse une vue d'ensemble des directives sur les normes qui concernent les organismes de promotion de l'égalité dans le contexte de l'égalité dans les systèmes d'IA, et propose des points d'entrée thématiques pour les secteurs du maintien de l'ordre, de la migration, de l'asile et du contrôle des frontières, de la protection et de la sécurité sociale, de l'emploi et également de l'éducation, dans lesquels le recours aux systèmes d'IA présente des risques importants en termes de discrimination. Chaque thème indique les articles de la législation sur l'IA qui sont pertinents pour le secteur concerné.

6. Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 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, ci-après dénommé « règlement général sur la protection des données » ou « RGPD ».
7. 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, et abrogeant la décision-cadre 2008/977/JAI du Conseil, ci-après la « directive Police-Justice ».

PARTIE I

1. Contexte général du règlement sur l'IA

Le règlement sur l'IA établit « un cadre juridique uniforme, en particulier pour le développement, la mise sur le marché, la mise en service et l'utilisation de systèmes d'intelligence artificielle (ci-après dénommés « systèmes d'IA ») dans l'Union » pour « promouvoir l'adoption de l'intelligence artificielle (IA) axée sur l'humain et digne de confiance tout en garantissant un niveau élevé de protection de la santé, de la sécurité et des droits fondamentaux consacrés dans la Charte des droits fondamentaux de l'Union européenne[8] ».

Le règlement sur l'IA est entré en vigueur le 2 août 2024, et ses différentes exigences s'appliqueront progressivement au fil du temps jusqu'à son application complète en 2030. De par sa nature, le règlement sur l'IA est directement applicable et ne nécessite aucune transposition en droit national. Toutefois, de nombreuses étapes, y compris au niveau de la gouvernance nationale et de l'appareil réglementaire, doivent être mises en place par le biais de lois nationales de mise en œuvre. Au moment de la rédaction de ces lignes directrices, la mise en œuvre du règlement sur l'IA n'en est donc qu'à ses débuts et de nombreux points restent à éclaircir.

8. Règlement sur l'IA, considérant 1.

Si le règlement sur l'IA offre aux organismes de promotion de l'égalité et aux SNDH de nouvelles possibilités de prévenir et de corriger les discriminations dans les systèmes d'IA, il convient de garder à l'esprit plusieurs de ses caractéristiques lors de la lecture des présentes lignes directrices.

Tout d'abord, la définition des systèmes d'IA. Des lignes directrices sur la définition d'un système d'intelligence artificielle établie par le règlement européen sur l'IA ont été publiées par la Commission le 6 février 2025. Comme l'ont précisé plusieurs observateurs et observatrices, ces lignes directrices ne sont pas contraignantes et peuvent donner lieu à quelques confusions[9]. Les organismes pourraient prétendre que certains systèmes de PDA utilisés dans le secteur public[10] échappent au champ d'application du règlement sur l'IA, bien qu'ils présentent des risques liés à la discrimination[11]. Ces affirmations pourraient reposer sur une interprétation erronée de la définition. Par exemple, le règlement sur l'IA stipule qu'un système d'IA « peut faire preuve d'une capacité d'adaptation[12] », posséder « différents niveaux d'autonomie[13] » et a la « capacité de déduire[14] ». Les organismes de promotion de l'égalité et les SNDH sont vivement encouragés à adopter une interprétation de la définition des systèmes d'IA qui ne fait pas cas de la technique d'IA employée (par exemple, apprentissage automatique ou traitement automatique du langage naturel) et qui considère l'autonomie et l'adaptabilité comme des caractéristiques facultatives[15].

Deuxièmement, la recherche et le développement n'entrent pas dans le champ d'application du règlement sur l'IA. « [L]es activités de recherche, d'essai et de développement [...] relatives aux systèmes ou modèles d'IA avant leur mise sur le marché ou leur mise en service[16] » ne sont pas régies par le règlement sur l'IA. Cela permet de rechercher, de tester et de développer des pratiques d'IA même interdites, à condition qu'elles ne soient pas mises sur le marché ou mises en service. Toutefois, une fois mises sur le marché ou mises en service, les pratiques de l'IA entrent dans le champ d'application du règlement. Le simple fait de qualifier un déploiement d'expérience ou d'étude ne suffirait pas à faire appliquer l'exception en faveur de la recherche et du développement. La réalisation d'« essais en conditions réelles »[17] sur des systèmes d'IA à haut risque en dehors des bacs à sables régle-

9. Kris Shrishak (2025), EU's AI Act : Tread the Guidelines Lightly, Tech Policy Press, accessible en anglais à l'adresse www.techpolicy.press/eu-ai-act-tread-the-guidelines-lightly/; Algorithm Audit (Février 2025) 'Implementation of the AI Act: definition of an AI system', accessible en anglais à l'adresse https://algorithmaudit.eu/knowledge-platform/knowledge-base/guidelines_ai_act_implementation
10. Lighthouse Reports (2023), France's Digital Inquisition, disponible en anglais sur : www.lighthousereports.com/investigation/frances-digital-inquisition/, consultés le 7 novembre 2025.
11. Pour accéder à une analyse détaillée, voir Xenidis 2025.
12. Règlement sur l'IA, art. 3 (1).
13. Règlement sur l'IA, art. 3 (1).
14. Règlement sur l'IA, considérant 12.
15. Kris Shrishak (2025), EU's AI Act : Tread the Guidelines Lightly, Tech Policy Press, accessible en anglais à l'adresse https://www.techpolicy.press/eu-ai-act-tread-the-guidelines-lightly/; voir également Algorithm Audit (2025), AI Act Implementation Tool, disponible en anglais à l'adresse https://algorithmaudit.eu/technical-tools/implementation-tool/.
16. Règlement sur l'IA, art. 2 (8).
17. Règlement sur l'IA, art. 3 (57).

mentaires sur l'IA, mais pas les pratiques interdites en matière d'IA, est autorisée temporairement et ne saurait être considérée comme une mise sur le marché ou une mise en service si des conditions spécifiques sont respectées[18]. Il est donc important que les organismes de promotion de l'égalité (OPE) règlent les questions liées aux tests des systèmes d'IA à haut risque en s'appuyant sur les instruments juridiques préexistants.

Troisièmement, le règlement sur l'IA ne s'applique qu'aux systèmes qui ont été mis sur le marché ou mis en service après la date générale d'application du règlement, exception faite pour toute « importante modification » de systèmes d'IA déjà déployés. Il existe une exception pour les systèmes d'IA à haut risque destinés à être utilisés par des autorités publiques, pour lesquels les opérateurs doivent se conformer aux exigences du règlement sur l'IA d'ici le 2 août 2030[19].

Quatrièmement, l'approche fondée sur les risques adoptée par le règlement sur l'IA établit des règles et des obligations différentes en fonction du niveau de risque des systèmes d'IA. En pratique, certaines pratiques seront interdites (voir article 5), certains systèmes d'IA seront considérés comme à haut risque et leurs opérateurs seront donc soumis à de nouvelles obligations (voir article 6), et des exigences de transparence s'appliqueront à certains systèmes d'IA (voir article 50). Les systèmes qui n'entrent pas dans ce champ d'application ne seront pas soumis aux mêmes exigences, au risque que les opérateurs d'IA adoptent des pratiques d' « éthique-washing » (Equinet, 2025) et de « dérisquage » (Xenidis, 2025) pour se soustraire à ces exigences. Toutefois, cette approche basée sur les risques doit être considérée à la lumière de l'obligation pour les opérateurs d'IA de respecter les droits fondamentaux et la législation anti-discrimination.

18. L'article 60 du règlement sur l'IA énonce toutes les conditions à remplir pour les essais en conditions réelles.
19. Règlement sur l'IA, considérant 177.

2. Interdictions

2.1. Introduction aux pratiques interdites en matière d'IA

2.1.1. Contexte et pertinence

L'article 5 établit la liste des systèmes d'IA interdits par le règlement sur l'IA, car ils « sont contraires aux valeurs de l'Union relatives au respect de la dignité humaine, à la liberté, à l'égalité, à la démocratie et à l'État de droit, ainsi qu'aux droits fondamentaux consacrés dans la Charte, y compris le droit à la non-discrimination, le droit à la protection des données et à la vie privée et les droits de l'enfant »[20]. Les pratiques d'IA qui ne sont pas considérées comme interdites en vertu du règlement sur l'IA peuvent l'être en vertu d'autres lois de l'Union européenne[21].

Ces interdictions sont en vigueur depuis le 2 février 2025. Le 6 février 2025, la Commission européenne a publié des lignes directrices sur ces pratiques interdites, conformément à l'article 96, paragraphe 1, point b), du règlement sur l'IA[22].

20. Règlement sur l'IA, considérant 28 (italique ajouté).
21. Règlement sur l'IA, art. 5 (8).
22. Commission européenne (2025), Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), disponible à l'adresse suivante https://digital-strategy.ec.europa.eu/fr/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act

Bien qu'elles puissent donner des indications sur les pratiques qui devraient être interdites ou non, ces lignes directrices ont été critiquées pour leur ambiguïté[23]. Le risque serait que cette ambiguïté permette aux opérateurs de systèmes d'IA de s'engager dans des « stratégies de dérisquage », c'est-à-dire, en cas de doute, à considérer que telle ou telle pratique n'est pas interdite. En outre, ces lignes directrices ne sont pas contraignantes et seule la jurisprudence permettra de clarifier les interdictions.

Les autorités de surveillance des marchés (ASM) rendent compte chaque année à la Commission européenne (la « Commission ») des recours aux pratiques interdites recensés au cours de l'année, ainsi que des mesures prises[24]. La Commission évalue une fois par an « la nécessité de modifier la liste des pratiques interdites en matière d'IA »[25], et soumet ses conclusions au Parlement européen et au Conseil[26].

Le Bureau de l'IA établi au sein de la Commission européenne est chargé de mettre au point « une méthode objective et participative pour l'évaluation des niveaux de risque fondée sur les critères décrits dans les articles pertinents et l'inclusion de nouveaux systèmes » à l'article 5[27].

Corrélations avec d'autres règlements

La Convention-cadre du Conseil de l'Europe prévoit également la possibilité pour une partie à la Convention de mettre en place des moratoires ou des interdictions « que ces utilisations sont incompatibles avec le respect des droits de l'homme, le fonctionnement de la démocratie ou l'État de droit »[28].

2.1.2. Rôle des organismes de promotion de l'égalité et des autres SNDH dans la gestion des applications interdites en matière d'IA

Cette partie présente le rôle que les organismes de promotion de l'égalité pourraient jouer au niveau de toutes les applications interdites en matière d'IA. Les parties suivantes mettent l'accent sur les actions que les organismes de promotion de l'égalité et les SNDH peuvent entreprendre et qui sont spécifiques à certaines interdictions.

23. Voir par exemple Autoriteit Persoonsgegevens (2025), Summary and next steps call for input on prohibition on AI systems for emotion recognition in the areas of workplace or education institutions, disponible à l'adresse https://www.autoriteitpersoonsgegevens.nl/en/documents/summary-and-next-steps-call-for-input-on-prohibition-on-ai-systems-for-emotion-recognition-in-the-areas-of-workplace-or-education-institutions (2025), consulté le 11 novembre 2025.
24. Règlement sur l'IA, art. 74 (2).
25. Règlement sur l'IA, art. 112 (1).
26. Si le règlement sur l'IA prévoit que la Commission européenne évalue la nécessité de modifier la liste des pratiques interdites, elle n'autorise toutefois pas la Commission à mettre à jour cette liste dans le cadre du règlement sur l'IA. Une telle mise à jour nécessiterait une procédure législative distincte.
27. Règlement sur l'IA, article 112 (11) (b).
28. Convention-cadre du Conseil de l'Europe, art. 16 (4).

Pour toutes les interdictions confondues, les organismes de promotion de l'égalité et les SNDH peuvent :

- surveiller les applications interdites, en consolidant les exemples de cas qui ont été ou devraient être évalués en vertu de l'article 5. Ces exemples permettront d'illustrer l'importance des utilisations actuellement interdites et la nécessité éventuelle d'étendre les actions à d'autres applications en raison des risques qu'elles font peser sur les droits fondamentaux, en particulier l'égalité et la non-discrimination. Ces exemples peuvent être recueillis en collaborant avec des organisations de la société civile et des universitaires qui ont étudié ces systèmes, en analysant les plaintes reçues par les organismes de promotion de l'égalité et les SNDH et en suivant les litiges, y compris les litiges privés. Des partenariats avec l'Agence des droits fondamentaux de l'Union européenne peuvent également être envisagés.
- contribuer aux évaluations annuelles de la Commission au titre de l'article 112. On peut par exemple partager avec la Commission des éléments de preuves attestant d'applications interdites et de systèmes d'IA dangereux qui ne sont pour le moment pas interdits par le règlement sur l'IA.
- contribuer à la méthodologie participative d'évaluation des niveaux de risque du Bureau de l'IA.
- promouvoir la mise en application des interdictions par les autorités de surveillance du marché, y compris en utilisant les exemples consolidés par les mesures de contrôle.
- apporter une expertise en matière d'égalité et de discrimination aux autorités de surveillance du marché, qui devront évaluer les interdictions à la lumière des atteintes aux droits fondamentaux.
- recevoir les plaintes du public, qui peuvent inclure des problématiques liées à des déploiements effectués par des acteurs privés.

2.2. Les systèmes d'IA qui manipulent, trompent ou exploitent les vulnérabilités des personnes

2.2.1. Contexte et pertinence

Le règlement sur l'IA interdit, à l'article 5, paragraphe 1, points a) et b), les systèmes d'IA préjudiciables qui manipulent, trompent ou exploitent les vulnérabilités des personnes. Cette interdiction concerne de nombreux droits fondamentaux menacés : la dignité humaine, l'autonomie des personnes, les droits des personnes handicapées, la non-discrimination en raison de l'âge (droits de l'enfant, droits des personnes âgées) ou de la situation socio-économique.

Il est peu probable que les organismes du secteur public déploient intentionnellement des systèmes d'IA manipulateurs. Mais ils pourraient le faire de manière accidentelle, par exemple via les chatbots qu'ils déploient.

Exemples

Les chatbots sont de plus en plus utilisés dans le secteur public, notamment pour donner au public des informations sur les services publics. En 2024, un chatbot déployé par la ville de New York a fourni des informations erronées en matière de droit du travail[29]. Ces types d'outils sont également susceptibles d'être utilisés dans les secteurs de la banque et de l'assurance.

Un chatbot pourrait être interdit en vertu de l'art. 5 (1) (a) et/ou (b) :

- Si un chatbot trompe les gens en fournissant des informations trompeuses qui amènent une personne à prendre une décision préjudiciable, par exemple en l'encourageant à se suicider[30], il pourrait être interdit en vertu de l'article 5, paragraphe 1, point a).
- Toutefois, si un chatbot exploite le statut socio-économique d'une personne et ne fournit pas d'informations (ou fournit des informations erronées) sur certains services essentiels tels que l'accès aux prestations sociales sous condition de ressources, et entraîne ainsi des difficultés financières, il pourrait être interdit en vertu de l'article 5, paragraphe 1, point b), étant donné que le système d'IA exploite la situation socio-économique de la personne (voir ci-dessous). Selon les circonstances, ce cas de figure pourrait également relever du champ d'application de l'article 5, paragraphe 1, point a).

Ces deux interdictions sont étudiées ensemble car l'article 5, paragraphe 1, point b), peut être traité comme une lex specialis lorsqu'il y a chevauchement avec l'article 5, paragraphe 1, point a)[31].

Évaluation des interdictions visées à l'article 5, paragraphe 1, points a) et b)

Pour déterminer si un système d'IA est interdit en vertu de l'article 5, paragraphe 1, point a), il faut procéder à une évaluation en cinq étapes comme suit :

1. Le système d'IA a-t-il été mis sur le marché, mis en service ou est-il utilisé[32] ?
2. Le système d'IA a-t-il recours à des « techniques subliminales, au-dessous du seuil de conscience d'une personne, ou à des techniques délibérément manipulatrices ou trompeuses » ou une combinaison de ces techniques ?

29. Offenhartz, J. (2024), NYC's AI chatbot was caught telling businesses to break law. The city isn't taking it down, AP News, accessible à l'adresse https://apnews.com/article/new-york-city-chatbot-misinformation-6ebc71db5b770b9969c906a7ee4fae21, consulté le 10 novembre 2025.
30. Walker, L. (2023), Belgian man dies by suicide following exchanges with chatbot, Brussels Times, disponible à l'adresse https://www.brusselstimes.com/430098/belgian-man-commits-suicide-following-exchanges-with-chatgpt, consulté le 10 novembre 2025.
31. Lignes directrices relatives aux interdictions du règlement sur l'IA.
32. Règlement sur l'IA, art. 3 (9)-(11). Pour l'interprétation par la Commission des termes « mis sur le marché, mis en service ou en utilisé », voir le « Guide bleu » relatif à la mise en œuvre de la réglementation de l'UE sur les produits 2022, 2022/C247/01, section 2.

3. Le système d'IA est-il déployé avec « pour objectif ou effet d'altérer substantiellement le comportement d'une personne ou d'un groupe de personnes » ?

4. En altérant ce comportement, le système d'IA a-t-il amené « la personne à prendre une décision qu'elle n'aurait pas prise autrement » ?

    a. Y a-t-il eu une compromission significative de l'autonomie de la personne ou du groupe de personnes (au-delà du simple fait d'être influencé par le biais d'une persuasion légale) ?

5. La décision d'une personne ou d'un groupe de personnes a-t-elle causé « ou est raisonnablement susceptible de causer un préjudice important à cette personne, à une autre personne ou à un groupe de personnes » ?

    a. Comment ont-ils été lésés ? Sur le plan physique, psychologique, financier ?

    b. Le préjudice causé était-il important[33] ? Niveau de gravité déterminé d'après la combinaison des préjudices, les effets cumulatifs des préjudices, l'ampleur du préjudice, la réversibilité du préjudice et la durée du préjudice[34].

Pour déterminer si un système d'IA est interdit en vertu de l'article 5, paragraphe 1, point b), il faut procéder à une évaluation en quatre étapes comme suit :

1. Le système d'IA a-t-il été mis sur le marché, mis en service ou utilisé ?

2. Le système d'IA exploite-t-il « 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 » ?

    a. Âge : enfants et personnes âgées;

    b. Handicap : comprend les déficiences physiques, mentales, intellectuelles et sensorielles qui empêchent la pleine participation des individus à la société[35];

    c. Situation sociale ou économique : comprend les personnes vivant dans une extrême pauvreté, les minorités ethniques ou religieuses[36].

3. Le système d'IA est-il déployé avec « pour objectif ou effet d'altérer substantiellement le comportement de la personne ou d'une personne appartenant au groupe de personnes en question » ?

4. Le système d'IA, en altérant le comportement, a-t-il causé « ou est-il raisonnablement susceptible de causer à cette personne ou à une autre un préjudice important » ?

---

33. Cour de justice du 7 septembre 2004, Waddenvereniging et Vogelbeschermingsvereniging, C-127/02, EU:C:2004:482 et du 11 avril 2013, Sweetman et autres, C-258/11, EU:C:2013:220.
34. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 92.
35. Directive (UE) 2019/882 du Parlement européen et du Conseil du 17 avril 2019 relative aux exigences en matière d'accessibilité applicables aux produits et services (JO L 151 du 7.6.2019, p. 70)
36. Règlement sur l'IA, considérant 29 : « En outre, des systèmes d'IA peuvent également exploiter les vulnérabilités d'une personne ou d'un groupe particulier de personnes en raison de leur âge, d'un handicap au sens de la directive (UE) 2019/882 du Parlement européen et du Conseil, ou d'une situation sociale ou économique spécifique susceptible de rendre ces personnes plus vulnérables à l'exploitation, telles que les personnes vivant dans une extrême pauvreté ou appartenant à des minorités ethniques ou religieuses. » (italique ajouté)

En ce qui concerne l'article 5, paragraphe 1, point a), certains concepts importants ne sont pas nécessairement évidents. Les expressions telles que « techniques subliminales, au-dessous du seuil de conscience d'une personne », « techniques délibérément manipulatrices » et « techniques trompeuses » ne sont pas définies dans le règlement sur l'IA.

- Les techniques subliminales, au-dessous du seuil de conscience d'une personne peuvent impliquer des stimuli qui ne sont pas perçus par l'humain de façon consciente mais qui sont tout de même traités par son cerveau et peuvent influencer son comportement. Ces stimuli peuvent être sonores ou visuels, ou venir modifier la perception du temps[37]. Il peut par exemple s'agir d'« interfaces machine-cerveau ou [de] réalité virtuelle »[38]. Dans le contexte des « communications commerciales audiovisuelles »[39], la directive sur les services de médias audiovisuels (SMA) interdit déjà les techniques subliminales[40]. Dans ce cadre, l'interdiction prévue par le règlement sur l'IA, qui ne s'applique que lorsqu'un système d'IA est en cause, constitue un cas particulier de l'interdiction posée par la directive SMA. Les techniques subliminales peuvent être appréhendées comme une illustration des techniques de manipulation.
- Des techniques délibérément manipulatrices pourraient être utilisées pour « persuader des personnes d'adopter des comportements indésirables » en exploitant leurs biais ou leur état émotionnel « pour les tromper en les poussant à prendre des décisions d'une manière qui met à mal et compromet leur autonomie, leur libre arbitre et leur liberté de choix »[41]. Ce type de manipulation peut inclure l'exploitation de données personnelles pour envoyer des messages personnalisés, par exemple des publicités ciblées.
- Les techniques trompeuses consistent à présenter des informations fausses et trompeuses. La tromperie peut prendre la forme de contenus audiovisuels, tels que des « deepfakes » (ou hypertrucages) ou encore des faux chatbots. Pour déterminer si une technique est trompeuse, les mesures de transparence prévues à l'article 50 sont pertinentes car elles peuvent empêcher la tromperie en informant les utilisateurs et utilisatrices qu'ils ou elles interagissent avec un système d'IA. Cependant, malgré ces mesures de transparence, ces techniques peuvent tout de même rester manipulatrices.

L'interdiction s'applique lorsqu'un système d'IA est déployé avec « pour objectif ou effet d'altérer substantiellement le comportement d'une personne ou d'un groupe de personnes ». Cela implique que l'intention d'altérer substantiellement le comportement n'est pas nécessaire. Mais l'effet suffit. Dans ses lignes directrices, la Commission européenne indique que même une probabilité[42] d'effet, par exemple

37. Lignes directrices relatives aux interdictions du règlement sur l'IA. paragraphes 64-66.
38. Règlement sur l'IA, considérant 29.
39. Directive 2010/13/UE du Parlement européen et du Conseil du 10 mars 2010 visant à la coordination de certaines dispositions législatives, réglementaires et administratives des États membres relatives à la fourniture de services de médias audiovisuels (JO L 95 du 15.4.2010, p. 1). Ci-après la « directive SMA », Art. 1 (h).
40. Directive SMA, art. 9 (1) (b).
41. Règlement sur l'IA, considérant 29.
42. Arrêt de la Cour de justice (cinquième chambre) du 26 octobre 2016. Canal Digital Danmark A/S. EU:C:2016:800, affaire C-611/14, paragraphe 73.

en raison de la présentation des informations[43], est suffisante, sans nécessité de présenter une preuve de l'effet faisant suite au préjudice. Selon la Commission, l'effet (ou la probabilité d'effet) pourrait peser sur :

1. la personne consommatrice moyenne[44], y compris avec la reconnaissance du fait que sa capacité de prise de décision peut être entravée par des contraintes telles que des biais cognitifs[45]; et
2. les personnes ou les groupes de personnes susceptibles d'avoir été spécifiquement ciblés ou discriminés.

Néanmoins, il est nécessaire qu'il existe une vraisemblance plausible d'un lien entre l'« objectif ou effet d'altérer substantiellement » et la technique subliminale, délibérément manipulatrice ou trompeuse déployée par le système d'IA.

Pour évaluer la probabilité raisonnable qu'un préjudice important soit causé, il convient de prendre en compte les différentes formes de préjudice physique, psychologique et financier[46]. En outre, l'importance du préjudice doit être évaluée[47], par exemple, sa gravité d'après la combinaison de préjudices, les effets cumulatifs des préjudices, l'ampleur du préjudice (par exemple pour les les chatbots déployés au niveau national dans le secteur public sont susceptibles d'être utilisés par de nombreuses personnes), la réversibilité du préjudice et la durée du préjudice[48].

2.2.2. Rôle des organismes de promotion de l'égalité et des autres SNDH dans la gestion des systèmes d'IA qui manipulent, trompent ou exploitent les vulnérabilités des personnes

- Sensibiliser et mettre l'accent sur le préjudice subi par la personne plutôt que par la personne consommatrice, étant donné que les droits de la personne vont au-delà de ceux du consommateur ou de la consommatrice, tandis que les lignes directrices de la Commission font spécifiquement référence au « consommateur moyen » pour ces interdictions (voir section 2.2 ci-dessus). Cela peut se faire dans les communications publiques des organismes de promotion de l'égalité et des SNDH et dans leurs échanges avec les autorités de surveillance des marchés.

---

43. Arrêt de la Cour de justice du 19 décembre 2013, Trento Sviluppo et Centrale Adriatica, C-281/12, EU:C:2013:859
44. Communication de la Commission - Orientations concernant l'interprétation et à l'application de la directive 2005/29/CE du Parlement européen et du Conseil relative aux pratiques commerciales déloyales des entreprises vis-à-vis des consommateurs dans le marché intérieur (JO C 526 du 29.12.2021, p. 1). Voir également le considérant 18 de la directive 2005/29/CE pour la définition d'un consommateur moyen : « normalement informé et raisonnablement attentif et avisé, compte tenu des facteurs sociaux, culturels et linguistiques ».
45. Compass Banca SpA contre Autorità Garante della Concorrenza e del Mercato (AGCM), Affaire C-646/22, EU:C:2024:957
46. Règlement sur l'IA, considérant 29.
47. Cour de justice du 7 septembre 2004, Waddenvereniging et Vogelbeschermingsvereniging, C-127/02, EU:C:2004:482 et du 11 avril 2013, Sweetman et autres, C-258/11, EU:C:2013:220.
48. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 92.

- Veiller de manière proactive à ce que l'« importance » du préjudice soit déterminée compte tenu des droits fondamentaux, et en particulier de l'égalité et de la non-discrimination. Les organismes de promotion de l'égalité (OPE) pourraient évaluer l'« importance » du préjudice sur la base de la jurisprudence existante en matière de non-discrimination. Il conviendrait d'évaluer le seuil correspondant à un préjudice suffisamment important pour que l'interdiction s'applique.
  - Il sera important pour les organismes de promotion de l'égalité de recueillir des exemples, les organisations de la société civile (OSC) pouvant être des alliés utiles à cet égard. En outre, les organismes de promotion de l'égalité doivent s'intéresser aux litiges privés et suivre leur résolution.
- Veiller à ce que l'« exploitation » des vulnérabilités soit déterminée compte tenu des droits fondamentaux, et en particulier de l'égalité et de la discrimination, en élaborant des orientations et en les communiquant aux autorités de surveillance des marchés qui seront chargées d'évaluer les préjudices. Une attention particulière sera portée à la question de savoir si ce critère pourrait inclure la discrimination indirecte.

2.3. La notation sociale

2.3.1. Contexte et pertinence

L'article 5, paragraphe 1, point c), du règlement sur l'IA interdit aux acteurs publics et privés de procéder à une notation sociale inacceptable des personnes au moyen de l'IA, car cela pourrait aboutir à des « résultats discriminatoires et à l'exclusion de certains groupes[49] », et porter atteinte au droit à la dignité et à la non-discrimination.

Exemples de systèmes de notation sociale

Cette interdiction est particulièrement pertinente pour le secteur public, où la classification et l'évaluation sont des pratiques très répandues, notamment dans les domaines de l'emploi et de la sécurité sociale, de la fiscalité, de la migration, de la police ou de la justice. Les systèmes de classification et d'évaluation sont également utilisés dans les domaines de l'assurance et de la banque, comme c'est le cas du système Schufa utilisé en Allemagne pour attribuer un score de crédit aux individus[50].

Par exemple, l'agence autrichienne pour l'emploi a mis au point un algorithme, qui permet de prédire la probabilité d'emploi, afin d'allouer des ressources d'aide aux demandeur.ses d'emploi. Le prototype s'est révélé discriminatoire à l'égard des femmes (en particulier les mères célibataires) et des demandeur.ses d'emploi issus de l'immigration[51]. Aux Pays-Bas, un système utilisé pour anticiper la fraude

49. Règlement sur l'IA, considérant 31.
50. AlgorithmWatch (2018), « SCHUFA, a black box: OpenSCHUFA results published », disponible à l'adresse https://algorithmwatch.org/en/schufa-a-black-box-openschufa-results-published/ et C-634/21 SCHUFA Holding (Scoring) EU:C:2023:957.
51. Allhutter, D. et al. (2020), « Algorithmic profiling of job seekers in Austria : How austerity politics are made effective », Frontiers in Big Data, 3, disponible à l'adresse https://doi.org/10.3389/fdata.2020.00005.

s'est revelé discriminatoire à l'égard de bénéficiaires en raison de leur race[52], de leur origine ethnique et de leur nationalité[53]. En Pologne, un système utilisé par l'agence pour l'emploi a finalement été abandonné après avoir été jugé contraire à la Constitution[54]. Des systèmes de PDA et d'IA sont ou ont également été utilisés pour contrôler les bénéficiaires d'aides sociales dans de nombreux pays (par exemple en France[55], aux Pays-Bas[56], au Danemark[57], en Belgique[58]), avec des effets discriminatoires et un contrôle excessif des personnes en situation de vulnérabilité. Ces systèmes reposent généralement sur une combinaison de données personnelles et de caractéristiques individuelles, de données relatives aux interactions entre les bénéficiaires et les services publics de l'emploi, et, dans certains cas, de données issues d'entreprises privées (telles que des fournisseurs d'électricité).

Les systèmes de classification et d'évaluation peuvent également reposer sur d'autres types de données. Par exemple, un système de surveillance partiellement automatisé dans un camp de réfugié.es pourrait analyser les données provenant de caméras et de détecteurs de mouvement afin de déterminer si des individus spécifiques (tels que des migrant.es) risquent d'essayer de fuir[59].

Évaluation des interdictions prévues par l'article 5, paragraphe 1, point c)

Pour déterminer si un système d'IA relève de l'article 5, paragraphe 1, point c), il faut procéder à une évaluation en cinq étapes comme suit :

1. Le système d'IA a-t-il été mis sur le marché, mis en service ou utilisé ?

52 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.

53. De Rechtspraak (2019),« SyRI legislation in breach of European Convention on Human Rights », De Rechtspraak, disponible à l'adresse https://www.rechtspraak.nl/Organisatie-en-contact/Organisatie/Rechtbanken/Rechtbank-Den-Haag/Nieuws/Paginas/SyRI-legislation-in-breach-of-European-Convention-on-Human-Rights.aspx, consulté le 11 novembre 2025. 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.

54. Szymielewicz (2015), Profiling the unemployed in Poland.: Social and Political Implications of Algorithmic Decision Making, Fundacja Panoptykon, disponible à l'adresse https://en.panoptykon.org/profiling-unemployed-poland-report, consulté le 11 novembre 2025.

55. Romain, M. et al. (2023).

56. Mehrotra, D. et al. (2023), « Inside the suspicion machine », WIRED, disponible à l'adresse https://www.wired.com/story/welfare-state-algorithms/.

57. Geiger, G. (2023), « How Denmark's welfare state became a surveillance nightmare », WIRED, disponible à l'adresse https://www.wired.com/story/algorithms-welfare-state-politics/ et Amnesty International (2024), « Denmark: AI-powered welfare system fuels mass surveillance and risks discriminating against marginalized groups – report », disponible à l'adresse https://www.amnesty.org/en/latest/news/2024/11/denmark-ai-powered-welfare-system-fuels-mass-surveillance-and-risks-discriminating-against-marginalized-groups-report/.

58. Degrave, E. (2020), The Use of Secret Algorithms to Combat Social Fraud in Belgium, European review of Digital Administration & Law 1-2 : 167–178.

59. Exemple donné dans les Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 155.

2. L'« évaluation ou la classification de personnes physiques ou de groupes de personnes » est-elle l'objectif ou l'utilisation prévue de ce système d'IA ?

3. L'évaluation ou la classification a-t-elle eu lieu « au cours d'une période donnée » ?

    a. Le classement ponctuel n'est pas interdit; toutefois, l'interdiction s'applique si les données analysées ponctuellement s'étendent sur une période de temps[60].

4. L'évaluation ou la classification repose-t-elle sur :

    a. le comportement social de personnes physiques ou de groupes de personnes, ou sur

    b. des « caractéristiques personnelles ou de personnalité connues, déduites ou prédites[61] » ?

5. L'évaluation ou la classification aboutit-elle à un score social conduisant à un « traitement préjudiciable ou défavorable de certaines personnes physiques ou de groupes de personnes » :

    a. « dans des contextes sociaux dissociés du contexte dans lequel les données ont été générées ou collectées à l'origine »; et/ou

    b. de manière injustifiée ou disproportionnée par rapport à leur comportement social ou à sa gravité[62] ?

La classification est plus large que l'évaluation et peut être « fondée sur des caractéristiques connues telles que l'âge, le sexe et la taille [qui] ne conduisent pas nécessairement à un profilage[63] ».

L'évaluation est plus étroitement liée au concept de 'profilage', qui signifie[64] :

> « La collecte d'informations sur un individu (ou un groupe d'individus) et l'évaluation de ses caractéristiques ou de ses schémas de comportement afin de le placer dans une certaine catégorie ou un certain groupe, notamment pour analyser et/ou faire des prédictions en ce qui concerne, par exemple :

---

60. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 155. Le paragraphe s'appuie sur l'exemple d'un système utilisé dans un camp de réfugiés, où les données analysées s'étendent sur une période donnée.
61. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 158 : les « caractéristiques personnelles » peuvent inclure toute une série d'informations relatives à une personne, par exemple le sexe, l'orientation sexuelle ou les caractéristiques sexuelles, le genre, l'identité de genre, la race, l'origine ethnique, la situation familiale, l'adresse, le niveau de revenus, la composition du foyer, la profession, l'emploi ou un autre statut juridique, les performances au travail, la situation économique, les liquidités financières, la santé, les préférences personnelles, les intérêts, la fiabilité, le comportement, la localisation ou les déplacements, le niveau d'endettement, le type de voiture, etc ».
62. Le scandale des allocations familiales aux Pays-Bas est un exemple où les deux conditions de l'art. 5 (1) (c) du règlement sur l'IA sont remplies. Voir Belastingdienst treft 232 gezinnen met onevenredig harde actie, 27.11.2019, (en néerlandais). Voir également Geen powerplay maar fair play. Onevenredig harde aanpak van 232 gezinnen met kinderopvangtoeslag, 2017, p. 32.
63. Groupe de travail « Article 29 », Lignes directrices sur la prise de décision individuelle automatisée et le profilage aux fins du règlement 2016/679, WP251rev.01, 6.2.2018, p. 7. Voir également le RGPD, Art. 4 (4).
64. Ibid

- sa capacité à effectuer une tâche;
- ses intérêts; ou
- son comportement probable ».

Par exemple, le système d'évaluation du crédit utilisé en Allemagne, Schufa, qui génère un « score de probabilité » pour estimer la capacité d'une personne à honorer des paiements, a été considéré comme relevant du « profilage » par la Cour de justice de l'Union européenne (CJUE)[65].

Une pratique de notation peut entraîner un traitement défavorable même si aucun préjudice spécifique n'est causé, par exemple en soumettant une personne à davantage de contrôles, alors qu'un traitement préjudiciable entraîne forcément un préjudice. Les traitements défavorables et préjudiciables peuvent déjà être interdit par la législation européenne en matière de non-discrimination, qui protège certains groupes en fonction, par exemple, de leur âge, de leur origine ethnique et raciale, de leur sexe et de leur religion. Toutefois, la portée de l'interdiction prévue par le règlement sur l'IA est plus large et concerne les traitements sortant du champ de la législation européenne en matière de non-discrimination[66].

Si le score entraîne un « traitement préjudiciable ou défavorable », cette interdiction s'applique même si le score est produit par une organisation (une société privée de notation de la solvabilité) autre que l'organisation (l'autorité publique) qui l'utilise[67]. Cette interdiction ne se limite pas non plus à l'évaluation ou à la classification effectuée uniquement par un système d'intelligence artificielle. L'interdiction concerne également les pratiques de notation qui peuvent impliquer des évaluations humaines si les résultats du système d'IA jouent « un rôle suffisamment important dans la production de la note sociale »[68]. Par exemple, une autorité publique qui utiliserait un système d'IA pour attribuer des notes et combinerait ces notes avec une évaluation humaine d'autres faits complémentaires ferait l'objet d'une interdiction s'il en résultait un traitement préjudiciable ou défavorable.

2.3.2. Rôle des organismes de promotion de l'égalité et des SNDH dans la gestion des interdictions liées à la notation sociale

- Contrôler et évaluer l'éventail des pratiques de notation qui sortent du champ de la non-discrimination encadrée par l'UE et qui aboutissent à un traitement défavorable et préjudiciable. Les OPE devront développer des compétences supplémentaires pour les contrôler efficacement.

65. Arrêt de la Cour de justice du 7 décembre 2023, SCHUFA Holding (Scoring), C-634/21, EU:C:2023:957, point 47.
66. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 165.
67. Arrêt de la Cour de justice du 7 décembre 2023, SCHUFA Holding (Scoring), C-634/21, EU:C:2023:957, points 42-51, 60-61.
68. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 161.

2.4. Évaluation du risque de criminalité

2.4.1. Contexte et pertinence

L'article 5, paragraphe 1, point d), du règlement sur l'IA interdit l'évaluation individuelle du risque de criminalité et les prédictions « sur la seule base du profilage de ces personnes physiques ou de l'évaluation de leurs traits de personnalité et caractéristiques »[69]. Cette interdiction vise à limiter les atteintes au droit à la dignité humaine, à la non-discrimination, au droit à un procès équitable, au droit à la présomption d'innocence, au droit à une défense, à un recours effectif, au respect de la vie privée et à la protection des données[70].

Pour déterminer si un système d'IA relève de l'article 5, paragraphe 1, point d), il faut procéder à une évaluation en trois étapes comme suit :

1. Le système d'IA a-t-il été mis sur le marché, mis en service ou est-il utilisé ?
2. L'objectif du système d'IA est-il de « mener des évaluations des risques des personnes physiques visant à évaluer ou à prédire le risque qu'une personne physique commette une infraction pénale » ?
3. L'évaluation ou les prédictions se font-elles uniquement sur la base d'un
   a. profilage[71], et/ou
   b. d'une évaluation de leurs traits de personnalité et de leurs caractéristiques telles que la nationalité, le lieu de naissance, le lieu de résidence, le nombre d'enfants, le niveau d'endettement, le type de voiture, etc[72].

Par exemple, un système d'IA utilisé par une autorité policière pour prédire un comportement criminel s'agissant de crimes tels que le terrorisme, uniquement sur la base de l'âge, de la nationalité, de l'adresse, du type de voiture et de l'état matrimonial des individus, serait interdit[73].

Les exemples suivants ne sont pas interdits :

- Évaluation du niveau de risque d'un groupe d'individus (et non d'un individu)[74];
- Toute autre approche de prévision policière qui n'est pas uniquement basée sur le profilage ou l'évaluation de traits de personnalité et de caractéristiques;

69. Règlement sur l'IA, art. 5 (1) (d).
70. Règlement sur l'IA, considérant 48.
71. Le profilage qui entraîne une discrimination indirecte ou directe est déjà interdit par la directive Police-Justice, article 11 (3). Voir également Groupe de travail « Article 29 », Lignes directrices sur la prise de décision individuelle automatisée et le profilage aux fins du règlement 2016/679, WP251rev.01, 6.2.2018, p. 7; Agence des droits fondamentaux, Preventing unlawful profiling today and in the future : a guide, Handbook, 2018, p.138.
72. Règlement sur l'IA, considérant 42.
73. Lignes directrices relatives aux interdictions du règlement sur l'IA, Paragraphe 202.
74. Lignes directrices relatives aux interdictions du règlement européen sur l'IA, paragraphe 196. Il est à noter que si un profil de groupe est utilisé pour évaluer et prédire le risque qu'une personne commette une infraction similaire, cela relève du profilage et peut donc tomber sous le coup de l'interdiction.

- Systèmes d'IA utilisés pour appuyer une évaluation humaine sur la base de faits objectifs et vérifiables directement liés à une activité criminelle[75];
- Pratiques de prévision policière basées sur la localisation[76];
- Les systèmes d'IA qui font des prédictions isolées autorisées par le droit national et européen en lien avec une infraction administrative (et non avec une infraction pénale), même si « des informations peuvent être recueillies en vue d'une éventuelle implication des personnes physiques dans des infractions pénales »[77].

Cette interdiction a une portée limitée et ne proscrit pas la prévision policière en soi. Lorsque l'interdiction s'applique, elle est plus large en ce qui concerne quand et à qui elle s'applique. Le règlement sur l'IA complète la directive (UE) 2016/343 pour protéger le droit à la présomption d'innocence jusqu'à ce que la culpabilité soit prouvée avant l'ouverture d'une enquête pénale formelle[78]. La directive (UE) 2016/343 ne s'applique que lorsqu'une personne est soupçonnée ou accusée d'avoir commis une infraction pénale.

Il est important de noter que cette interdiction ne se limite pas aux autorités répressives ou aux entités agissant en leur nom. Toute entité ayant l'obligation légale d'« évaluer ou [de] prédire le risque qu'une personne physique commette une infraction pénale » est concernée. Une autorité fiscale qui établit des profils de personnes sur la base de leur nationalité ou d'autres caractéristiques à l'aide de systèmes d'IA entre dans le champ d'application de cette interdiction[79].

Une institution bancaire, une entité privée, qui est chargée par la loi de filtrer les clients ayant commis une infraction pénale telle qu'un blanchiment d'argent, tomberait également sous le coup de cette interdiction si elle utilise des systèmes d'IA et ne respecte pas le règlement (UE) 2024/1624[80].

Il est également important de souligner que l'application de systèmes de « profilage de groupe[81] » à des individus entre dans le champ de cette interdiction. Le profilage de groupe consiste à établir le profil d'un groupe spécifique, notamment autour de catégories telles que les terroristes, les criminels, etc. Ces profils peuvent être utilisés

75. Lignes directrices relatives aux interdictions du règlement sur l'IA, Paragraphe 214.
76. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphes 212 et 213. Cela signifie que les patrouilles pourraient être fortement déployées dans des zones déterminées par des algorithmes prédictifs « basés sur des données historiques, et perpétuer la discrimination et les inégalités dans le domaine de l'application de la loi. » Voir également « Cop out: automation in the criminal legal system », Georgetown Law Centre on Privacy & Technology, disponible en anglais sur https://copout.tech/, consulté le 10 novembre 2025.
77. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 217. Voir également la note de bas de page 143 des lignes directrices relatives aux interdictions du règlement sur l'IA concernant les critères permettant d'évaluer si une infraction est pénale ou non.
78. Directive (UE) 2016/343 du Parlement européen et du Conseil du 9 mars 2016 portant renforcement de certains aspects de la présomption d'innocence et du droit d'assister à son procès dans le cadre des procédures pénales.
79. L'interdiction ne s'applique pas lorsqu'une autorité fiscale évalue le niveau de risque d'une entité juridique telle qu'une société.
80. Règlement (UE) 2024/1624 du Parlement européen et du Conseil du 31 mai 2024 relatif à la prévention de l'utilisation du système financier aux fins du blanchiment de capitaux ou du financement du terrorisme, articles 20 et 76 (5) (b).
81. Agence des droits fondamentaux (2018), Preventing unlawful profiling today and in the future : a guide, Handbook, p. 21.

pour évaluer et prédire le risque que d'autres personnes commettent des crimes similaires. Cette pratique est interdite.

2.4.2. Rôle des organismes de promotion de l'égalité et des SNDH dans la gestion des interdictions liées à l'évaluation du risque de criminalité

- Explorer l'ensemble du périmètre d'interdiction, en menant ou en commandant des études sur les applications relevant de cette interdiction en dehors de l'utilisation par les services répressifs de l'évaluation des risques pour la prévention de la criminalité individuelle (par exemple : autorités fiscales ou institutions en charge du blanchiment d'argent).
- Contribuer à l'application de l'interdiction en formant et en sensibilisant les autorités compétentes à ces applications, qui peuvent ne pas être des autorités compétentes en vertu du règlement sur l'IA, afin qu'elles s'attaquent aux préjudices causés par la discrimination.

2.5. Le recours au moissonnage pour construire ou développer des bases de données de reconnaissance faciale

2.5.1. Contexte et pertinence

L'article 5, paragraphe 1), point e) du règlement sur l'IA interdit aux fournisseurs et déployeurs de développer et d'enrichir les « bases de données de reconnaissance faciale par le moissonnage non ciblé d'images faciales provenant de l'internet ou de la vidéosurveillance[82] », et porte atteinte au droit à la dignité humaine, à la non-discrimination, au respect de la vie privée et à la protection des données[83]. Cette interdiction s'applique quelle que soit la structure de stockage de la base de données. Il n'est pas nécessaire que la base de données de reconnaissance faciale soit centralisée en un seul lieu ou sous le contrôle d'une seule entité juridique. Elle peut être décentralisée. L'interdiction s'applique également si la base de données est temporaire ou existe sur une courte durée.

Toutefois, le moissonnage non ciblé d'images faciales est déjà illégal en vertu de la législation de l'UE sur la protection des données[84]. C'est son application, notamment à l'échelle extraterritoriale, qui pose problème.

82. Règlement sur l'IA, considérant 43.
83. Lignes directrices relatives aux interdictions du règlement européen sur l'IA, paragraphe 226. De Autoriteit Persoonsgegevens (2024), Dutch DPA imposes a fine on Clearview because of illegal data collection for facial recognition, disponible à l'adresse https://www.autoriteitpersoonsgegevens.nl/en/current/dutch-dpa-imposes-a-fine-on-clearview-because-of-illegal-data-collection-for-facial-recognition, consulté le 10 novembre 2025.
84. De Autoriteit Persoonsgegevens (2024), Dutch DPA imposes a fine on Clearview because of illegal data collection for facial recognition, disponible à l'adresse https://www.autoriteitpersoonsgegevens.nl/en/current/dutch-dpa-imposes-a-fine-on-clearview-because-of-illegal-data-collection-for-facial-recognition, consulté le 10 novembre 2025.

Par exemple, l'application de reconnaissance faciale commercialisée par l'entreprise américaine Clearview AI, qui repose sur le moissonnage non ciblé d'images faciales sur les réseaux sociaux, aurait été utilisée par plusieurs services répressifs dans toute l'Europe[85].

Pour déterminer si un système d'IA est concerné par l'article 5, paragraphe 1, point e), il faut procéder à une évaluation en quatre étapes comme suit :

1. Le système d'IA a-t-il été mis sur le marché, mis en service ou est-il utilisé ?
2. Ce système d'IA a-t-il recours au « moissonnage non ciblé » (voir ci-dessous) ?
3. Les « images faciales » proviennent-elles de « l'internet ou de la vidéosurveillance » ?
4. Le système d'IA est-il utilisé pour « créer ou développer des bases de données de reconnaissance faciale » ?

Le règlement sur l'IA n'interdit pas tout moissonnage quel qu'il soit. Il n'interdit pas non plus la constitution de bases de données autres que des images faciales. Ces activités, en particulier celles qui impliquent des données biométriques[86], sont déjà limitées par la législation européenne sur la protection des données[87].

Le fait qu'une entreprise qui explore un site web respecte ou non les mécanismes techniques d'exclusion[88] n'a aucune incidence sur le caractère non ciblé ou non du moissonnage. Le moissonnage ciblé d' « images ou de vidéos contenant uniquement des visages humains de personnes spécifiques ou d'un groupe de personnes prédéfini » n'est pas interdit[89]. Toutefois, si ce moissonnage ciblé est effectué pour plusieurs personnes ou groupes au cours d'une période donnée, il équivaudrait à un moissonnage non ciblé et serait donc interdit[90].

En outre, les données accessibles au public, même lorsqu'une personne a publié son image faciale sur un réseau social, sont protégées par la législation européenne sur la protection des données.

---

85. Conseil européen de la protection des données (2020), disponible à l'adresse https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_letter_out_2020-0052_facialrecognition.pdf, consulté le 10 novembre 2025.
86. Il est à noter que la définition des « données biométriques » dans le règlement sur l'IA diffère de celle du RGPD et de la directive Police-Justice. L'article 3 (34) du règlement sur l'IA définit les « données biométriques » : « données à caractère personnel résultant d'un traitement technique spécifique, relatives aux caractéristiques physiques, physiologiques ou comportementales d'une personne physique, telles que des images faciales ou des données dactyloscopiques » Le RGPD, Art 4 (14) et la directive Police-Justice, Art 3 (13), eux, définissent les « données biométriques » comme : « les données à caractère personnel résultant d'un traitement technique spécifique, relatives aux caractéristiques physiques, physiologiques ou comportementales d'une personne physique, qui permettent ou confirment son identification unique, telles que des images faciales ou des données dactyloscopiques »;
87. EDPB (2024), Avis 28/2024 relatif à certains aspects de la protection des données liés au traitement de données à caractère personnel dans le contexte des modèles d'IA, adopté le 17 décembre 2024, paragraphes 104-106.
88. Koster, M., Illyes, G., Zeller, H. et L. Sassman (2022), « Robots Exclusion Protocol », RFC 9309, DOI 10.17487/RFC9309, disponible à l'adresse https://www.rfc-editor.org/info/rfc9309, consulté le 10 novembre 2025.
89. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 229.
90. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 230.

'[L]e simple fait que des données à caractère personnel soient accessibles au public n'implique pas que « la personne concernée a manifestement rendu ces données publiques »[91].

2.5.2. Rôle des organismes de promotion de l'égalité et des SNDH dans la gestion des interdictions liées au recours au moissonnage pour construire ou développer des bases de données de reconnaissance faciale

- Appuyer la mise en application de cette interdiction, en particulier au niveau des entreprises basées en dehors de l'UE mais qui ont recours au moissonnage d'images faciales de personnes dans l'UE, par le biais d'un engagement continu auprès des autorités de surveillance et des autorités chargées de la protection des données.

2.6. La reconnaissance des émotions

2.6.1. Contexte et pertinence

Malgré les « conséquences hautement discriminatoires et compromettantes pour la dignité, les effets manipulateurs »[92] et le manque de preuves scientifiques concernant l'efficacité de la reconnaissance des émotions[93] [94], l'article 5 (1) (f) du règlement sur l'IA n'interdit la reconnaissance des émotions que sur le lieu de travail et dans les établissements d'enseignement. Toutes les autres utilisations de la reconnaissance des émotions sont considérées comme à haut risque[95], mais ne sont pas interdites.

Examples

La reconnaissance des émotions comprend « différentes technologies et opérations de traitement permettant de détecter, collecter, analyser, catégoriser, réagir à, interagir avec et apprendre les émotions des personnes »[96]. Ces technologies peuvent être utilisées dans le domaine de l'emploi, lors des processus

91. EDPB (2024), Report of the work undertaken by the ChatGPT Taskforce, adopted on 23 May 2024, paragraphe 18.
92. Codagnone, C. et al. (2022), Identification and assessment of existing and draft EU legislation in the digital field, étude pour la commission spéciale sur l'intelligence artificielle à l'ère numérique (AIDA), Département thématique des politiques économiques, scientifiques et de la qualité de la vie, Parlement européen, Luxembourg, p. 62.
93. Règlement sur l'IA, art. 3 (39) : un « système de reconnaissance des émotions » désigne un « système d'IA servant à identifier les émotions ou les intentions de personnes physiques ou à faire des déductions quant à leurs émotions ou intentions, sur la base de leurs données biométriques ».
94. Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019), Emotional Expressions Reconsidered : Challenges to Inferring Emotion From Human Facial Movements, Psychological Science in the Public Interest, 20(1), 1-68 disponible sur https://doi.org/10.1177/1529100619832930, consulté le 10 novembre 2025.
95. Règlement sur l'IA, annexe III (1) (c).
96. Lignes directrices sur les interdictions du règlement sur l'IA, paragraphe 240.

de recrutement ou pour suivre les émotions des employés, dans le domaine de la santé, pour la prévention du suicide, ou par les forces de l'ordre en tant que « détecteurs de mensonges » lors des contrôles aux frontières[97].

Pour déterminer si un système d'IA relève de l'article 5, paragraphe 1, point f), il faut procéder à une évaluation en trois étapes comme suit :

1. Le système d'IA a-t-il été mis sur le marché, mis en service ou utilisé ?
2. Le système d'IA identifie-t-il ou déduit-il des émotions ou est-il capable de déduire les émotions ou les intentions de personnes à partir de données biométriques ?
3. Le système d'IA est-il déployé sur le lieu de travail ou dans des établissements d'enseignement ?

Il est important de noter que la notion de reconnaissance des émotions est restreinte dans le règlement sur l'IA :

> « Cette notion renvoie à des émotions ou des intentions telles que le bonheur, la tristesse, la colère, la surprise, le dégoût, la gêne, l'excitation, la honte, le mépris, la satisfaction et l'amusement. Cette notion ne recouvre pas les états physiques, tels que la douleur ou la fatigue, qui comprennent, par exemple, des systèmes utilisés pour déceler l'état de fatigue des pilotes ou des conducteurs professionnels aux fins de la prévention des accidents. Elle ne recouvre pas non plus la simple détection d'expressions, de gestes ou de mouvements dont l'apparence est immédiate, à moins que ceux-ci ne soient utilisés pour identifier ou déduire des émotions. Ces expressions peuvent être des expressions faciales toutes simples telles qu'un froncement de sourcils ou un sourire, ou des gestes tels qu'un mouvement de mains, de bras ou de tête, ou encore des caractéristiques de la voix d'une personne, comme le fait de parler fort ou de chuchoter ».[98]

Pour autant, le paragraphe ci-dessus ne signifie pas qu'il s'agit de pratiques légales. La reconnaissance des émotions implique le traitement de données biométriques qui, en l'absence de base juridique valable, serait illicite au regard de la législation européenne en matière de protection des données[99].

Si la notion de lieu de travail comprend le processus de recrutement et protège les salariés ainsi que les travailleur.ses indépendant.es, selon les lignes directrices de la Commission européenne, elle ne couvre pas le reste des personnes. Par exemple, « l'utilisation de webcams et de systèmes de reconnaissance vocale par un centre d'appel pour suivre les émotions de ses clients, telles que la colère ou l'impatience, n'est pas interdite », tandis que leur utilisation « par un supermarché [...] pour suivre les émotions de ses employés » est interdite, mais le suivi des émotions des clients et clientes ne l'est pas[100].

La notion d'établissement d'enseignement concerne tous les niveaux d'établissements d'enseignement, qu'ils soient publics ou privés, « accrédités ou agréés par les

97. Boffey, D. (2018), EU border "lie detector" system criticised as pseudoscience, The Guardian, disponible à l'adresse https://www.theguardian.com/world/2018/nov/02/eu-border-lie-detection-system-criticised-as-pseudoscience, consulté le 10 novembre 2025.
98. Règlement sur l'IA, considérant 18 (italique ajouté).
99. RGPD, art. 6 (licéité du traitement) et 9 (catégories particulières de données) en particulier.
100. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 254.

autorités nationales compétentes en matière d'éducation ou par des autorités équivalentes ». L'utilisation de la reconnaissance des émotions dans les établissements d'enseignement, y compris lors des procédures d'admission et des examens, est concernée; mais l'interdiction ne s'applique pas aux cours, y compris les cours en ligne, qui sont proposés par des entités non considérées comme des établissements d'enseignement[101].

Même sur le lieu de travail et dans les établissements d'enseignement, la reconnaissance des émotions est autorisée pour des raisons médicales ou de sécurité si elle est strictement nécessaire et proportionnée. Les systèmes de reconnaissance des émotions utilisés pour des raisons médicales doivent être conformes au règlement (UE) 2017/745 (règlement sur les dispositifs médicaux), au droit de l'Union européenne et au droit national en matière d'emploi et de conditions de travail, notamment en ce qui concerne la santé et la sécurité au travail, qui peuvent restreindre leur utilisation. En d'autres termes, un marquage CE[102] pour un système de reconnaissance des émotions concerné par le règlement sur les dispositifs médicaux est nécessaire mais pas suffisant pour pouvoir utiliser ce système sur le lieu de travail et dans les établissements d'enseignement en vertu du règlement sur l'IA.

En outre, toute utilisation de cet ordre nécessite un « avis d'expert préalable, écrit et motivé, concernant le cas d'utilisation en question... [dont] la nécessité doit être évaluée de manière objective par rapport à l'objectif médical et de sécurité, sans égard aux 'besoins' de l'employeur ou de l'établissement d'enseignement. Cette évaluation doit permettre de déterminer s'il existe d'autres moyens moins intrusifs qui permettraient d'atteindre le même objectif ».[103]

Les systèmes de reconnaissance des émotions non interdits sont considérés comme présentant un risque élevé en vertu de l'annexe III, paragraphe 1, point c).

Quelques exemples[104], si et seulement s'il existe une base juridique valable :

- Autorités statistiques ayant recours à la reconnaissance des émotions dans les isoloirs pour connaître le positionnement des citoyens à l'égard de la démocratie (par exemple : colère, satisfaction)
- Entreprise ayant recours à un chatbot qui se sert de la reconnaissance des émotions pour réagir de manière appropriée aux clients très mécontents;
- Autorité policière utilisant un système de reconnaissance des émotions lors de l'interrogatoire d'un suspect.

Il est à noter que dans l'ensemble de ces cas, on a affaire à une utilisation de données biométriques pour déduire ou identifier des émotions.

---

101. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 257.
102. Commission européenne (n.d.), Marquage CE, disponible sur https://single-market-economy.ec.europa.eu/single-market/ce-marking_en?prefLang=fr, consulté le 10 novembre 2025
103. Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 259.
104. Wendehorst C., et Duller. Y (2021), Biometric Recognition and Behavioural Detection Assessing the ethical aspects of biometric recognition and behavioural detection techniques with a focus on their current and future use in public spaces, étude pour la commission des affaires juridiques et des pétitions du Parlement européen, Parlement européen, p. 66.

2.6.2. Rôle des organismes de promotion de l'égalité et des SNDH dans la gestion des interdictions liées à la reconnaissance des émotions

- Recueillir des preuves des préjudices et des risques liés à la discrimination dus au recours à la reconnaissance des émotions dans des domaines autres que le lieu de travail et les établissements d'enseignement. En particulier, les organismes de promotion de l'égalité et les SNDH peuvent contrôler les systèmes de reconnaissance des émotions considérés comme présentant un risque élevé en vertu de l'annexe III, paragraphe 1, point c), du règlement sur l'IA, qui seront enregistrés dans la base de données des systèmes à haut risque (voir article 49).
- Recueillir des preuves des motifs avancés par les opérateurs d'IA pour étayer le caractère nécessaire et proportionné du recours à la reconnaissance des émotions pour des raisons médicales et de sécurité afin de vérifier le respect de la loi.
- Protéger les client.es, les usager.ères et les autres personnes interagissant avec les travailleurs et travailleuses sur leur lieu de travail, cette finalité n'étant pas interdite par l'article 5, paragraphe 1, point f).

2.7. Catégorisation biométrique

2.7.1. Contexte et pertinence

L'article 5, paragraphe 1, point (g), du règlement sur l'IA interdit de catégoriser des personnes physiques individuellement « 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 »[105], car cette catégorisation porte atteinte au droit à la dignité humaine, à la non-discrimination, au respect de la vie privée et à la protection des données. Il est à noter que la catégorisation biométrique telle qu'expliquée ci-dessous diffère de l'identification biométrique (à distance), qui consiste à identifier des personnes (cf. 2.8.).

Contrairement à l'expression « système de catégorisation biométrique », l'expression « catégorisation biométrique » n'est pas défini dans le règlement sur l'IA, mais peut être compris comme le processus consistant à :

> « Établir si les données biométriques d'une personne appartiennent à un groupe qui présente certaines caractéristiques prédéfinies en vue de prendre des mesures spécifiques. Dans ce cas, l'important n'est pas l'identification ou la vérification d'une personne, mais son inscription automatique dans une certaine catégorie Par exemple, un panneau publicitaire peut afficher différentes annonces selon la personne qui le regarde, en fonction de l'âge ou du sexe de cette dernière[106]. ». [non en italique dans l'original]

105. Règlement sur l'IA, art. 5 (1) (g).
106. Groupe de travail « Article 29 », avis 3/2012 sur l'évolution des technologies biométriques, WP193, 27.4.2012, p. 6.

Exemples

Parmi les autres exemples de catégorisation biométrique, on peut citer l'utilisation de logiciels pour catégoriser automatiquement les personnes en fonction de leur race ou de leur sexe, ou un système d'IA qui analyse les photos d'une personne sur les réseaux sociaux pour en déduire son orientation politique et lui envoyer des messages ciblés, ou pour en déduire son orientation sexuelle et lui envoyer de la publicité ciblée.

La catégorisation biométrique est une forme de profilage[107]. L'article 22, paragraphe 1, du RGPD et l'article 11, paragraphe 3, de la directive Police-Justice interdisent la discrimination indirecte ou directe fondée sur le profilage.

Pour déterminer si un système d'IA relève de l'article 5, paragraphe 1, point g), il faut procéder à une évaluation en cinq étapes comme suit :

1. Le système d'IA a-t-il été mis sur le marché, mis en service ou utilisé ?
2. Le système d'IA est-il basé sur des « données biométriques » ?
3. Le système d'IA est-il un « système de catégorisation biométrique » au sens de l'article 3, paragraphe 40 ?
4. Des individus sont-ils catégorisés par ce « système de catégorisation biométrique » ?
    a. L'objectif premier du système d'IA est-il de classer des « personnes physiques dans certaines catégories sur la base de leurs données biométriques » ?
5. Le système d'IA est-il utilisé pour « 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 » ?

Hors du champ d'application de l'interdiction, mais considérés comme à haut risque en vertu de l'annexe III, paragraphe 1, point b) :

- « L'étiquetage ou le filtrage d'ensembles de données biométriques acquis légalement, tels que des images, fondés sur des données biométriques ou la catégorisation de données biométriques dans le domaine répressif »[108].
- La catégorisation biométrique qui est « accessoire à un autre service commercial et strictement nécessaire pour des raisons techniques objectives »[109].

Il est à noter que ces utilisations peuvent tout de même être interdites en vertu du RGPD et de la directive Police-Justice si la finalité du traitement n'est pas licite ou si l'atteinte aux droits fondamentaux des personnes due au traitement des données biométriques n'est pas nécessaire et proportionnelle.

---

107. Groupe de travail « Article 29 », Lignes directrices sur la prise de décision individuelle automatisée et le profilage aux fins du règlement 2016/679, WP251rev.01, 6.2.2018, p. 7. Voir également RGPD, Art. 4 (4)
108. Règlement sur l'IA, art. 5 (1) (g).
109. Règlement sur l'IA, art. 3 (40).

2.7.2. Rôle des organismes de promotion de l'égalité et des SNDH dans la gestion des interdictions liées à une catégorisation biométrique

- Accumuler des preuves de préjudices liés à la discrimination en raison d'une catégorisation biométrique et les partager avec les ASM et la Commission européenne afin de contribuer à l'évaluation annuelle de l'article 5 et de l'annexe III, conformément à l'article 112, à l'aide d'exemples concrets. En particulier, les organismes de promotion de l'égalité et les SNDH peuvent collaborer avec les organisations de la société civile et les autorités de surveillance des marchés pour accorder une attention particulière aux utilisations à haut risque dans le domaine répressif, qui seront enregistrées dans la version non publique de la base de données des systèmes d'IA à haut risque (voir l'article 49).

2.8. L'identification biométrique à distance

2.8.1. Contexte et pertinence

Le règlement sur l'IA aborde les systèmes d'identification biométrique à distance (IBD[110]) dans plusieurs de ses articles. En vertu de l'annexe III, paragraphe 1, point a), tous les systèmes d'IBD[111] autorisés par la législation nationale ou européenne sont considérés comme des systèmes d'IA à haut risque. Étant donné que les systèmes d'IBD traitent des données biométriques, ce qui « constitue en toutes circonstances une atteinte grave en soi [aux droits garantis par la Charte de l'Union européenne] »[112], une base juridique est nécessaire pour permettre une telle ingérence. Les technologies de reconnaissance faciale et vocale sont deux exemples courants de systèmes d'IBD. Le recours à l'identification biométrique à des fins autres que répressives est déjà interdit de manière générale[113].

Dans le domaine répressif, une « simple transposition en droit interne de la clause générale inscrite à l'article 10 »[114] de la directive Police-Justice ne suffit pas à établir une base juridique. Le règlement sur l'IA ne fournit pas non plus de base juridique pour une ingérence de cet ordre[115]. Ainsi, le traitement des données biométriques, y compris par le biais de systèmes d'IBD dans le domaine répressif, est interdit dans les

110. Règlement sur l'IA, art. 3 (41) : « 'système d'identification biométrique à distance' [désigne] un système d'IA destiné à identifier des personnes physiques sans leur participation active, généralement à distance, en comparant les données biométriques d'une personne avec celles qui figurent dans une base de données ». Voir également le règlement sur l'IA, article 3 (35) et les Lignes directrices relatives aux interdictions du règlement sur l'IA, paragraphe 306.
111. Ces systèmes peuvent être des systèmes de reconnaissance faciale, des systèmes de reconnaissance vocale à distance, des systèmes de reconnaissance de la démarche, etc.
112. EDPB (2022), Lignes directrices 05/2022 sur l'utilisation de la technologie de reconnaissance faciale dans le domaine répressif, p.5
113. Règlement sur l'IA, considérant 39. Voir également RGPD, art. 9 (1) et RGPD UE, art. 10 (1).
114. EDPB (2022), Lignes directrices 05/2022 sur l'utilisation de la technologie de reconnaissance faciale dans le domaine répressif, p.5
115. Règlement sur l'IA, considérant 38.

pays de l'UE, sauf si une base juridique spécifique a été établie. Au moins deux pays (les Pays-Bas[116] et l'Italie[117]) n'ont pas encore établi une base juridique de cet ordre.

Systèmes d'IBD en temps réel

Le règlement sur l'IA fait la distinction entre les systèmes d'IBD en temps réel et les systèmes d'IBD a posteriori[118]. Ces derniers désignent les systèmes d'IBD qui ne fonctionnent pas en temps réel. L'utilisation de systèmes d'IBD en temps réel peut « susciter un sentiment de surveillance constante et dissuader indirectement l'exercice de la liberté de réunion et d'autres droits fondamentaux » et donner lieu à des cas de discrimination fondée sur l'âge, l'appartenance ethnique, la race, le sexe ou le handicap[119].

S'agissant de cette problématique, l'article 5, paragraphe 1, point h), interdit l'IBD en temps réel dans les espaces accessibles au public à des fins répressives. Cela signifie que l'interdiction s'applique non seulement lorsque les autorités répressives[120] ont recours à l'IBD en temps réel, mais également dans le cas où une autre entité, telle qu'une société de transport public ou un club sportif, y a recours à des fins répressives, comme ce serait le cas si une autorité répressive leur déléguait le déploiement de ce type de système[121]. En ce qui concerne cette interdiction, le règlement sur l'IA s'applique en tant que lex specialis à l'article 10 de la directive Police-Justice[122].

Pour déterminer si un système d'IA relève de l'article 5, paragraphe 1, point h), il faut procéder à une évaluation en cinq étapes comme suit :

1. Le système d'IA est-il un système d'IBD au sens de l'article 3 (41) ?
2. Le système d'IBD est-il utilisé ?
   - Cela implique que l'interdiction ne s'applique qu'aux déployeurs (et non aux fournisseurs comme dans le cas d'autres interdictions)
3. Le système d'IBD est-il utilisé dans des « espaces accessibles au public » ?
4. Le système d'IBD est-il un système « en temps réel » ?
5. Le système d'IBD en temps réel est-il utilisé à des fins répressives ?

---

116. Galič, M., & Stevens, L. (2023), Regulating police use of facial recognition technology in the Netherlands: The complex interplay between criminal procedural law and data protection law, New Journal of European Criminal Law, 14(4), 459-478, disponible à l'adresse https://doi.org/10.1177/20322844231212834, consulté le 10 novembre 2025.
117. Garante per la protezione dei dati personali (2021), Riconoscimento facciale: Sari Real Time non è conforme alla normativa sulla privacy, disponible sur https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9575842, consulté le 10 novembre 2025.
118. Règlement sur l'IA, art. 3 (42) et (43). L'art. 3 (42) précise que « 'système d'identification biométrique à distance en temps réel' [désigne] un système d'identification biométrique à distance dans lequel l'acquisition des données biométriques, la comparaison et l'identification se déroulent sans décalage temporel important et qui comprend non seulement l'identification instantanée, mais aussi avec un léger décalage afin d'éviter tout contournement des règles ».
119. Règlement sur l'IA, considérant 32. (italique ajouté)
120. Règlement sur l'IA, art. 3 (45).
121. Règlement sur l'IA, art. 3 (46) : « 'activités répressives' [désigne] des activités menées par les autorités répressives ou pour leur compte pour la prévention et la détection des infractions pénales, les enquêtes et les poursuites en la matière ou l'exécution de sanctions pénales, y compris la protection contre les menaces pour la sécurité publique et la prévention de telles menaces »;
122. Règlement sur l'IA, considérant 38.

Cette interdiction n'est toutefois pas absolue. L'utilisation de systèmes d'IBD en temps réel à des fins répressives pourrait être autorisée si cela s'avère strictement nécessaire dans les cas suivants :

1. Recherche ciblée de victimes de trois crimes graves spécifiques[123] et de personnes disparues;
2. Prévention de menaces imminentes pour la vie ou la sécurité physique ou d'une menace réelle d'attaques terroristes; et/ou
3. Localisation ou identification de suspects et d'auteurs de certaines infractions pénales[124].

Le règlement sur l'IA n'établit pas de base juridique pour ce type d'utilisation. Dans ces cas de figure exceptionnels, l'IBD en temps réel ne peut être utilisé à des fins répressives que si :

- une loi nationale est adoptée, qui établit une base juridique pour l'IBD en temps réel et autorise un ou plusieurs des trois cas;
- l'autorité répressive a procédé à une analyse d'impact sur les droits fondamentaux (AIDF) conformément à l'article 27 à des fins d'évaluation de la nécessité et de la proportionnalité[125];
- pour chaque utilisation, une autorité répressive (ou une entité agissant en son nom) souhaite avoir recours à l'IBD en temps réel, l'autorité répressive doit[126] :
  - obtenir une autorisation d'une autorité judiciaire ou d'une autorité administrative indépendante dont la décision sera contraignante dans le pays de l'UE concerné (sauf en cas d'urgence où une autorisation ad hoc est requise);
  - notifier l'autorité de surveillance du marché (ASM) et l'autorité de protection des données (APD);
  - ajouter les informations relatives à l'utilisation à la base de données non publique de l'UE conformément à l'article 49, paragraphe 4.
- L'APD et l'ASM doivent soumettre à la Commission européenne un rapport annuel indiquant la fréquence, etc. du recours à l'IBD en temps réel dans leur pays. Elles peuvent choisir d'envoyer un rapport commun[127].
- Sur la base des rapports des APD et des ASM des pays de l'UE, la Commission européenne doit publier des rapports annuels[128].

Si une loi nationale est créée pour établir une base juridique, les règles relatives aux systèmes d'IA à haut risque s'appliqueront à cette utilisation à des fins répressives.

---

123 Les trois crimes graves sont l'enlèvement, la traite d'êtres humains et l'exploitation sexuelle d'êtres humains.

124 Règlement sur l'IA, art. (5) (1) (h) (iii) : Les infractions pénales entrent dans le champ d'application concerné dès lors qu'elles remplissent les deux conditions : (1) elles sont répertoriées dans le règlement sur l'IA, à l'annexe II, et (2) elles sont « passibles, dans l'État membre concerné, d'une peine ou d'une mesure de sûreté privative de liberté d'une durée maximale d'au moins quatre ans ».

125. Règlement sur l'IA, art. 5 (2).

126. Règlement sur l'IA, art. 5 (2)-(5).

127. Règlement sur l'IA, art. 5 (6).

128. Règlement sur l'IA, art. 5 (7).

Si une loi nationale n'établit pas de base juridique pour le recours à l'IBD en temps réel, qui nécessite d'abord une base juridique pour permettre le traitement de données biométriques, il existe alors une interdiction générale, effective à partir du 2 février 2025, de recourir à l'IBD en temps réel dans les espaces publics à des fins répressives. Les États membres n'ont pas de délai pour adopter une loi nationale de cet ordre. À la date du présent rapport, aucun pays de l'UE n'a encore adopté une loi nationale de cet ordre. Les pays de l'UE peuvent appuyer l'interdiction générale en n'adoptant pas de loi nationale à cet égard.

En résumé, les systèmes d'IBD en temps réel sont interdits, à moins qu'ils ne soient autorisés à des fins répressives par une loi nationale qui en autorise l'utilisation pour des objectifs spécifiques, avec des contrôles de proportionnalité et de nécessité, car ce type d'utilisation porte atteinte à une série de droits fondamentaux, y compris le droit à la non-discrimination.

Affaire critique portée devant la Cour européenne des droits de l'homme : *Glukhin c. Russie* – Requête n° 11519/20 (2023)

Lors de leurs activités de surveillance de routine d'Internet, la police a découvert des photographies et une vidéo du requérant effectuant une manifestation individuelle dans le métro de Moscou, qui avaient été publiées sur un canal Telegram public. D'après le requérant, la police a eu recours à une technologie de reconnaissance faciale pour l'identifier à partir de captures d'écran du canal, a récupéré des enregistrements vidéo issus des caméras de vidéosurveillance installées dans les stations du métro de Moscou par lesquelles il était passé, et plusieurs jours plus tard, a eu recours à une technologie faciale en direct afin de le localiser et procéder à son arrestation alors qu'il prenait le métro.

Les différentes captures d'écran ont servi de preuves dans le cadre de la procédure d'infraction administrative engagée contre le requérant, qui a été condamné.

La Cour européenne des droits de l'homme a constaté une violation des articles 10 et 8.

Le recours à une technologie de reconnaissance faciale dans le cadre d'une procédure administrative en vue d'identifier, de localiser et d'arrêter un manifestant pacifique était susceptible d'avoir un effet dissuasif vis-à-vis des droits à la liberté d'expression et de réunion (article 10). Pour mettre en œuvre une technologie de reconnaissance faciale, il faut :

- des règles détaillées régissant la portée et l'application des mesures,
- des mesures de protection robustes contre le risque d'abus et de décision arbitraire.

La Cour a estimé que cela était d'autant plus vrai dans le cadre du recours à une technologie de reconnaissance faciale, et sans aller jusqu'à interdire le recours à ces technologies de manière générale, la Cour a jugé que dans le cas de Glukhin, le fait de recourir à cette technologie pour retrouver, localiser et arrêter le requérant ne correspondait pas à un « besoin social impérieux » et ne pouvait pas être considéré comme « nécessaire dans une société démocratique ».

Fiche thématique de la Cour pour les affaires impliquant de nouvelles technologies (en anglais)

Systèmes d'IBD a posteriori

Les systèmes d'IBD qui ne sont pas interdits, soit par le règlement européen sur l'IA, soit par une autre loi nationale ou européenne, sont considérés comme étant à haut risque. Par définition, il s'agit de tous les systèmes d'IBD a posteriori (en plus des systèmes d'IBD en temps réel autorisés) qui sont autorisés par le droit national ou de l'Union européenne.

Toutes les règles applicables aux systèmes d'IA à haut risque s'appliquent à ces systèmes d'IBD a posteriori. L'utilisation de systèmes d'IBD a posteriori par des autorités répressives est soumise à une obligation supplémentaire pour les déployeurs.

Les déployeurs de systèmes d'IBD a posteriori « pour la recherche ciblée d'une personne soupçonnée d'avoir commis une infraction pénale ou condamnée pour avoir commis une infraction pénale » [129]sont tenus d'obtenir une autorisation d'une autorité judiciaire ou d'une autorité administrative indépendante dont la décision sera contraignante dans le pays de l'UE concerné. Cette autorisation doit être obtenue avant chaque utilisation du système, et au moins 48 heures avant. Chaque utilisation de ce type doit être « limitée à ce qui est strictement nécessaire pour enquêter sur une infraction pénale spécifique »[130].

2.8.2. Rôle des organismes de promotion de l'égalité et des autres SNDH dans la gestion des interdictions liées à l'identification biométrique à distance

- Encourager les gouvernements à maintenir l'interdiction générale d'utiliser les systèmes d'IBD en temps réel à des fins répressives.

129. Règlement sur l'IA, art. 26 (10).
130. Règlement sur l'IA, art. 26 (10).

3. Les systèmes d'IA à haut risque

3.1. Classification des systèmes d'IA à haut risque

3.1.1. Contexte et pertinence

L'article 6 définit les règles de classification des systèmes d'IA à haut risque qui entrent dans le champ d'application du règlement sur l'IA. Les applications répertoriées à l'annexe III, qui couvrent les domaines suivants, présentent un intérêt particulier pour les OPE :

- Biométrie[131];
- Infrastructures critiques (infrastructures numériques, trafic routier, fourniture d'eau, gaz, électricité)[132];
- Éducation et formation professionnelle[133];

131. Règlement sur l'IA, annexe III (1) et considérant 54.
132. Règlement sur l'IA, annexe III (1) et considérant 55.
133. Règlement sur l'IA, annexe III (1) et considérant 56.

- Emploi, gestion de la main d'œuvre et accès à l'emploi indépendant[134];
- Accès et jouissance de services privés essentiels et des services et prestations publics essentiels (par exemple : sécurité sociale, évaluation du crédit, assurance, soins de santé, services d'urgence)[135];
- Répression[136];
- Migration, asile et gestion des contrôles aux frontières[137];
- Administration de la justice et processus démocratiques[138].

Les considérants du règlement sur l'IA relatifs à ces domaines soulignent que la notion de risque ne se limite pas aux caractéristiques technologiques, mais qu'elle est étroitement liée au contexte sociétal, et mettent en garde contre le risque de renforcement des inégalités structurelles.

Comparés aux autres systèmes d'IA, les systèmes d'IA à haut risque sont soumis à des obligations plus en matière de documentation, d'assurance qualité et de transparence, ce qui constitue autant de garanties contre la discrimination.

Toutefois, les fournisseurs des systèmes d'IA de ces secteurs peuvent choisir de ne pas être considérés comme à haut risque s'ils estiment que leur cas particulier ne présente pas de « risque significatif » (voir ci-dessous). Ce faisant, les fournisseurs qui choisissent de ne pas être répertoriés comme à haut risque n'auront pas à se conformer aux obligations relatives aux systèmes d'IA à haut risque.

Ce choix repose sur une auto-évaluation des fournisseurs, qui n'ont aucune obligation de publier cette évaluation (voir ci-dessous). Ainsi, il y a un risque que certains systèmes à haut risque non réglementés soient déployés à tort comme n'étant pas à haut risque.

Les règles relatives aux systèmes d'IA à haut risque pour les systèmes d'IA relevant des domaines répertoriés à l'annexe III entreront en vigueur en août 2026. Le Bureau de l'IA devrait publier des lignes directrices pour la classification à haut risque des systèmes d'IA d'ici le 2 février 2026[139]. Ces lignes directrices devraient inclure des « exemples pratiques de cas d'utilisation de systèmes d'IA qui sont à haut risque et de cas d'utilisation qui ne le sont pas »[140]. Conformément à l'article 6, paragraphes 6 à 8, sur les règles de classification des systèmes d'IA à haut risque, des actes délégués peuvent ajouter, supprimer ou modifier les conditions de l'article 6, paragraphe 3, mais ne doivent pas diminuer « le niveau global de protection de la santé, de la sécurité et des droits fondamentaux »[141].

134. Règlement sur l'IA, annexe III (1) et considérant 57.
135. Règlement sur l'IA, annexe III (1) et considérant 58.
136. Règlement sur l'IA, annexe III (1) et considérant 59.
137. Règlement sur l'IA, annexe III (1) et considérant 60.
138. Règlement sur l'IA, annexe III (1) et considérant 61.
139. Règlement sur l'IA, art. 6 (5).
140. Ibid.
141. Règlement sur l'IA, art. 6 (8).

Risque important et conditions d'exclusion

Bien que l'expression « risque important » ne soit pas définie dans le règlement sur l'IA, elle peut être interprétée comme désignant un risque qui a un impact négatif sur les droits fondamentaux, notamment sur « le droit à la dignité humaine, le respect de la vie privée et familiale, la protection des données à caractère personnel, la liberté d'expression et d'information, la liberté de réunion et d'association, le droit à la non-discrimination, le droit à l'éducation, la protection des consommateurs, les droits des travailleurs, les droits des personnes handicapées, l'égalité de genre, les droits de propriété intellectuelle, le droit à un recours effectif et à accéder à un tribunal impartial, les droits de la défense et la présomption d'innocence, et le droit à une bonne administration »[142], les droits des enfants et le droit à un niveau élevé de protection de l'environnement[143]. L'évaluation peut prendre en compte la gravité du préjudice et la probabilité qu'il se produise.

En particulier, le « risque important » doit être évalué à la lumière des quatre conditions suivantes, dont chacune pourrait être invoquée par les fournisseurs de systèmes d'IA pour choisir d'être exclu de ce champ d'application[144] :

a. Le système d'IA est destiné à exécuter une tâche procédurale étroite (par exemple, un système d'IA qui classe les documents entrants par catégories);

b. Le système d'IA est destiné à améliorer le résultat d'une activité humaine déjà réalisée (par exemple, un système d'IA destiné à améliorer la langue employée dans des documents déjà rédigés);

c. Le système d'IA vise à détecter des schémas de prise de décision ou des écarts par rapport à des schémas de prise de décision antérieurs, et n'est pas destiné à remplacer ou à influencer l'évaluation humaine effectuée antérieurement, sans examen humain approprié (par exemple, un système d'IA qui « compte tenu de certaines constantes habituelles observées chez un enseignant au niveau de la notation, [peut] être [utilisé] pour vérifier a posteriori si l'enseignant s'est éventuellement écarté de ces constantes, de manière à signaler d'éventuelles incohérences ou anomalies »); ou

d. Le système d'IA est destiné à effectuer une tâche préparatoire à une évaluation pertinente aux fins des cas d'utilisation répertoriés à l'annexe III (par exemple, systèmes d'IA utilisés pour la traduction de documents initiaux).

Le « profilage »[145], tel qu'il est défini dans la législation européenne sur la protection des données, est déjà considéré comme un « risque important ». Les systèmes d'IA relevant de l'annexe III et ayant recours au profilage sont toujours considérés comme étant à haut risque[146].

---

142. Règlement sur l'IA, considérant 48 (italique ajouté).
143. Charte de l'UE, art. 37.
144. Règlement sur l'IA, art. 6 (3) et considérant 53. Ces conditions n'empêchent pas l'interdiction d'un système d'IA.
145. RGPD, Art. 4(4) et directive Police-Justice, Art 3 (4).
146. Règlement sur l'IA, art. 6 (3) troisième alinéa.

Conséquences de l'exclusion du régime des systèmes d'IA à haut risque

1. Si les fournisseurs de systèmes d'IA choisissent eux-mêmes de s'exclure de cette catégorie, ils doivent :
   - documenter leur évaluation pour s'exclure de cette catégorie avant la mise sur le marché ou la mise en service du système d'IA;
   - enregistrer leur système dans la base de données de l'UE conformément à l'article 49, paragraphe 2, et à l'article 71, en indiquant les informations requises à l'annexe VIII, section B. Les informations requises sont moins détaillées que celles exigées à l'annexe VIII, section A (conformément à l'article 49, paragraphe 1) pour les systèmes considérés comme étant à haut risque. Toutefois, les informations comprennent « la ou les conditions visées à l'article 6, paragraphe 3, sur la base desquelles le système d'IA est considéré comme n'étant pas à haut risque »[147];
2. Les fournisseurs de systèmes d'IA ne sont pas tenus de publier leur auto-évaluation. Toutefois, ils sont tenus de fournir les documents de l'évaluation à la demande des autorités nationales compétentes[148].
3. Les ASM peuvent réaliser une évaluation des systèmes d'IA pour déterminer s'ils sont à haut risque et exiger du prestataire qu'il prenne des mesures correctives[149]. L'ASM peut infliger une amende au fournisseur pour avoir classé à tort le système d'IA comme n'étant pas à haut risque dans le but de contourner le règlement sur l'IA[150]. (voir article 77)

Problématiques liées aux pratiques de dérisquage

Ensemble, l'auto-évaluation pour permettre l'exclusion, l'absence de définition de l'expression « risque important » et les conditions d'exclusion énoncées à l'article 6, paragraphe 3, présentent le risque que les fournisseurs se livrent à des « pratiques de dérisquage », c'est-à-dire qu'ils excluent leur système du régime à haut risque bien que leurs systèmes présentent un risque important, échappant ainsi aux obligations fixées pour les systèmes d'IA à haut risque.

L'absence de liste exhaustive d'exemples pratiques exacerbe ce risque. Par exemple, les systèmes d'IA utilisés pour analyser les CV[151], qui peuvent avoir un impact sur le choix des candidat.es à recevoir en entretien, pourraient ne pas être réglementés s'ils sont considérés comme des tâches procédurales étroites ou des tâches préliminaires. Il en va de même pour le cas d'un système d'IA utilisé pour traduire des documents dans le cadre de demandes d'asile. Ces systèmes peuvent être moins performants

147. Annexe VIII, section B (6).
148. Règlement sur l'IA, art. 6 (4). Les organismes de promotion de l'égalité ne sont pas considérés comme des autorités nationales compétentes au sens de l'article 6, paragraphe 4, du règlement sur l'IA. Toutefois, les organismes de promotion de l'égalité ont la possibilité d'accéder aux mêmes informations en vertu des pouvoirs conférés par le règlement sur l'IA, art. 77. En outre, les directives sur les normes, art. 8, leur confèrent des pouvoirs d'investigation et un droit d'accès à l'information pour remplir leur mandat.
149. Règlement sur l'IA, art. 80 (1)-(2).
150. Règlement sur l'IA, art. 80 (7).
151. HrFlow.ai, disponible à l'adresse https://hrflow.ai/parsing/, consulté le 12 novembre 2025.

pour certaines langues, ce qui entraîne des malentendus et a des effets préjudiciables pour les demandeurs d'asile[152]. Les organismes de promotion de l'égalité et les SNDH joueront donc un rôle déterminant en veillant à ce que les systèmes d'IA à haut risque ne contournent pas le règlement. Ce devrait être à eux d'élaborer les lignes directrices et les actes délégués à venir.

3.1.2. Rôle des organismes de promotion de l'égalité et des SNDH en matière de classification des systèmes d'IA à haut risque

- Évaluer les informations fournies dans les auto-évaluations par les fournisseurs qui ont exclu leur système du régime à haut risque, afin de vérifier si les informations fournies permettent une analyse correcte des incidences sur l'égalité et la non-discrimination.
- Dresser une liste des systèmes d'IA que les fournisseurs ont choisi d'exclure du régime à haut risque prévu à l'article 6, paragraphe 3, mais qui présentent un risque important. Ces exemples peuvent être tirés de la base de données de l'UE sur les systèmes à haut risque[153], des plaintes reçues par les OPE et des informations obtenues grâce à une collaboration avec les organisations de la société civile (voir Coopération).
- Publier cette liste et la communiquer directement aux ASM, aux APD et à la Commission européenne afin d'accroître la sensibilisation sur le sujet et de contribuer aux lignes directrices de la Commission européenne.
- Élaborer des orientations sur la manière dont les systèmes d'IA peuvent présenter un risque important pour l'égalité et la discrimination. Ces orientations peuvent être élaborées sur la base de la liste d'exemples susmentionnée.
- Les OPE devraient coopérer avec les ASM et les aider à déterminer si un système d'IA exclu du régime à haut risque présente ou non un risque significatif pour les droits fondamentaux.

3.2. Modifier la liste des cas d'utilisation à haut risque

3.2.1. Contexte et importance

L'article 7 confère à la Commission européenne le pouvoir d'adopter des actes délégués pour

1. ajouter ou modifier les cas d'utilisation des systèmes d'IA à haut risque répertoriés à l'annexe III[154], et
2. supprimer les systèmes d'IA à haut risque de la liste fournie à l'annexe III[155].

152. Bhuiyan, J. (2023), Lost in AI translation : growing reliance on language apps jeopardizes some asylum applications, The Guardian, disponible à l'adresse https://www.theguardian.com/us-news/2023/sep/07/asylum-seekers-ai-translation-apps, consulté le 10 novembre 2025.
153. La possibilité pour les OPE d'accéder à la partie non publique de la base de données reste à clarifier.
154. Règlement sur l'IA, art. 7 (1) premier alinéa.
155. Règlement sur l'IA, art. 7 (1) premier alinéa.

Pour comprendre la différence entre ces deux possibilités, il peut être utile de prendre un exemple issu de l'annexe III[156]:

> Emploi, gestion de main d'œuvre et accès à l'emploi indépendant :
>
> (a) Systèmes d'IA destinés à être utilisés pour le recrutement ou la sélection de personnes physiques, notamment pour diffuser des offres d'emploi ciblées, analyser et filtrer les candidatures et évaluer les candidat.es;
> (b) Systèmes d'IA destinés à être utilisés pour prendre des décisions affectant les conditions des relations professionnelles, la promotion ou la résiliation des relations professionnelles contractuelles, pour attribuer des tâches sur la base du comportement individuel ou de traits ou caractéristiques personnels, ou pour suivre et évaluer les performances et le comportement de personnes dans le cadre de ces relations.

Dans cet exemple, un système d'IA est considéré comme à haut risque dans le domaine « Emploi, gestion de main d'œuvre et accès à l'emploi indépendant » pour deux cas d'utilisation figurant aux alinéas (a) et (b).

Ajouter ou modifier des cas d'utilisation à haut risque

Par un acte délégué, la Commission peut modifier l'un des cas d'utilisation visés aux alinéas (a) et (b), ou ajouter de nouveaux cas d'utilisation en ajoutant des alinéas c), d), etc. Les systèmes d'IA présentent « un risque d'incidence négative sur les droits fondamentaux, et ce risque est équivalent ou supérieur au risque de préjudice ou d'incidence négative que présentent les systèmes d'IA à haut risque déjà visés à l'annexe III »[157]. Toutefois, la Commission ne peut pas ajouter un nouveau domaine aux huit domaines répertoriés comme systèmes d'IA à haut risque à l'annexe III.

Supprimer des cas d'utilisation à haut risque

Par un acte délégué, si la Commission estime que l'utilisation de systèmes d'IA dans le domaine de l'emploi, de la gestion de la main d'œuvre et de l'accès à l'emploi indépendant ne présente plus de risques importants et n'altère plus la protection des droits fondamentaux, de la santé ou de la sécurité[158], et que dès lors, les systèmes d'IA utilisés dans ce domaine et dans tous les cas d'utilisation peuvent donc être supprimés de l'annexe III.

Pour ajouter ou modifier des cas d'utilisation, ou supprimer des systèmes d'IA à haut risque de l'annexe III, la Commission doit prendre en compte de nombreux critères, parmi lesquels[159]:

- la mesure dans laquelle les personnes « ayant potentiellement subi un préjudice ou une incidence négative dépendent des résultats obtenus au moyen d'un système d'IA » en raison d'un manque de viabilité pratique ou de raisons juridiques qui font qu'il n'est « pas raisonnablement possible [...] de s'affranchir de ces résultats »;
- un déséquilibre des pouvoirs : « les personnes ayant potentiellement subi un préjudice ou une incidence négative se trouvent dans une situation vulnérable par rapport au déployeur d'un système d'IA, notamment en raison du statut,

156. Règlement sur l'IA, annexe III (4).
157. Règlement sur l'IA, art. 7 (1) (b).
158. Règlement sur l'IA, art. 7 (3) (a) et (b).
159. Règlement sur l'IA, art. 7 (2).

de l'autorité, de connaissances, de circonstances économiques ou sociales ou de l'âge »;

- la capacité de correction ou de réversibilité d'un résultat produit par un système d'IA, « les résultats qui ont une incidence négative sur la santé, la sécurité ou les droits fondamentaux ne devant pas être considérés comme facilement corrigibles ou réversibles »; et
- si la législation européenne existante (et non la législation nationale) prévoit « des mesures de réparation efficaces en ce qui concerne les risques posés par un système d'IA, à l'exclusion des réclamations en dommages-intérêts » et « des mesures efficaces destinées à prévenir ou à réduire substantiellement ces risques ».

Ces critères prouvent clairement que les droits fondamentaux, y compris la non-discrimination, sont au cœur de toute modification de l'annexe III.

3.2.2. Rôle des organismes de promotion de l'égalité et SNDH vis-à-vis de la modification de la liste des cas d'utilisation à haut risque

- Documenter et publier des rapports mettant en évidence les incidences négatives des systèmes d'IA sur les droits fondamentaux. Ces rapports doivent faire la distinction entre :

  1. les systèmes d'IA déjà considérés comme à haut risque selon l'annexe III, et
  2. les cas d'utilisation non encore répertoriés à l'annexe III, y compris ceux qui sont exemptés par l'article 6, paragraphe 3).

  Les données relatives au point (1) soulignent l'importance de maintenir les domaines des systèmes d'IA à haut risque dans l'annexe III et de ne pas les supprimer. Les données relatives au point (2), soulignent, quant à elles, la nécessité de modifier ou d'ajouter des cas d'utilisation à l'annexe III. S'agissant du point (2), les organismes compétents devraient accorder la priorité aux systèmes d'IA déployés dans des cas d'usage caractérisés par un déséquilibre de pouvoir entre les entités déployantes, telles que les autorités publiques, et les personnes concernées, lorsque celles-ci ne disposent pas d'une possibilité raisonnable de ne pas recourir au système (par exemple dans le domaine de la gestion des migrations ou des systèmes de sécurité sociale), et lorsque les effets des décisions portant atteinte aux droits fondamentaux ne peuvent être aisément corrigés ou inversés. Ces rapports peuvent se fonder sur des éléments recueillis par des organisations de la société civile, en particulier concernant des systèmes qui ne sont pas encore répertoriés à l'annexe III. Les organisations de la société civile peuvent également contribuer à identifier des domaines prioritaires de recherche et d'investigation.

- Envoyer ces rapports aux autorités nationales compétentes des États membres afin de les sensibiliser, ainsi qu'à la Commission européenne pour qu'ils soient pris en considération lorsque la Commission envisagera de rédiger et d'adopter des actes délégués.
- Mener ou commander des recherches qui mettent en évidence les lacunes du droit européen « pour prévenir ou réduire substantiellement[160]» les risques de discrimination dus au recours à des systèmes d'IA et pour proposer des mécanismes efficaces pour les personnes touchées en ce qui concerne les risques posés par un système d'IA, ce qui est l'un des critères que la Commission doit prendre en compte. Ces recherches pourraient être menées en collaboration avec différents organismes de promotion de l'égalité, via une éventuelle coordination assurée par le réseau européen des organismes de promotion de l'égalité (Equinet).

3.3. Exigences du système de gestion des risques

3.3.1. Contexte et pertinence

L'article 9 du règlement sur l'IA impose aux fournisseurs de systèmes d'IA à haut risque d'établir, de mettre en œuvre, de documenter et de tenir à jour un système de gestion des risques « sur l'ensemble du cycle de vie d'un système d'IA à haut risque et qui doit périodiquement faire l'objet d'un examen et d'une mise à jour méthodiques »[161]. Ce système doit comprendre :

- « l'identification et l'analyse des risques connus et raisonnablement prévisibles »[162] pour les droits fondamentaux, notamment toute « incidence négative sur des personnes âgées de moins de 18 ans et, le cas échéant, sur d'autres groupes vulnérables »[163], y compris l'évaluation des risques survenant après le déploiement[164].
- « l'adoption de mesures appropriées et ciblées de gestion des risques, conçues pour répondre aux risques identifiés »[165], de sorte que « le risque résiduel global lié aux systèmes d'IA à haut risque »[166] soit jugé acceptable. Ces mesures peuvent, entre autres, être les suivantes :
  - « éliminer ou réduire les risques identifiés et évalués [...] autant que la technologie le permet grâce à une conception et à un développement appropriés du système d'IA à haut risque »[167]

160. Règlement sur l'IA, art. 7 (2) (k) (ii).
161. Règlement sur l'IA, art. 9 (1).
162. Règlement sur l'IA, art. 9 (2) (a).
163. Règlement sur l'IA, art. 9 (9).
164. Règlement sur l'IA, art. 9 (2) (c).
165. Règlement sur l'IA, art. 9 (2) (d).
166. Règlement sur l'IA, art. 9 (5) premier alinéa.
167. Règlement sur l'IA, art. 9 (5) (a).

- « mettre en œuvre [...] des mesures adéquates d'atténuation et de contrôle répondant aux risques impossibles à éliminer »[168]
- transparence et information à l'égard des déployeurs[169].

Il est important de souligner que l'article 9 ne concerne que les risques « qui peuvent être raisonnablement atténués ou éliminés dans le cadre du développement ou de la conception du système d'IA à haut risque, ou par la fourniture d'informations techniques appropriées »[170]. 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 ». Dans l'ensemble, il est possible que les préjudices liés à la discrimination ne puissent pas être légitimement inclus dans les systèmes de gestion des risques visés à l'article 9 s'ils ne peuvent pas être atténués ou éliminés par des moyens techniques. Cela dit, les risques non atténués et persistants devraient être inclus dans la documentation technique visée à l'article 11 du règlement européen sur l'IA[171]. Cela est d'autant plus vrai qu'il peut ne pas exister de risque résiduel acceptable en matière de discrimination.

Corrélations avec d'autres cadres juridiques

L'article 16 de la Convention-cadre du Conseil de l'Europe prévoit également l'adoption et la mise à jour d'un cadre de gestion des risques et des impacts. À cette fin, le Conseil de l'Europe a mis au point la méthodologie d'évaluation des risques et des impacts du point de vue des droits des humains, de la démocratie et de l'État de droit (HUDERIA) pour 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 humains, le fonctionnement de la démocratie et le respect de l'État de droit »[172]. Elle propose 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 », par le biais d'une analyse des risques fondée sur le contexte, d'un processus d'engagement des parties prenantes, d'une évaluation des risques et des impacts et d'un plan d'atténuation, avec des évaluations à caractère itératif.

168. Règlement sur l'IA, article 9 (5) (b).
169. Règlement sur l'IA, art. 9 (5) (c).
170. Règlement sur l'IA, art. 9 (3).
171. Règlement européen sur l'IA, art. 11 et annexe IV, paragraphes 2 (g) et (3).
172. Comité du Conseil de l'Europe sur l'intelligence artificielle (2022), Contours de la méthodologie pour l'évaluation des risques et de l'impact Huderia, Strasbourg, disponible à l'adresse https://rm.coe.int/cai-bu-2022-03-fr-contours-de-la-methodologie-huderia/1680a81e15, consulté le 10 novembre 2025; et Comité du Conseil de l'Europe sur l'intelligence artificielle (2024), 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), Strasbourg, disponible à l'adresse https://rm.coe.int/cai-2024-16rev2-fr-methodologie-pour-l-evaluation-des-risques-et-des-i/1680b2a09e, consulté le 10 novembre 2025.

Toutefois, il ne s'agit pas d'un instrument juridiquement contraignant, ni d'un guide d'interprétation de la Convention-cadre, et il n'est pas obligatoire de se conformer à l'HUDERIA pour satisfaire aux obligations de la Convention.

3.3.2. Rôle des organismes de promotion de l'égalité et des SNDH en ce qui concerne les exigences du système de gestion des risques

- Collaborer avec des juristes pour déterminer si les préjudices liés à la discrimination entrent dans le champ d'application de l'article 9, compte tenu des références aux risques résiduels acceptables et aux risques qui peuvent être gérés par des moyens techniques. Cette collaboration sera utile jusqu'à ce qu'une interprétation juridiquement contraignante soit donnée par la CJUE.
- Fournir des éclairages provisoires sur la nécessité d'inclure le risque de discrimination dans l'analyse de risques de l'article 9. Ces éclairages pourraient établir la nécessité d'expliquer l'identification du risque, l'évaluation du risque et l'adoption de mesures adéquates de gestion des risques.
- Les éclairages provisoires devraient préciser que même si les risques de discrimination identifiés au niveau du système d'IA à haut risque ne peuvent être atténués, ils devraient être inclus dans la documentation technique.

3.4. Exigences liées à la gouvernance des données

3.4.1. Contexte et pertinence

Les données font partie intégrante du développement des systèmes d'IA. L'article 10 prévoit des obligations en matière de gouvernance des données pour les fournisseurs de systèmes d'IA à haut risque. Les fournisseurs pourraient concevoir leur propre approche, couvrant notamment la mise en œuvre technique, pour remplir ces obligations. Quelle que soit l'approche adoptée, le respect de ces obligations est nécessaire mais insuffisant pour limiter le risque de non-discrimination dû aux systèmes d'IA. En effet, les données ne sont que l'une des sources de biais et de discrimination dans les systèmes d'IA. Les systèmes d'IA peuvent être discriminatoires en raison des algorithmes et de leurs évaluations[173].

Les fournisseurs de systèmes d'IA à haut risque sont tenus d'examiner et de prendre « des mesures appropriées pour détecter, prévenir et atténuer les éventuels biais »[174] repérés dans les ensembles de données[175] qui sont susceptibles « d'avoir

173. Shrishak, K. (2025), Bias Evaluation, AI-Complex Algorithms and effective Data Protection Supervision, EDPB, disponible à l'adresse https://www.edpb.europa.eu/our-work-tools/our-documents/support-pool-experts-projects/ai-complex-algorithms-and-effective-data_en, consulté le 10 novembre 2025.
174. Règlement sur l'IA, art. 10 (2) (g).
175. Le règlement sur l'IA, article 3, paragraphes 29 à 33, définit les ensembles de données d'entraînement, de validation et de test. Les systèmes d'IA peuvent être développés à partir de ces trois ensembles de données ou d'un seul, en fonction de la technique utilisée. Dans tous les cas, sur la base de l'article 10, paragraphes 1 et 6, les obligations de l'article 10, paragraphes 2 à 5, s'appliquent.

une incidence négative sur les droits fondamentaux ou de se traduire par une discrimination interdite par le droit de l'Union »[176]. Pour respecter cette obligation, il est essentiel que les fournisseurs

- fassent des « choix de conception pertinents »[177];
- collectent et exploitent des données de haute qualité basées sur des hypothèses bien formulées[178];
- s'assurent que les ensembles de données sont « suffisamment représentatifs »[179] des « personnes ou groupes de personnes pour lesquels le système d'IA à haut risque est destiné à être utilisé »[180], ont les « caractéristiques ou éléments qui sont propres à l'environnement géographique, contextuel, comportemental ou fonctionnel spécifique »[181]; et
- effectuent des « opérations de traitement pertinentes pour la préparation des données, telles que l'annotation, l'étiquetage, le nettoyage »[182], etc.

Il est important de souligner que l'annotation et l'étiquetage, qui font partie intégrante du fonctionnement de nombreux systèmes d'IA, sont souvent confiés à des personnes anonymes qui ne font pas partie de la population représentative où le système d'IA est déployé[183]. Il n'est donc pas rare que des étiquettes soient incorrectes dans les ensembles de données. Dans certains cas, des insultes raciales et des expressions péjoratives sont intégrées au cours du processus d'étiquetage[184]. Ce problème peut également se poser dans les systèmes d'IA à usage général et lorsqu'ils sont déployés pour des applications spécifiques[185]. Par conséquent, l'annotation et l'étiquetage peuvent être une source de biais dans les données utilisées pour développer des systèmes d'IA.

Bien que l'élaboration des différentes approches techniques « pour détecter, prévenir et atténuer les biais possibles »[186] n'entre pas dans le cadre de ce travail[187], il est important de souligner que

176. Règlement sur l'IA, art. 10 (2) (f).
177. Règlement sur l'IA, art. 10 (2) (a).
178. Règlement sur l'IA, art. 10 (2) (d).
179. Règlement sur l'IA, art. 10 (3).
180. Ibid.
181. Règlement sur l'IA, art. 10 (4).
182. Règlement sur l'IA, art. 10 (2) (c).
183. C. G. Northcutt, et al (2021). Pervasive label errors in test sets destabilize machine learning benchmarks, disponible sur le site https://labelerrors.com, consulté le 10 novembre 2025.
184. Birhane, A., & Prabhu, V. U. (2021), « Large image datasets : A pyrrhic win for computer vision? » WACV, (2021 IEEE Winter Conference on Applications of Computer Vision), 1536-1546; Crawford, K. et Paglen, T. (2021), Excavating AI : The politics of images in machine learning training sets. AI & SOCIETY, 36, 1105-16.
185. Même si aucun étiquetage n'est effectué lors de l'entraînement initial des systèmes d'IA à usage général, certaines entreprises s'appuient sur un processus appelé apprentissage par renforcement et rétroaction humaine à un stade ultérieur du processus de développement, ce qui introduit des biais dans le système.
186. Règlement sur l'IA, art. 10 (2) (g).
187. Pour plus d'informations sur l'évaluation des biais, voir Shrishak, K. (2025), Bias Evaluation, AI-Complex Algorithms and effective Data Protection Supervision, EDPB.

- les systèmes d'IA sont sociotechniques et les approches purement techniques ne permettent pas d'éliminer totalement les biais[188]
- Il ne suffit pas de supprimer les variables sensibles des ensembles de données, car les données comportent généralement des variables de substitution qui peuvent elles aussi contribuer à créer des biais[189].

L'article 10(5) du règlement sur l'IA

De manière générale, le RGPD interdit le traitement d'une catégorie particulière de données à caractère personnel[190], sauf si une exception s'applique. L'article 10, paragraphe 5, du règlement sur l'IA met en œuvre une exception prévue par le RGPD lorsque « le traitement est nécessaire pour des motifs d'intérêt public important »[191]. L'intérêt important invoqué dans le règlement sur l'IA correspond à une stricte nécessité « aux fins de la détection et de la correction des biais »[192].

L'article 10, paragraphe 5, ne s'applique qu'aux fournisseurs de systèmes d'IA, et non aux déployeurs ou à tout autre tiers. Afin d'en comprendre les implications, les deux scénarios suivants peuvent être envisagés :

- Une autorité publique qui agit en tant que fournisseur et déployeur : Si une autorité publique développe (en tant que fournisseur) un système d'IA en interne et le déploie pour évaluer les prestations sociales, elle peut, en tant que fournisseur, se prévaloir de l'article 10, paragraphe 5.
- Une autorité publique qui agit uniquement en tant que déployeur : Si une entreprise (fournisseur) développe des systèmes d'IA destinés à évaluer les prestations sociales, et que l'autorité publique (déployeur) achète l'un de ces systèmes d'IA et l'utilise, seule l'entreprise peut se prévaloir de l'article 10, paragraphe 5. Le système d'IA développé par une entreprise peut être utilisé pour évaluer les prestations sociales dans plusieurs pays par différentes autorités publiques. Cependant, les autorités publiques, qui peuvent avoir davantage de connaissances géographiques et contextuelles, ne seraient pas en mesure de collecter et de traiter des données de catégories spéciales pour détecter et corriger les biais. Cela signifie que les autorités publiques peuvent être incitées à ne pas développer en interne des systèmes d'IA à haut risque, car cela leur confèrerait une plus grande responsabilité (en tant que fournisseur).

Les fournisseurs de systèmes d'IA ne peuvent s'appuyer sur l'article 10, paragraphe 5, pour traiter des données relevant de catégories particulières que lorsque cela est strictement nécessaire pour détecter et corriger des biais, et que d'autres données ne

188. Buyl, M., & De Bie, T. (2024), Inherent Limitations of AI Fairness, Communications of the ACM, 67(2), 48-55, disponible à l'adresse suivante https://doi.org/10.1145/3624700 schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., & Hall, P. (2022). Towards a standard for identifying and managing bias in artificial intelligence (NIST SP 1270).
189. Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. S. (2012), Fairness through awareness, ITCS, 214-226; Kamiran, F., & Calders, T. (2012), Data preprocessing techniques for classification without discrimination, Knowledge and Information Systems, 33(1), 1-33.
190. RGPD, Art. 9 (1) et règlement (UE) 2018/1725, art. 10 (1).
191. RGPD, Art. 9 (2) (g) et Règlement (UE) 2018/1725, Art. 10 (2) (g).
192. Règlement sur l'IA, art. 10 (5), premier alinéa.

relevant pas de catégories spéciales sont inefficaces[193]. L'approche des prestataires doit être « précisément encadrée par des dispositions garantissant qu'elle est effectivement limitée au strict nécessaire »[194]. Il pourrait s'agir d'une documentation montrant comment le traitement de données spécifiques à une catégorie particulière permet de détecter et de corriger des biais et de prévenir efficacement la discrimination.

En outre, le fournisseur ne peut pas utiliser cette catégorie particulière de données à caractère personnel à d'autres fins, doit conserver les données en toute sécurité, ne pas les partager avec d'autres parties ni leur permettre d'y accéder, et les supprimer "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 »[195]. Cela signifie également que ces données ne seront pas accessibles aux organismes de promotion de l'égalité et aux SNDH, tout au moins, sur la base de l'article 10, paragraphe 5.

Toutefois, il est important de souligner que les catégories particulières de données à caractère personnel du RGPD ne recoupent que quatre des caractéristiques protégées dans la législation sur la non-discrimination : (1) le handicap, (2) la religion ou les convictions, (3) l'origine raciale ou ethnique et (4) l'orientation sexuelle. Le RGPD n'interdit pas aux fournisseurs et aux déployeurs de collecter des informations sur les autres caractéristiques, dont l'âge et le sexe.

En outre, les fournisseurs peuvent déjà traiter des données de catégorie particulière si « la personne concernée a donné son consentement explicite »[196]. L'article 10, paragraphe 5, du règlement sur l'IA n'est utile que lorsque le fournisseur ne peut pas obtenir le consentement explicite de la personne concernée.

Corrélations avec d'autres cadres juridiques

Bien que l'article 10, paragraphe 5, du règlement sur l'IA ne s'applique qu'aux fournisseurs, les organismes de promotion de l'égalité pourraient s'appuyer sur l'article 21 des directives sur les normes[197], qui autorise les organismes de promotion de l'égalité à traiter des catégories particulières de données à caractère personnel. Pour que les organismes de promotion de l'égalité puissent se saisir de cette possibilité, il est essentiel que leur État membre transpose la directive sur les normes dans leur législation nationale avec la mise en œuvre d'une exception au titre de l'article 9, paragraphe 2, point g), du RGPD, à condition que cette exception ne soit pas déjà garantie dans la législation nationale en matière de protection des données ou dans une autre législation. Alors que la directive sur les normes autorise les organismes de promotion de l'égalité à traiter des catégories particulières de données à caractère personnel, ni le règlement sur l'IA ni la directive sur les normes n'autorisent explicitement les organismes de promotion de l'égalité à collecter ce type de données auprès de fournisseurs d'IA.

193. Règlement sur l'IA, art. 10 (2) (a) et (f).
194. Arrêt de la Cour de justice du 8 avril 2014, Digital Rights Ireland Ltd, C-293/12 et C-594/12, ECLI:EU:C:2014:238, point 65.
195. Règlement sur l'IA, art. 10 (5) (e).
196. RGPD, Art. 9 (2) (a) et Règlement (UE) 2018/1725, Art. 10 (2) (a).
197. Directives sur les normes, art. 21 et considérant 48.

3.4.2. Rôle des organismes de promotion de l'égalité et des SNDH en ce qui concerne les exigences liées à la gouvernance des données

- Être conscient.es des limites des approches techniques de détection et de correction des biais dans le cadre de l'évaluation des risques de discrimination associés aux systèmes d'IA à haut risque. Le respect du règlement sur l'IA, et notamment de son article 10, peut ne pas suffire à remédier aux risques de discrimination associé aux systèmes d'IA à haut risque. Dans ces scénarios, lorsque les OPE sont consultés par une autorité de surveillance du marché exerçant ses pouvoirs au titre de l'article 82, paragraphe 1, les OPE pourraient demander à l'autorité de surveillance du marché d'obliger les entreprises à « prendre toutes les mesures appropriées pour faire en sorte que le système d'IA [...] ne présente plus ce risque »[198].
- Assurer la transposition en droit national de l'article 21 des directives européennes sur les normes, qui autorise les organismes de promotion de l'égalité à traiter des catégories particulières de données à caractère personnel si cela n'est pas déjà permis par d'autres lois nationales.

3.5. L'analyse d'impact sur les droits fondamentaux (AIDF)

3.5.1. Contexte et pertinence

Les organismes de promotion de l'égalité et les SNDH devraient considérer l'AIDF comme le point de départ de leur évaluation des systèmes d'IA déployés par les administrations publiques ou en leur nom. L'analyse d'impact sur les droits fondamentaux (AIDF) doit inclure une série d'informations qui mettent en évidence les lacunes et les atteintes potentielles aux droits, y compris au droit à la non-discrimination. L'AIDF devrait être effectuée par le déployeur (et non par le fournisseur) d'un système d'IA à haut risque, car on suppose que le déployeur dispose d'informations contextuelles pertinentes pour cette évaluation.

L'AIDF doit inclure des informations sur la manière dont le déployeur utilisera le système d'IA à haut risque selon l'usage prévu par le fournisseur, mais également sur la durée et la fréquence d'utilisation du système d'IA et sur la manière dont les mesures de surveillance humaine ont été mises en œuvre[199]. Dans l'idéal, le déployeur devrait expliquer en des termes compréhensibles le fonctionnement du système d'IA sur la base de la documentation reçue du fournisseur et du lien avec le système d'IA à haut risque enregistré dans la base de données de l'UE. Les autorités publiques ne doivent pas utiliser de systèmes d'IA à haut risque que les fournisseurs n'auraient pas enregistrés dans la base de données de l'UE[200].

198. Règlement sur l'IA, art. 82 (1).
199. Règlement sur l'IA, art. 27 (1) (a) et (e).
200. Règlement sur l'IA, art. 26 (8). Cette disposition s'applique également aux institutions, organes et organismes de l'Union européenne.

En outre, l'AIDF doit décrire[201] :

- « les catégories de personnes physiques et de groupes susceptibles d'être concernés »;
- « les risques spécifiques de préjudices » portant atteinte à leurs droits fondamentaux, y compris à la non-discrimination, en tenant compte des informations fournies par le prestataire;
- les mesures « y compris les dispositifs relatifs à la gouvernance interne et aux mécanismes de plainte internes » que le déployeur entend prendre si ces risques se concrétisent.

Lorsque des données à caractère personnel sont traitées, l'AIDF complète les analyses d'impact sur la protection des données (AIPD) réalisées en vertu de la législation européenne sur la protection des données[202]. On peut s'attendre à ce que les déployeurs combinent AIDF et AIPD.

Le considérant 96 du règlement sur l'IA stipule que « 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 » (italique ajouté).

Corrélations avec d'autres cadres juridiques

Le règlement sur la protection des données prévoit la réalisation d'analyses d'impact des systèmes d'IA. La législation européenne sur la protection des données prévoit des analyses d'impact sur la protection des données (AIPD)[203]. La Convention modernisée du Conseil de l'Europe pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel (Convention 108+) prévoit une obligation générale d'analyser l'impact probable du traitement des données sur les droits et les libertés fondamentales des personnes avant leur utilisation[204].

Les dispositions relatives aux analyses d'impact sur les droits fondamentaux du règlement sur l'IA peuvent être liées aux obligations d'évaluation des risques prévues par la Convention-cadre du Conseil de l'Europe et la méthodologie HUDERIA. La Convention-cadre du Conseil de l'Europe exige que chacune des parties « adopte ou [maintienne] des mesures afin d'identifier, d'évaluer, de prévenir et d'atténuer les risques posés par les systèmes d'intelligence artificielle en tenant compte des impacts réels et potentiels sur les droits de l'homme, la démocratie et l'État de

201. Règlement sur l'IA, art. 27 (1) (c), (d) et (f).
202. Règlement sur l'IA, art. 27 (4). Voir également RGPD, article 35 et directive Police-Justice, article. 27. L'article 35 du RGPD stipule qu'une EIDP doit être réalisée lorsque le traitement de données à caractère personnel est « susceptible d'engendrer un risque élevé pour les droits et libertés d'une personne physique », en particulier lorsqu'il s'agit d'un « traitement automatisé, y compris le profilage ».
203. RGPD, Art. 35 et directive Police-Justice, art. 27.
204. Convention modernisée du Conseil de l'Europe pour la protection des personnes à l'égard du traitement des données à caractère personnel, article 10, (2).

droit[205] ». L'article 16, paragraphe 2, point f), de la Convention impose « la documentation des risques, des impacts réels et potentiels, et de l'approche de la gestion des risques ». En outre, ces mesures d'évaluation des risques doivent prendre « en compte, le cas échéant, le point de vue des parties prenantes pertinentes, en particulier les personnes dont les droits pourraient être affectés[206] » et s'appliquer « de manière itérative tout au long des activités menées dans le cadre du cycle de vie du système d'intelligence artificielle »[207].

Limites des AIDF

Champ d'application limité : Malgré l'importance potentielle de l'AIDF, l'obligation d'effectuer une AIDF avant le déploiement ne s'applique qu'à un sous-ensemble de déployeurs de systèmes d'IA à haut risque :

1. Pour tous les systèmes d'IA à haut risque, sauf dans le domaine des infrastructures critiques[208], l'obligation concerne les déployeurs qui sont :
   - des organismes de droit public, ou
   - des entités privées fournissant des services publics[209].
2. Pour les systèmes d'IA à haut risque, dans deux cas d'utilisation (évaluation de la solvabilité ou attribution d'un score de crédit à des personnes physiques[210], et évaluation des risques et tarification des assurances vie et des assurances maladie[211]), cette obligation s'applique à tous les déployeurs.

En pratique, cela signifie que les systèmes utilisés dans le secteur de l'emploi par des entités privées (qui, dans ce cas, seraient des déployeurs) sont exemptés de l'obligation de mener une AIDF.

En outre, les déployeurs pourraient ne pas effectuer d'AIDF si le prestataire a choisi de ne considérer son système comme étant à haut risque. Il est également possible que le déployeur puisse « dans des cas similaires, s'appuyer sur des analyses d'impact sur les droits fondamentaux effectuées précédemment ou sur des analyses d'impact existantes réalisées par le fournisseur[212]». Toutefois, cela ne doit pas être interprété comme libérant les déployeurs de leur obligation de mener une AIDF, mais uniquement comme la possibilité de s'appuyer sur les AIDF antérieures au lieu de repartir de zéro.

Manque d'information du public

Les autorités publiques ou les personnes agissant en leur nom qui déploient des systèmes d'IA à haut risque « s'enregistrent, sélectionnent le système et enregistrent

205. Convention-cadre du Conseil de l'Europe, art. 16 (1).
206. Convention-cadre du Conseil de l'Europe, art. 16 (2) (c).
207. Convention-cadre du Conseil de l'Europe, art. 16 (2) (d).
208. Règlement sur l'IA, annexe III (2).
209. Règlement sur l'IA, art. 27 (1), premier alinéa.
210. Règlement sur l'IA, annexe III (5) (b).
211. Règlement sur l'IA, annexe III (5) (c).
212. Règlement sur l'IA, art. 27 (2). Le fournisseur peut également être déployeur du système d'IA qu'il a développé. Dans ce cas, il est possible que le fournisseur ait effectué une AIDF.

son utilisation dans la base de données de l'UE »[213] et doivent inclure un résumé des conclusions de l'AIDF [214].

Toutefois, « dans les domaines des actions répressives, de la migration, de l'asile et de la gestion des contrôles aux frontières[215] »

- La base de données de l'UE sera non publique et son accès sera limité à la Commission européenne et aux autorités chargées de la protection des données[216].
- Les résultats de l'AIDF ne seront pas inclus dans la base de données de l'UE[217].

En outre, même si, dans les cas d'utilisation d'une évaluation de la solvabilité ou du score de crédit de personnes physiques[218], et de l'évaluation des risques et la tarification des assurances vie et des assurances maladie[219], tous les déployeurs ont l'obligation de réaliser une AIDF (voir ci-dessous), seuls « les déployeurs qui sont des autorités publiques, des institutions, organes ou organismes de l'Union ou des personnes agissant en leur nom »[220] ont l'obligation d'enregistrer l'utilisation d'un système d'IA dans la base de données. Ainsi, les sociétés privées qui déploient des systèmes d'IA correspondant à ces cas d'utilisation n'ont pas l'obligation de s'enregistrer, et la synthèse de leur AIDF ne sera pas accessible.

Manque d'efficacité dans l'évaluation des risques de non-discrimination

Bien que les déployeurs doivent notifier les résultats de l'AIDF à l'ASM, l'analyse elle-même pourrait se limiter à cocher quelques cases sans véritablement consister en une analyse d'impact significative, puisque le déployeur pourrait remplir un « modèle de questionnaire » élaboré par le Bureau de l'IA[221] ultérieurement. L'un des risques est que les AIDF ne fournissent pas d'informations suffisamment précises pour permettre d'évaluer comment la discrimination a été évitée et/ou atténuée, quels motifs de discrimination ont été testés et comment les groupes de personnes protégés par la législation ont été définis au cours des procédures d'analyse des risques.

213. Règlement sur l'IA, art. 49 (3).
214. Annexe VIII, section C (4).
215. Règlement sur l'IA, art. 49 (4) premier alinéa.
216. L'article 74, paragraphe 8, du règlement sur l'IA dispose que l'autorité nationale doit être l'autorité compétente au titre du RGPD ou de la directive Police-Justice, ou « toute autre autorité désignée en application des mêmes conditions que celles prévues aux articles 41 à 44 » de la directive Police-Justice.
217. Règlement sur l'IA, art. 49 (4) (c).
218. Règlement sur l'IA, annexe III (5) (b).
219. Règlement sur l'IA, annexe (5) (c).
220. Règlement sur l'IA, article 49 (3).
221. Règlement sur l'IA, art. 27 (3) et (5).

3.5.2. Rôle des organismes de promotion de l'égalité et des SNDH vis-à-vis des obligations d'analyse d'impact sur les droits fondamentaux (AIDF)

- **Se servir des synthèses des AIDF de la base de données de l'UE comme point de départ pour étudier les systèmes d'IA.** Les OPE doivent cartographier le recours à des systèmes d'IA à haut risque d'après la base de données de l'UE et analyser la synthèse de l'AIDF. Ensuite, les OPE doivent demander à l'ASM de leur donner accès aux résultats de l'AIDF. Sur la base de ces informations, les OPE peuvent procéder à une première évaluation des violations de la législation en matière de non-discrimination et déterminer s'il y a lieu de mener une enquête approfondie, qui nécessiterait la réalisation d'une évaluation après le déploiement (ce que l'AIDF ne propose pas).
- **Contribuer à l'élaboration du modèle de questionnaire AIDF.** Le Bureau de l'IA est tenu d'élaborer un modèle de questionnaire AIDF. Il sera important que ce modèle pose des questions appropriées et spécifiques sur les différentes dimensions de l'égalité et de la non-discrimination afin que les informations recueillies et leur analyse permettent une évaluation et des garanties adéquates, plutôt que de devenir un simple exercice consistant à cocher quelques cases. Le Bureau de l'IA n'a pas de délai pour créer ce modèle. Les organismes de promotion de l'égalité et les SNDH pourraient contribuer de manière proactive à cette mise en œuvre en élaborant des orientations spécifiques. Ces orientations pourraient s'inspirer de la méthodologie HUDERIA mise au point par le Conseil de l'Europe.
  - Élaborer des orientations sur :
  - la teneur de synthèses significatives d'AIDF; et
  - qui devrait mener les AIDF. Les organismes de promotion de l'égalité et les SNDH devraient souligner l'importance et la nécessité que les organismes chargées des AIDF disposent d'une expertise en matière de droits fondamentaux.
- **Publier des études de cas** pour mettre en valeur l'importance d'analyser « les risques spécifiques de préjudice » dus à la discrimination dans le cadre des AIDF. En particulier, ces études devraient être axées sur des exemples spécifiques et aborder les questions relatives à la gravité et à l'ampleur des préjudices causés. Ces études pourraient être réalisées en collaboration avec une autorité publique qui le souhaite.

3.6. La base de données de l'UE relative aux systèmes d'IA à haut risque répertoriés à l'annexe III

3.6.1. Contexte et importance

L'article 71, paragraphe 1, du règlement sur l'IA impose à la Commission, « en collaboration avec les États membres » de créer et de tenir à jour « une base de données

de l'UE contenant [certaines] informations ». Cette base de données devrait recenser des informations sur :

1. les systèmes d'IA à haut risque répertoriés à l'annexe III;
2. les systèmes d'IA qui ont été autoévalués par les fournisseurs comme n'étant pas à haut risque au titre de l'article 6, paragraphe 3.

Les informations contenues dans la base de données doivent être « consultables grâce à une navigation aisée et lisibles par machine » et, pour la partie publique de la base de données, « accessibles et mises à la disposition du public d'une manière conviviale »[222]. Une interprétation large de l'accessibilité signifierait que les informations devraient être accessibles aux personnes lectrices porteuses d'un handicap et, conformément au considérant 72 (qui traite des exigences relatives aux notices d'utilisation préparées par les fournisseurs d'IA pour les déployeurs d'IA), qu'elles devraient être significatives et compréhensibles pour différents publics[223].

Une opportunité pour la transparence

Les obligations d'enregistrement et la base de données de l'UE recensant les systèmes d'IA à haut risque seront particulièrement utiles aux organismes de promotion de l'égalité et aux SNDH pour bénéficier d'une visibilité sur les systèmes d'IA utilisés par les organismes publics, ce qui n'est pas le cas actuellement (Xenidis, 2025). De nombreuses parties prenantes s'accordent à dire que ces bases de données constituent une étape nécessaire mais insuffisante pour garantir que les systèmes d'IA respectent les droits fondamentaux, et en particulier le droit à la non-discrimination[224].

La base de données de l'UE sera utile aux organismes de promotion de l'égalité et aux SNDH pour surveiller la mise sur le marché ou la mise en service de systèmes d'IA, aux organisations de la société civile et aux journalistes travaillant sur ces questions, ainsi qu'aux personnes potentiellement concernées pour établir des liens entre leur situation et le recours potentiel à des systèmes d'IA[225].

Types de systèmes à enregistrer

On dénombre quatre cas dans lesquels les systèmes d'IA doivent être enregistrés dans la base de données de l'UE :

---

222. Règlement sur l'IA, art. 71 (4).
223. Equinet 2025 souligne que le considérant porte sur les exigences de transparence des fournisseurs d'IA à l'égard des déployeurs l'IA (règlement sur l'IA, article 13) et que « cette définition peut donc ne pas s'appliquer dans ce contexte ». Toutefois, l'auteur souligne qu'il s'agit de « la seule explication de ce que signifie 'accessible' s'agissant de la documentation du règlement sur l'AI ».
224. IA Ciudadana (2025), Making Algorithm Registers Work for Meaningful Transparency, disponible à l'adresse https://iaciudadana.org/2025/03/13/making-algorithm-registers-work-for-meaningful-transparency/; Ada Lovelace Institute (2020), Meaningful transparency and (in)visible algorithms : Can transparency bring accountability to public-sector algorithmic decision-making (ADM) systems?, disponible sur https://www.adalovelaceinstitute.org/blog/meaningful-transparency-and-invisible-algorithms/; et Cath, C. et Jansen, F. (2022), Dutch Comfort: The Limits of AI Governance through Municipal Registers, Techné Research in Philosophy and Technology, 26(3), pp. 395-412, disponible à l'adresse https://doi.org/10.5840/techne202323172, consultés le 10 novembre 2025.
225. IA Ciudadana (2025), Making Algorithm Registers Work for Meaningful Transparency, disponible à l'adresse https://iaciudadana.org/2025/03/13/ making-algorithm-registers-work-for-meaningful-transparency

Entité soumise à l'obligation d'enregistrement	*Aperçu des informations à enregistrer*[226]
1. Enregistrement par les **fournisseurs de systèmes d'IA à haut risque** (ou, le cas échéant, les mandataires de ces fournisseurs) liés aux domaines énumérés à l'annexe III, à l'exception des systèmes d'IA à haut risque visés au point 2 de l'annexe III (« infrastructures critiques »), qui seront enregistrés au niveau national[227], avant la mise sur le marché ou la mise en service du système[228].	Nom, adresse et coordonnées du fournisseur; Objectif du système et de ses composants; Informations techniques de base et concises; Statut du système (par exemple, si le système est en cours d'utilisation ou s'il a été abandonné); Notice d'utilisation en format électronique communiquée par le fournisseur aux déployeurs conformément à l'article 13, paragraphe 2 (y compris les « caractéristiques, les capacités et les limites du système, telles que son niveau d'exactitude », et les mesures de contrôle humain)[229].
2. Par les **fournisseurs de systèmes d'IA** (ou, le cas échéant, leur mandataire) **dont ils ont conclu qu'ils n'étaient pas à haut risque bien qu'ils relèvent des domaines énumérés à l'annexe III** (en vertu de l'article 6, paragraphe 3), avant la mise sur le marché ou la mise en service du système[230].	Informations similaires aux systèmes d'IA à haut risque, avec toutefois des exigences réduites. Par exemple, la « notice d'utilisation en format électronique » n'est pas requise. Les informations comprennent également un bref résumé des raisons pour lesquelles le système d'IA est considéré comme n'étant pas à haut risque en application de la procédure prévue à l'article 6, paragraphe 3;

226. Les annexes VIII et IX du règlement sur l'IA en recensent la liste exhaustive.
227. Règlement sur l'IA, article 49 (5).
228. Règlement sur l'IA, art. 49 (1).
229. Règlement sur l'IA, art. 13 (2).
230. Règlement sur l'IA, art. 49 (2).

Entité soumise à l'obligation d'enregistrement	*Aperçu des informations à enregistrer*[226]
3. Par les **déployeurs de systèmes d'IA à haut risque relevant des domaines énumérés à l'annexe III** (à l'exception des systèmes d'IA à haut risque dans le domaine des infrastructures critiques) qui sont **« des autorités publiques, des institutions, organes ou des organismes de l'Union ou des personnes agissant en leur nom »** avant la mise en service ou l'utilisation du système[231].	Nom, adresse et coordonnées du déployeur; URL de l'entrée du système d'IA dans la base de données de l'UE par le fournisseur; Un résumé des conclusions de l'analyse d'impact sur les droits fondamentaux qui doit être réalisée conformément à l'article 27; Un résumé de l'AIPD réalisée conformément à la réglementation applicable en matière de protection des données.
4. Par les **fournisseurs** de systèmes d'IA à haut risque visés à l'annexe III, qui effectuent des **essais en conditions réelles** en dehors des bacs à sable réglementaires de l'IA[232].	Le nom et les coordonnées du fournisseur ou du fournisseur potentiel et des déployeurs impliqués dans les essais en conditions réelles; Une brève description du système d'IA, de son objectif et des autres informations nécessaires à l'identification du système; Un résumé des principales caractéristiques du plan de test en conditions réelles; Des informations sur la suspension ou l'arrêt des essais en conditions réelles réelles.

Les déployeurs de systèmes d'IA qui ne sont pas soumis à des obligations « devraient être autorisés » à enregistrer volontairement leur système dans la base de données, y compris lorsque ces déployeurs sont des entités privés[233].

Les autorités publiques qui ont la fonction de déployeurs doivent vérifier si le fournisseur a enregistré le système à haut risque dans la base de données. Si le système n'est pas enregistré, ils ne doivent pas l'utiliser et en informer le fournisseur ou le distributeur[234].

231. Règlement sur l'IA, art. 49 (3). Voir également le règlement sur l'IA, annexe III (1), (6) et (7).
232. Règlement sur l'IA, art. 60 (4) (c).
233. Règlement sur l'IA, considérant 131 (italique ajouté)
234. Règlement sur l'IA, art. 26 (8).

Versions publique et non publique de la base de données

La base de données comprendra deux sections : une section publique et une version « non publique sécurisée »[235]. La version non publique de la base de données contiendra des informations sur :

- Les systèmes d'IA à haut risque utilisés dans les domaines de la biométrie, des activités répressives et de la gestion des migrations, de l'asile et des contrôles aux frontières[236];
- Les systèmes d'IA à haut risque qui sont testés dans des conditions réelles conformément à l'article 60 (à moins que le fournisseur n'ait consenti à rendre ces informations accessibles au public)[237].

Calendrier de mise en place et de déploiement

- La Commission est chargée d'établir les spécifications fonctionnelles de la base de données en consultation avec les « experts en la matière », et la mise à jour des spécifications fonctionnelles de la base de données sera effectuée par la Commission, en consultation avec le Comité IA[238].
- Au moment de la rédaction du présent document, la base de données n'a pas encore été mise en place par la Commission européenne.
- L'enregistrement dans la base de données deviendra obligatoire le 2 août 2026. Toutefois, l'enregistrement volontaire avant le 2 août 2026 est également encouragé[239].

Corrélations avec d'autres cadres juridiques

Les obligations d'enregistrement prévues par le règlement sur l'IA peuvent être liées à l'article 8 de la Convention-cadre du Conseil de l'Europe, qui prévoit que les parties adoptent ou maintiennent des mesures visant à garantir « des exigences de transparence et de contrôle adaptées », et à l'article 9, qui porte sur les mesures visant à «'obligation de rendre des comptes et d'assumer la responsabilité pour les impacts négatifs sur les droits humains, la démocratie et l'État de droit ». L'article 14 de la Convention-cadre exige également que les Parties mettent en place « des mesures garantissant que des informations pertinentes concernant les systèmes d'intelligence artificielle susceptibles d'avoir une incidence significative sur les droits de l'homme 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 ».

Le Rapport explicatif de la Convention-cadre convient que les parties sont tenues de « trouver un juste équilibre entre [la transparence et] les divers intérêts concurrents »[240], parmi lesquels il cite la vie privée, la confidentialité (y compris, par

235. Règlement sur l'IA, art. 49 (4) et règlement sur l'IA, art. 71 (4).
236. Règlement sur l'IA, art. 49 (4)
237. Règlement sur l'IA, art. 60 (4) (c).
238. Règlement sur l'IA, art. 71 (1).
239. Règlement sur l'IA, considérant 179.
240. Conseil de l'Europe (2024), Rapport explicatif à 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, paragraphe 62

exemple, les secrets commerciaux), la sécurité nationale, la protection des droits des tiers, l'ordre public, l'indépendance judiciaire. Toutefois, le terme « adaptées » suggère qu'un niveau minimum de transparence, en particulier vis-à-vis du grand public, devrait être envisagé, même vis-à-vis des systèmes d'IA pour lesquels certaines informations doivent rester privées.

Difficultés

La base de données présente plusieurs limites qui peuvent entraver son utilité pour identifier et atténuer la discrimination dans les systèmes d'IA.

- **Certaines informations se trouveront dans la partie non publique de la base de données :**
  - Informations sur les systèmes d'IA utilisés dans les domaines de la biométrie, des activités répressives et de la migration, du contrôle des frontières et de l'asile.
  - Informations sur les systèmes d'intelligence artificielle testés en conditions réelles.
- **Certaines informations ne seront pas du tout enregistrées dans la base de données :**
  - La base de données ne concerne que les systèmes d'IA à haut risque. Cela pourrait exclure des applications telles que les chatbots utilisés par des fonctionnaires ou interagissant directement avec les citoyens et citoyennes.
  - Les fournisseurs et les déployeurs de systèmes d'IA à haut risque dans les domaines de la biométrie, des activités répressives et de la migration, de l'asile et du contrôle des frontières doivent enregistrer moins d'informations sur leurs systèmes d'IA. Par exemple, les fournisseurs ne sont pas tenus d'enregistrer la notice d'utilisation en format électronique fournie aux déployeurs. Les déployeurs ne sont pas tenus d'enregistrer une synthèse de leur AIDF.
  - Lorsque les fournisseurs de systèmes d'IA ne les considèrent pas comme étant à haut risque au sens de l'article 6, paragraphe 3, leurs déployeurs ne sont pas tenus d'enregistrer leur mise en service ou la réalisation de tests sur leurs systèmes d'IA dans la base de données.
  - Seuls les déployeurs qui sont « des autorités publiques, des institutions, organes ou organismes de l'Union ou des personnes agissant en leur nom »[241] ont l'obligation de s'enregistrer. En revanche, les déployeurs de certains systèmes d'IA à haut risque peuvent être des entités privées, telles que des compagnies d'assurance : ces déployeurs n'ont pas l'obligation d'enregistrer leur système d'IA dans la base de données de l'UE. Ainsi, certaines informations, telles que les résumés d'AIDF, ne seront pas accessibles concernant ces systèmes d'IA à haut risque.

241. Règlement sur l'IA, art. 49 (3).

- Même pour les systèmes d'IA à haut risque enregistrés dans la version publique de la base de données, il n'est pas garanti que les informations requises soient suffisantes pour évaluer la discrimination. Par exemple, les déployeurs sont tenus de soumettre uniquement une synthèse des AIDF, sans avoir à les communiquer dans leur intégralité.
- L'enregistrement dans la base de données ne sera obligatoire qu'à partir du 2 août 2026. Cela signifie que les systèmes d'IA mis en service avant cette date n'auront pas obligatoirement à être enregistrés, sauf en cas de « modification importante »[242].

Accès des organismes de promotion de l'égalité et des SNDH aux informations figurant dans la partie non publique de la base de données

Les systèmes d'IA utilisés dans le cadre d'activités répressives et dans le domaine de la migration, de l'asile et des contrôles aux frontières ont des effets discriminatoires sur certaines personnes (cf. 7.1 : Thématique centrale : Activités répressives, migration, asile et contrôle des frontières) et l'accès à la documentation contenue dans la version non publique de la base de données est nécessaire pour permettre aux organismes de promotion de l'égalité et aux SNDH de remplir efficacement leur mandat. Sans cet accès, les OPE et les SNDH ne pourront pas avoir une vue d'ensemble des systèmes d'IA à haut risque utilisés dans ces domaines. Toutefois, la mesure dans laquelle les OPE et les SNDH peuvent accéder aux informations contenues dans la version non publique de la base de données n'est pas claire.

L'article 71, paragraphe 4, dispose que « les informations enregistrées conformément à l'article 60 [concernant les systèmes d'IA à haut risque ne relevant pas de bacs à sables règlementaires] 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 ». L'article 49, paragraphe 4, sur l'enregistrement prévoit également que « [s]eules la Commission et les autorités nationales visées à l'article 74, paragraphe 8, ont accès » aux informations relatives aux systèmes d'IA à haut risque utilisés dans les domaines de la biométrie, des activités répressives et de la migration, de l'asile et des contrôles aux frontières.

L'article 77, paragraphe 1, prévoit que « [l]es 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. » (voir l'article 77).

Si des bases juridiques pour le partage d'informations entre les organismes de promotion de l'égalité/SNDH, les autorités de surveillance du marché et d'autres

242. Règlement sur l'IA, considérant 177.

organismes relevant de l'article 77 sont établies (voir Coopération), les OPE et les SNDH pourraient alors accéder aux informations obtenues par les autorités de surveillance du marché en vertu de l'article 71, paragraphe 4, et de l'article 49, paragraphe 4.

3.6.2. Rôle des organismes de promotion de l'égalité et des SNDH concernant la base de données de l'UE relative aux systèmes d'IA à haut risque répertoriés à l'annexe III

Lors de la mise en place de la base de données

- Participer en tant qu'expert.es compétent.es[243] (éventuellement via Equinet) à la mise en place de la base de données, notamment pour s'assurer que la base de données répond aux exigences en matière d'accessibilité.

Une fois la base de données en place

- Plaider en faveur d'un remplissage volontaire de la base de données par les administrations publiques lorsqu'elles déploient des systèmes d'IA qui ne sont pas considérés comme à haut risque, ainsi que par les déployeurs privés de systèmes d'IA à haut risque (en particulier dans les secteurs de l'assurance et de la banque).
- Élaborer un modèle de coordination avec les autorités nationales compétentes qui auront accès à la version non publique de la base de données.
- Plaider pour que les lois nationales mettant en œuvre le règlement sur l'IA indiquent clairement que les OPE et les SNDH devraient avoir accès aux informations transmises aux autorités nationales et à la base de données non publique de l'UE.
- Contrôler les systèmes d'IA enregistrés dans la base de données afin d'identifier les systèmes d'IA à haut risque qui devraient faire l'objet d'une enquête plus approfondie en raison de risques de discrimination, soit par les organismes de promotion de l'égalité, soit par les autorités de surveillance du marché (éventuellement en collaboration avec d'autres parties prenantes telles que des organisations de la société civile).
- Contrôler les systèmes d'IA enregistrés qui n'ont pas été considérés comme présentant un risque élevé au titre de l'article 6, paragraphe 3, et consulter la synthèse des auto-évaluations pour déterminer s'ils doivent être considérés comme à haut risque. Si c'est le cas, remonter l'information à l'ASM et demander la réalisation d'une évaluation au titre de la procédure de l'article 80 concernant la gestion des systèmes d'IA classés comme non à haut risque par leur fournisseur.

243. Règlement sur l'IA, art. 71 (1).

- Contrôler les systèmes d'IA liés à l'évaluation du crédit et à l'assurance enregistrés par les fournisseurs, afin de demander aux déployeurs des informations qui ne figurent pas dans la base de données, telles que les AIDF.
- Utiliser la version publique de la base de données comme outil de sensibilisation à l'utilisation des systèmes d'IA dans le secteur public.
- Contrôler l'utilisation de la base de données par les organisations de la société civile, les journalistes, les particuliers et les autres institutions concernées, et proposer des mises à jour à la Commission et au Comité IA. Il convient notamment de vérifier si les informations contenues dans les synthèses d'AIDF sont suffisantes pour évaluer les risques de discrimination.

4. Transparence des exigences liées aux systèmes d'IA

4.1. Contexte et importance

L'article 50 prévoit des obligations de transparence pour les fournisseurs et les déployeurs de systèmes d'IA, qui doivent être communiquées clairement aux personnes physiques « au plus tard au moment de la première interaction ou de la première exposition »[244] et être conformes aux exigences en matière d'accessibilité.

Les déployeurs doivent informer les personnes physiques dans les cas suivants :

- Déploiement d'un système de reconnaissance des émotions[245] ou d'un système de catégorisation biométrique[246];

244. Règlement européen sur l'IA, art. 50 (5).
245. Règlement européen sur l'IA, art. 3 (39).
246. Règlement européen sur l'IA, art. 3 (40).

- Déploiement d'un système d'IA qui génère des deep fakes (ou hypertrucages)[247];
- Déploiement d'« un système d'IA qui génère ou manipule des textes publiés dans le but d'informer le public sur des questions d'intérêt public »[248].

Les fournisseurs devraient concevoir et développer les systèmes d'IA qui interagissent avec des personnes physiques de manière à ce que ces dernières soient conscientes de la nature de cette interaction, sans penser qu'elles interagissent avec un être humain. Cette obligation s'applique « sauf si cela ressort clairement du point de vue d'une personne physique normalement informée et raisonnablement attentive et avisée, compte tenu des circonstances et du contexte d'utilisation[249] ».

Pour les contenus générés par l'IA, les fournisseurs de systèmes d'IA doivent proposer des solutions techniques « efficaces, interopérables, solides et fiables »[250] pour assurer un marquage lisible par une machine afin de permettre de détecter que le contenu a été généré ou manipulé artificiellement. Toutefois, cette obligation est subordonnée à la faisabilité technique, à l'état actuel de la technique et aux types de contenus générés. Cette obligation s'étend aux fournisseurs de systèmes d'IA à usage général[251].

Le Bureau de l'IA de l'UE basé à la Commission européenne peut également encourager et faciliter « l'élaboration de codes de bonne pratique au niveau de l'Union »[252] en ce qui concerne la détection et l'étiquetage de contenus générés ou manipulés artificiellement, et peut adopter des actes d'exécution pour adopter ces codes de pratique. Le 4 septembre 2025, une consultation a été organisée pour amorcer le processus d'élaboration de lignes directrices et de codes de pratique[253]. Si les éventuels codes de pratique sont inadéquats, la Commission peut adopter des actes d'exécution qui établissent des règles communes.

Trois points importants sont à retenir :

1. L'article 50 s'applique à tous les systèmes d'IA, qu'ils soient ou non à haut risque[254];
2. Le respect de l'article 50 ne rend pas pour autant licite l'utilisation ou la production d'un système d'IA[255], par exemple s'agissant de systèmes d'IA interdits en vertu de l'article 5.

---

247. Règlement européen sur l'IA, art. 3 (60) : L'expression « deep fake » ou « hypertrucage » désigne « une image ou un contenu audio ou vidéo généré ou manipulé par l'IA, présentant une ressemblance avec des personnes, des objets, des lieux, des entités ou événements existants et pouvant être perçu à tort par une personne comme authentiques ou véridiques ».
248. Le règlement sur l'IA, art. 50 (4), deuxième alinéa, précise que les déployeurs ne sont pas tenus d'informer le public au sujet d'un contenu généré artificiellement si ce contenu a « fait l'objet d'un processus d'examen humain ou de contrôle éditorial et lorsqu'une personne physique ou morale assume la responsabilité éditoriale de la publication du contenu ».
249. Règlement européen sur l'IA, art. 50 (1).
250. Règlement européen sur l'IA, art. 50 (2).
251. Règlement européen sur l'IA, art. 3 (66).
252. Règlement européen sur l'IA, art. 50 (7).
253. « La Commission lance une consultation en vue d'élaborer des lignes directrices et un code de bonnes pratiques sur les systèmes d'IA transparents », 4 septembre 2025, accessible à l'adresse https://digital-strategy.ec.europa.eu/fr/news/commission-launches-consultation-develop-guidelines-and-code-practice-transparent-ai-systems, consulté le 11 novembre 2025.
254. Règlement sur l'IA, art. 50 (6) et considérant 132.
255. Règlement sur l'IA, considérant 137.

3. Les obligations découlant de l'article 50 ne s'appliquent pas « aux systèmes d'IA dont la loi autorise l'utilisation à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière[256].

Corrélations avec d'autres cadres juridiques

La transparence et le contrôle comptent parmi les principes énoncés par la Convention-cadre du Conseil de l'Europe[257]. L'article 15 (2) de la Convention-cadre prévoit que « [c]haque Partie cherche à veiller à ce que, en fonction du contexte, les personnes qui interagissent avec des systèmes d'intelligence artificielle soient informées du fait qu'elles interagissent avec de tels systèmes et non avec un humain ».

4.1.1. Rôle des organismes de promotion de l'égalité et des SNDH concernant la transparence des exigences liées aux systèmes d'IA

Cet article peut être examiné sous l'angle des utilisations interdites énumérées à l'article 5 :

- Considérer l'article 50 en lien avec l'Article 5, paragraphe 1, points (a) et (b) et préconiser l'interdiction dans les articles où l'obligation de transparence prévue par cet article est insuffisante pour prévenir les atteintes aux droits fondamentaux; cela devrait inclure les situations où l'étiquetage de données générées et manipulées par l'IA selon l'état actuel de la technique ne constitue pas une solution technique « efficace, interopérable, solide et fiable »[258].
- Lire l'article 50 en lien avec l'article 5, paragraphe 1, point (f) qui interdit les systèmes de reconnaissance des émotions sur le lieu de travail et dans les établissements d'enseignement, et l'article 5, paragraphe 1, point g), qui interdit les 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 »[259]. L'article 50 ne rend pas licites les systèmes d'IA interdits.
- Prévenir les stratégies de « dérisquage » en réalisant ou en faisant réaliser une étude juridique pour identifier les différentes situations et conditions dans lesquelles les obligations de transparence de l'article 50 pourraient être invoquées pour contourner les interdictions de l'article 5. Cette étude devrait être publiée pour informer les fournisseurs et les déployeurs qu'ils ne doivent pas essayer de contourner les interdictions.

256. Règlement européen sur l'IA, art. 50 (1).
257. Convention-cadre du Conseil de l'Europe, art. 8.
258. Règlement européen sur l'IA, art. 50 (2).
259. Règlement sur l'IA, art. 5 (1) (g).

5. Application

5.1. Compétences des organes de protection des droits fondamentaux

5.1.1. Contexte et importance

L'article 77 confère aux OPE de nouveaux pouvoirs pour évaluer les discriminations causées par le biais des systèmes d'IA en leur octroyant un droit d'accès à la documentation, un droit de test et un droit de collaboration de la part des ASM et des opérateurs d'IA. Toutefois, les OPE ne sont pas automatiquement considérés comme des organismes compétents au sens de l'article 77.

Conditions cumulatives pour qu'un organisme de promotion de l'égalité constitue une autorité au sens de l'article 77

1. Ils doivent être considérés comme des « 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 »[260].

260. Règlement sur l'IA, art. 77 (1).

2. Leur État membre aurait dû identifier explicitement les OPE ou tout autre organisme de protection des droits fondamentaux, les inclure dans une liste transmise à la Commission et aux autres États membres, et rendre cette liste publique au plus tard le 2 novembre 2024. Les États membres doivent tenir « cette liste à jour »[261], ce qui indique que les OPE ou autres organismes de protection des droits fondamentaux qui n'ont pas été répertoriés par les États membres avant l'échéance du premier délai de notification peuvent encore être ajoutés. Les OPE et autres organismes peuvent également être retirés de cette liste.

Dans cette partie et dans le reste du document, lorsque nous faisons référence aux OPE, nous désignons spécifiquement les OPE désignés comme organismes visés par l'article 77 du règlement sur l'IA dans leur État membre.

Rôle des organismes de promotion de l'égalité dans la Convention-cadre du Conseil de l'Europe

Les organismes de promotion de l'égalité et les structures de protection des droits humains sont explicitement mentionnés dans la Convention-cadre du Conseil de l'Europe comme étant compétents pour faire partie des mesures de contrôle visant à garantir le respect des obligations de la Convention[262]. La Convention-cadre prévoit également que ces mécanismes « exercent leurs fonctions de manière indépendante et impartiale, et à ce qu'ils disposent des compétences, de l'expertise et des ressources nécessaires pour s'acquitter efficacement de leur mission de contrôle du respect des obligations nées de la présente Convention »[263]. Il s'agit d'une différence notable entre la Convention-cadre du Conseil de l'Europe et le règlement européen sur l'IA : au regard du règlement européen sur l'IA, les autorités de surveillance du marché sont chargées des activités de surveillance et d'application, tandis que les organismes de promotion de l'égalité et les SNDH sont identifiés comme des organismes relevant de l'article 77, mais ne font pas partie des mesures de supervision. Dans la Convention-cadre du Conseil de l'Europe, en revanche, les organismes de promotion de l'égalité et les SNDH pourraient être désignés comme des autorités chargées de la supervision et de l'application.

Pouvoirs des OPE visés par l'article 77

► Droit de demander et d'accéder à la documentation

Les OPE et les SNDH « 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 »[264]. Les OPE et les SNDH doivent interpréter et tester l'invocation

261. Règlement sur l'IA, art. 77 (2).
262. Conseil de l'Europe (2024), Rapport explicatif de 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, paragraphe 63.
263. Convention-cadre du Conseil de l'Europe, art. 26.
264. Règlement sur l'IA, art. 77 (1).

de cet article pour demander des documents prévus dans le règlement européen sur l'IA qui sont nécessaires aux OPE et aux SNDH pour remplir efficacement leur mandat. Les OPE et les SNDH doivent informer les ASM de leur pays lorsqu'ils font une demande de documentation. Cette documentation pourrait être utilisée pour identifier les préjudices causés par la discrimination à des individus ou à des groupes, par exemple lorsque le système d'IA est peu performant pour les femmes noires et qu'il entraîne la suppression des prestations sociales qui leur sont destinées, même si le système d'IA est globalement performant pour les personnes noires dans leur ensemble[265].

► Droit de tester

Si la documentation est insuffisante « pour déterminer s'il y a eu violation des obligations au titre du droit de l'Union protégeant les droits fondamentaux[266] », les OPE et les SNDH peuvent demander aux ASM d'organiser « des tests du système d'IA à haut risque par des moyens techniques [...] dans un délai raisonnable[267] ». Les cas d'insuffisance pourraient concerner des fournisseurs qui ne communiquent pas les résultats de plusieurs indicateurs de mesure de l'équité, ou qui définissent de manière étroite les groupes concernés, rendant ainsi impossible l'évaluation de types spécifiques de discrimination, en particulier la discrimination intersectionnelle (Equinet, 2025). Les OPE et les SNDH doivent collaborer étroitement avec les ASM pendant les tests.

► Droit à la collaboration avec les ASM et les opérateurs d'IA :

- Lorsqu'une ASM est informée par les fournisseurs de systèmes d'IA à haut risque d'un incident grave[268] qui porte atteinte à la législation européenne destinée à protéger les droits fondamentaux, dont le droit à la non-discrimination[269], l'ASM doit en informer les OPE et les SNDH[270].
- L'ASM évalue les « systèmes d'IA présentant un risque »[271] pour les droits fondamentaux, y compris des risques pour les groupes vulnérables.
  - L'ASM est tenue d'« informer et de coopérer pleinement[272] » avec les OPE et les SNDH lors de la réalisation de l'évaluation.
  - Les opérateurs concernés des systèmes d'IA sont également tenus de coopérer avec les ASM, les OPE et les SNDH[273].

265. Pour une analyse approfondie de (1) la nature de la documentation à laquelle les OPE pourraient avoir accès et de (2) la manière dont il faut comprendre les paramètres de base de cette documentation technique regroupée sur la base des normes techniques mises en œuvre, voir Equinet 2025.
266. Règlement sur l'IA, art. 77 (3).
267. Ibid.
268. Règlement sur l'IA, art. 73 (1).
269. Règlement européen sur l'IA, art. 3 (49) : Un incident grave désigne « un incident ou dysfonctionnement d'un système d'IA entraînant directement ou indirectement [...] (c) la violation des obligations au titre du droit de l'Union visant à protéger les droits fondamentaux ».
270. Règlement sur l'IA, art. 73 (7).
271. Règlement sur l'IA, art. 79 (1) et Règlement (UE) 2019/1020, Art. 3 (19).
272. Règlement sur l'IA, art. 79 (2).
273. Ibid.

- D'après leur évaluation, les ASM peuvent imposer aux opérateurs d'IA de prendre des mesures correctives, de retirer le système d'IA ou de le rappeler[274].

- Les OPE et les SNDH devraient coopérer avec les ASM et les aider à évaluer si les systèmes exclus du régime à haut risque présentent des risques significatifs pour les droits fondamentaux.
- Les ASM devraient déterminer si « un système d'IA classé par le fournisseur comme n'étant pas à haut risque [...] est en réalité à haut risque »[275] en raison d'un risque pour les droits fondamentaux. Les OPE peuvent aider les ASM dans cette évaluation. L'ASM doit consulter les OPE et les SNDH même lorsqu'un système d'IA à haut risque est conforme au règlement sur l'IA, mais présente tout de même un risque pour les droits fondamentaux[276].

Les OPE et les SNDH peuvent ainsi contribuer aux enquêtes des ASM et potentiellement suggérer des mesures correctives appropriées, y compris le retrait du système d'IA du marché, puisqu'ils peuvent proposer leur aide aux ASM pour déterminer si un système d'IA potentiellement discriminatoire est conforme ou non aux obligations prévues par le règlement européen sur l'IA.

Ces pouvoirs, lorsqu'ils sont utilisés, peuvent être efficaces pour évaluer et remédier aux discriminations causées par les systèmes d'IA.

En outre, il est important de souligner que le règlement sur l'IA est « sans préjudice des compétences, des tâches, des pouvoirs et de l'indépendance des autorités ou organismes publics nationaux compétents qui contrôlent l'application du droit de l'Union en matière de protection des droits fondamentaux, y compris les organismes chargés des questions d'égalité[277] ». En d'autres termes, les enquêtes visant à évaluer le respect du règlement sur l'IA n'affectent pas les enquêtes visant à évaluer le respect des lois sur l'égalité[278]. De même, les personnes touchées ont le droit de déposer une plainte au titre de l'article 85 du règlement européen sur l'IA, ainsi qu'en vertu des lois en matière d'égalité.

274. Note : règlement (UE) 2019/1020, Art. 18 : Les opérateurs d'IA ont le droit d'être entendus avant qu'une telle mesure ne soit prise.
275. Règlement européen sur l'IA, art. 80 (1).
276. Règlement sur l'IA, art. 82 (1).
277. Règlement sur l'IA, considérant 157.
278. De même, les personnes concernées ont le droit de déposer une plainte en vertu de l'article 85 du règlement sur l'IA et des lois sur l'égalité

5.1.2. Rôle des organismes de promotion de l'égalité et des SNDH vis-à-vis de l'article 77 du règlement sur l'IA

- Plaider en faveur de l'obtention du statut d'organisme visé par l'article 77.
- En qualité d'organismes visés par l'article 77, par le biais de leur droit de demander et d'accéder à la documentation :
  - Préparer des arguments solides en faveur de leur droit d'accès à l'information, en s'appuyant sur les travaux d'Equinet (cf. Equinet 2025). Préparer des arguments solides en faveur de leur droit d'accès à l'information, en s'appuyant sur les travaux d'Equinet (cf. Equinet 2025). Sans cela, les OPE et les SNDH sont susceptibles d'être confrontés à des problèmes posés par les entreprises et les autorités publiques dans leurs tentatives d'empêcher l'accès aux documents.
  - Élaborer des scénarios de test sur la base des informations disponibles, afin de déterminer si ces informations sont pertinentes pour l'exécution de leur mandat.
  - Utiliser la documentation reçue pour enquêter sur les préjudices liés à la discrimination.
  - Vérifier l'absence d'« éthique-washing » et évaluer la façon dont les groupes ont été définis par les opérateurs d'IA lors des tests de biais.
- En qualité d'organismes visés par l'article 77, dans le contexte des tests :
  - Veiller à disposer de ressources suffisantes pour pouvoir coopérer efficacement avec les ASM, qui mèneront les tests. Il peut notamment s'agir de disposer d'une expertise technique interne, de personnel non technique formé pour collaborer avec des experts techniques, ou d'établir des partenariats avec des partenaires extérieurs (pour plus de détails, voir la partie Coopération).
  - Élaborer des modèles de coopération avec les ASM.
- En qualité d'organismes visés par l'article 77, dans le cadre des actions correctives menées par les ASM :
  - Sensibiliser aux risques de discrimination associés aux systèmes d'IA, par exemple en publiant leurs décisions.
  - Suivre et faire connaître les actions liées à leur rôle en tant qu'organismes relevant de l'article 77, afin de sensibiliser le grand public à leur rôle et d'illustrer la nécessité de promouvoir l'égalité et la non-discrimination dans les systèmes d'IA.

- ► Publier des orientations et plaider en faveur de la prise en compte des droits fondamentaux par la Commission européenne :
  - La Commission européenne « établit un formulaire de documentation technique simplifié ciblant les besoins des petites entreprises et des microentreprises »[279]. L'atteinte aux droits des personnes n'est pas liée à la taille des entreprises qui développent des systèmes d'IA à haut risque. Les OPE et les SNDH doivent veiller à ce que les informations essentielles de la documentation ne soient pas omises.
  - La Commission européenne est habilitée à adopter des actes délégués conformément à l'article 97 (exercice de la délégation) afin de modifier l'annexe IV (documentation technique des systèmes d'IA à haut risque).
  - Contribuer aux orientations de la Commission européenne sur la déclaration des incidents graves. La Commission européenne devrait élaborer des orientations sur la déclaration des incidents graves d'ici le 2 août 2025[280]. Ces orientations incluront probablement une coordination et des protocoles pour le transfert d'informations des ASM aux OPE et SNDH. Les OPE devraient envisager d'influencer ces orientations par l'intermédiaire d'Equinet ou d'autres organismes.
  - Voir également les recommandations répertoriées dans ces lignes directrices en ce qui concerne la classification, la modification des listes des systèmes d'IA à haut risque e interdits, les analyses d'impact sur les droits fondamentaux, etc.

5.2. Recours

5.2.1. Contexte et pertinence

Le règlement sur l'IA confère aux personnes physiques et morales le droit de déposer une plainte auprès d'une ASM et, aux particuliers, un droit à l'explication. Toutefois, aucun de ces droits n'est en soi un droit effectif.

Droit de déposer une plainte

Une personne touchée peut déposer une plainte auprès d'une ASM (article 85). La plainte sera prise « en compte aux fins de l'exercice des activités de surveillance du marché »[281] et traitée conformément aux « procédures concernant les suites à donner aux plaintes »[282]. Toutefois, l'ASM n'a aucune obligation générale de donner suite à une plainte déposée par une personne touchée. L'auteur de la plainte n'a pas de statut juridique dans la procédure, même si la plainte est prise « en compte

279. Règlement sur l'IA, art. 11 (1) deuxième alinéa.
280. Règlement sur l'IA, art. 73 (7).
281. Règlement sur l'IA, art. 85.
282. Règlement (UE) 2019/1020, Art. 11 (7) (a).

aux fins des activités de surveillance du marché »[283]. Ainsi, le droit de plainte prévu par le règlement européen sur l'IA peut, au mieux, être considéré comme un droit partiel. Néanmoins, ce droit s'applique sans préjudice des autres recours prévus par d'autres lois nationales ou européennes[284].

Lorsqu'un système d'IA porte atteinte aux droits fondamentaux d'une personne, celle-ci devrait plutôt envisager de déposer une plainte auprès d'un organisme de défense des droits fondamentaux, par exemple un organisme de promotion de l'égalité dans le cas d'une discrimination. Lorsqu'il s'agit de données à caractère personnel, une personne concernée peut également avoir la possibilité de déposer une plainte auprès de l'autorité compétente en matière de protection des données[285].

Droit à une explication

Le règlement sur l'IA confère également à toute personne concernée le droit d'obtenir une explication (article 86) lorsque le déployeur prend une décision « sur la base des sorties d'un système d'IA à haut risque mentionné à l'annexe III [...] qui produit des effets juridiques ou affecte significativement cette personne de façon similaire d'une manière qu'elle considère comme ayant eu des conséquences négatives sur sa santé, sa sécurité ou ses droits fondamentaux »[286]. L'explication doit comprendre « des explications claires et pertinentes sur le rôle du système d'IA dans la procédure décisionnelle et sur les principaux éléments de la décision prise »[287].

Quand le droit à l'explication prévu par le règlement sur l'IA s'applique-t-il ?

Toutefois, le droit à l'explication prévu par le règlement sur l'IA ne s'applique que si ce même droit « n'est pas prévu par ailleurs dans e droit de l'Union »[288]. Lorsque des données à caractère personnel sont traitées par des systèmes d'IA à haut risque, le droit à l'information prévu à l'article 15, paragraphe 1, point h), du RGPD, qui comprend des « informations utiles concernant la logique sous-jacente »[289], s'applique.

En outre, il est important de souligner que l'article 22, paragraphe 1, du RGPD interdit de manière générale toute « décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou

283. Règlement sur l'IA, art. 85. Une partie touchée peut également envisager d'autres voies de recours, notamment au titre de la directive (UE) 2020/1828 du Parlement européen et du Conseil du 25 novembre 2020 relative aux actions représentatives visant à protéger les intérêts collectifs des consommateurs et abrogeant la directive 2009/22/CE. Voir également le règlement sur l'IA, art. 110.
284. Règlement sur l'IA, considérant 170.
285. RGPD, art. 77-79.
286. Règlement sur l'IA, art. 86 (1) (italique ajouté)
287. Règlement sur l'IA, art. 86 (1).
288. Règlement sur l'IA, art. 86 (3).
289. RGPD, Art. 15 (1) (h) : « (1) La personne concernée a le droit d'obtenir du responsable du traitement la confirmation que des données à caractère personnel la concernant sont ou ne sont pas traitées et, lorsqu'elles le sont, l'accès auxdites données à caractère personnel ainsi que les informations suivantes :... h) 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. »

l'affectant de manière significative de façon similaire ». Cela signifie que les systèmes d'IA à haut risque utilisant des données à caractère personnel qui produisent « des effets juridiques » ou « l'affectant de manière significative de façon similaire » ne peuvent être utilisés que lorsque des exceptions spécifiques au titre de l'article 22, paragraphe 2, du RGPD s'appliquent, c'est-à-dire lorsque leur utilisation :

(a) est nécessaire à la conclusion ou à l'exécution d'un contrat entre la personne concernée et un responsable du traitement;

(b) est autorisée par la législation de l'Union européenne ou d'un État membre qui s'applique au responsable du traitement, et qui prévoit également des mesures adéquates pour protéger les droits et libertés de la personne concernée et ses intérêts légitimes; ou

(c) repose sur le consentement explicite de la personne concernée.

En outre, le terme « exclusivement » a été interprété dans l'affaire Schufa comme incluant un traitement automatisé qui peut n'être qu'une partie de la décision mais qui joue néanmoins un « rôle déterminant »[290].

Ainsi, l'article 86 du règlement sur l'IA peut être utile dans deux cas de figure :

1. Lorsque les données à caractère personnel ne sont pas impliquées dans le système d'IA à haut risque (sans quoi l'article 15, paragraphe 1, point h) du RGPD s'appliquerait); et

2. Lorsque les données à caractère personnel sont impliquées et que le traitement automatisé via le système d'IA à haut risque ne joue pas un « rôle déterminant » dans la décision (cf. décision de la CJUE concernant Schufa plus haut).

Qu'est-ce qu'une explication pertinente ?

Le règlement sur l'IA exige « des explications claires et pertinentes sur le rôle du système d'IA dans la procédure décisionnelle et sur les principaux éléments de la décision prise »[291]. L'Arrêt CK contre Dun & Bradstreet pourrait aider à clarifier cette question. Il y a en effet été précisé que « les informations utiles concernant la logique sous-jacente »[292] mentionnées dans le RGPD signifiaient « d'une façon concise, transparente, compréhensible et aisément accessible, en des termes clairs et simples, la procédure et les principes concrètement appliqués pour exploiter, par la voie automatisée, les données à caractère personnel relatives à cette personne aux fins d'en obtenir un résultat déterminé »[293]. Cette explication « n'est pas nécessairement une explication complexe des algorithmes utilisés ou la divulgation de l'algorithme complet »[294].

290. Arrêt de la Cour de justice du 7 décembre 2023, SCHUFA Holding (Scoring), C-634/21, EU:C:2023:957, point 50.
291. Règlement sur l'IA, art. 86 (1).
292. RGPD, Article 15 (1) (h)
293. Arrêt de la Cour (première chambre) du 27 février 2025. CK contre Dun & Bradstreet Austria GmbH et Magistrat der Stadt Wien, C-203/22, ECLI:EU:C:2025:117, paragraphe 66.
294. Ibid. paragraphe.

Protection du secret commercial

Le règlement sur l'IA prévoit que « le droit d'obtenir une explication ne devrait pas s'appliquer à l'utilisation de systèmes d'IA pour lesquels des exceptions ou des restrictions découlent du droit de l'Union ou du droit national »[295]. L'une de ces exceptions pourrait être la protection des secrets commerciaux[296]. Comme dans l'affaire CK contre Dun & Bradstreet, le droit à l'explication devrait être mis en balance avec la protection de la propriété intellectuelle, y compris les secrets commerciaux. Dans cette affaire, le tribunal a jugé que l'équilibre des droits devait être assuré par le régulateur ou un tribunal compétent, et non par l'entreprise[297]. Dans le cas du règlement sur l'IA, un argument analogue peut être avancé.

5.2.2. Rôle des organismes de promotion de l'égalité et des SNDH concernant les recours

- Informer les citoyens et citoyennes de leurs droits. Dans le cadre de leur devoir d'assistance aux victimes de discrimination (tel que décrit dans les directives sur les normes), les OPE devraient informer les personnes de leur droit à l'explication et de leur droit à déposer une plainte auprès de l'autorité de surveillance du marché en vertu du règlement sur l'IA, et auprès de l'autorité de protection des données en vertu de la législation de l'UE sur la protection des données. Cela nécessite de former le personnel à reconnaître les cas où le droit à l'explication s'applique.
- Il s'agit d'avertir les personnes que ce droit vient s'ajouter aux recours existants, et non s'y substituer.

5.3. Mécanismes de coopération

5.3.1. Contexte et pertinence

Pour être efficaces et exercer les pouvoirs que leur confère le règlement sur l'IA, les OPE et les SNDH doivent coopérer avec les ASM de leur pays ainsi qu'avec d'autres organismes relevant de l'article 77, notamment les autorités chargées de la protection des données. En outre, la consultation des organisations de la société civile serait bénéfique pour le travail des OPE.

295. Règlement sur l'IA, considérant 171.
296. Directive 2016/943, article 2 (1).
297. Arrêt de la Cour (première chambre) du 27 février 2025. CK contre Dun & Bradstreet Austria GmbH et Magistrat der Stadt Wien, C-203/22, ECLI:EU:C:2025:117, paragraphe 76.

5.3.2. Rôle des organismes de promotion de l'égalité et des SNDH concernant la coopération avec différentes parties prenantes

Coopération entre organismes relevant de l'article 77

Les OPE et les SNDH doivent convaincre leur État membre qu'il existe une base juridique claire en ce qui concerne le partage d'informations entre les organismes relevant de l'article 77, si ce n'est pas déjà le cas. Le règlement sur l'IA prévoit le partage d'informations entre les ASM et les OPE/SNDH.[298] Toutefois, dans certaines situations, les OPE et les SNDH devront également coopérer avec d'autres organismes relevant de l'article 77 dans leur pays, ce qui nécessitera une base juridique distincte pour régir le partage d'informations.

Coopération avec les ASM

Les OE et les SNDH doivent établir une relation de travail étroite avec les ASM dès leur création. La structure précise de gouvernance dans chaque État membre peut varier. Il est possible que certains pays choisissent un seul régulateur pour faire office d'autorité de surveillance du marché, tandis que d'autres opteront pour plusieurs régulateurs. Les Pays-Bas et l'Irlande ont indiqué leur préférence pour des régulateurs multiples dans un réseau en étoile, avec un régulateur principal épaulant les autres régulateurs. La France a mis en place un système de gouvernance coordonné, avec la direction générale de la Concurrence, de la Consommation et de la Répression des Fraudes, qui est chargée de la coordination opérationnelle et assure la liaison avec la direction générale des Entreprises à des fins de coordination stratégique. Plus de dix autres entités ont ensuite la charge d'assurer des dispositions spécifiques du règlement européen sur l'IA : il s'agit notamment de la Commission nationale de l'informatique et des libertés (CNIL) et de l'Autorité de régulation de la communication audiovisuelle et numérique (Arcom) (voir direction générale des Entreprises, 2025). Dans ce contexte, les OPE devraient établir une collaboration avec toutes les autorités de surveillance du marché en accordant une priorité particulière à l'autorité de surveillance du marché principale.

Établir des bases juridiques pour le partage d'informations

Les OPE et les SNDH doivent convaincre le gouvernement de leur État membre et s'assurer que l'État membre a établi une large base juridique pour le partage d'informations entre organismes relevant de l'article 77, mais également entre les ASM et les organismes relevant de l'article 77[299]. L'existence de ces bases juridiques est une condition préalable à la création d'accords de coopération bilatéraux et multilatéraux permettant le partage d'informations. Ces accords devraient être globaux, de sorte qu'ils couvrent les attributions des organismes relevant de l'article 77 et des ASM. Ces accords devraient permettre la conclusion d'accords bilatéraux complémentaires spécifiques, le cas échéant[300].

298. Règlement sur l'IA, art. 73 (7) et 79 (2).
299. De Autoriteit Persoonsgegevens (2024), Final recommendation on supervision of AI : sector and centrally coordinated, paragraphe 37, disponible en anglais à l'adresse www.autoriteitpersoonsgegevens.nl/en/current/final-recommendation-on-supervision-of-ai-sector-and-centrally-coordinated, consulté le 11 novembre 2025.
300. Ibid.

Mettre en place des groupes de travail consacrés aux questions d'égalité

Les OPE et les SNDH devraient, aux côtés des ASM, former des groupes de travail qui pourraient mettre en commun leur expertise sur des sujets spécifiques. Dans un premier temps, ces groupes de travail pourraient viser à établir une compréhension commune entre organismes. Par exemple, un groupe de travail pourrait être axé sur « Les meilleures pratiques en matière d'identification des préjudices liés à la discrimination par les systèmes d'IA ». Cela pourrait éclairer les OPE dans l'élaboration d'orientations à l'intention des ASM concernant l'identification des préjudices liés à la discrimination et leur signalement aux OPE.

Créer des systèmes d'éléments déclencheurs automatiques qui obligent l'ASM à coopérer avec les organismes de promotion de l'égalité et les SNDH

Il est essentiel que les ASM soient pleinement conscientes de l'expertise des OPE et des SNDH, notamment en matière de non-discrimination et d'égalité. Les ASM devraient disposer d'un système d'éléments déclencheurs automatiques pour informer les OPE lorsqu'un incident grave portant atteinte au droit à la non-discrimination est signalé conformément à l'article 73, paragraphe 7. Lorsque les ASM ne sont pas certaines de l'existence d'un risque de discrimination, elles devraient consulter immédiatement les OPE, par exemple dans les situations visées aux articles 79 et 80 du règlement européen sur l'IA, afin que les OPE puissent apporter leur expertise. Il est préférable de recevoir des faux signalements plutôt que de passer à côté d'infractions que les ASM auraient négligées en raison d'un manque d'expertise.

Signalement de cas de non-conformité par d'autres pays de l'UE

Lorsque les ASM constatent un cas de non-conformité qui ne se limite pas à leur territoire national, elles sont tenues d'informer « la Commission et les autres États membres, sans retard injustifié, des résultats de l'évaluation et des mesures qu'elle a exigées de l'opérateur »[301] et de transmettre également un signalement détaillé si l'opérateur ne se conforme pas aux mesures requises[302]. Les OPE et les SNDH devraient insister auprès des ASM et du Comité IA sur le fait que les OPE et les SNDH des autres États membres devraient être informés et notifiés lorsque l'évaluation d'une non-conformité fait apparaître des risques pour les droits fondamentaux tels que le droit à la non-discrimination.

Accès aux documents

Les OPE et les SNDH peuvent demander des documents aux fournisseurs de systèmes d'IA à haut risque[303]. Les OPE peuvent également accéder à la documentation des ASM si une ASM a déjà accédé à cette documentation du fournisseur. Cela pourrait accélérer le processus d'accès pour les OPE et les SNDH. Dans les deux cas, il est important de souligner que les entreprises ne peuvent pas empêcher les ASM de communiquer la documentation aux OPE et aux SNDH au titre de la protection du secret commercial. En particulier si les OPE et les SNDH demandent la documentation par l'intermédiaire de l'ASM, le récent arrêt indique

301. Règlement sur l'IA, art. 79 (3).
302. Règlement sur l'IA, art. 79 (5)-(6).
303. Règlement sur l'IA, art. 77 (1).

que l'entreprise ne peut pas décider si la documentation peut être partagée ou non[304]. Cette décision appartient au régulateur ou au tribunal.

Participation aux essais

Lorsque les OPE et les SNDH estiment que la documentation fournie par les fournisseurs est insuffisante pour attester qu'il y a eu violation de la législation européenne protégeant les droits fondamentaux, ils peuvent demander à l'ASM d'« organiser des tests du système d'IA à haut risque par des moyens techniques [...] dans un délai raisonnable »[305]. Le règlement sur l'IA prévoit que les ASM organisent les tests. Cela laisse une marge de manœuvre pour que les tests soient effectués par un fournisseur ou un expert tiers. Les tests doivent être menés avec « la participation étroite »[306] des OPE et/ou SNDH.

Il est important pour les OPE et les SNDH d'établir :

- qui effectuera les tests ?
- qu'est-ce qui est testé ?
- la nature exacte du « délai raisonnable ». Ce délai doit être établi le plus tôt possible et non pas lorsqu'un cas pertinent se présente;
- la compréhension mutuelle précise des OPE/SNDH et des ASM du concept de « participation étroite ». Cela nécessite-t-il un point de contact directement impliqué dans les tests ? Les OPE/SNDH doivent-ils être consultés au préalable sur les différentes étapes des tests ?
- Mettre au point un protocole pour la collaboration inter-agences ?

Exemples et bonnes pratiques

Il est possible de s'inspirer des coopérations passées entre des organismes de promotion de l'égalité et des autorités chargées de la protection des données. Au Royaume-Uni, la Commission pour l'égalité et les droits (EHRC) a collaboré avec l'Information Commissioner's Office dans le cadre du défi de l'équité et de l'innovation (« Fairness and Innovation Challenge »), sur des cas pilotes dans l'enseignement supérieur, la finance, la santé et le recrutement, afin de trouver de nouvelles façons de traiter les biais et les discriminations statistiques, humains et structurels dans les systèmes d'IA. Ce défi était financé par le ministère britannique dédié à la science[307]. En Norvège, l'organisme de promotion de l'égalité a coopéré avec l'autorité chargée de la protection des données dans le cadre d'un bac à sable réglementaire consacré à l'intelligence artificielle[308].

En France, la coopération entre le Défenseur des droits et la CNIL est facilitée par plusieurs mécanismes : les deux institutions ont conclu un protocole d'accord et le Défenseur des droits est également membre avec voix consultative du Collège

304. Arrêt de la Cour (première chambre) du 27 février 2025. CK contre Dun & Bradstreet Austria GmbH et Magistrat der Stadt Wien, C-203/22, ECLI:EU:C:2025:117, p. 76.
305. Règlement sur l'IA, art. 77 (3).
306. Ibid.
307. Gov.uk (2024), communiqué de presse : "AI Fairness Innovation Challenge winners announced", consultable en anglais sur www.gov.uk/government/news/ai-fairness-innovation-challenge-winners-announced, consulté le 11 novembre 2025.
308. Voir www.datatilsynet.no/en/regulations-and-tools/sandbox-for-artificial-intelligence/, consulté le 11 novembre 2025.

de la CNIL. Grâce à ce lien structurel, le Défenseur des droits est informé de tous les projets soumis aux réunions plénières de la CNIL[309].

Coopération avec d'autres agences publiques

La coopération doit également être envisagée comme un moyen de mettre en commun des ressources spécifiques. Par exemple, la France a mis en place un groupe de travail sur la science des données appelé PEReN (Pôle d'expertise de la régulation numérique), composé d'une vingtaine de scientifiques des données, qui est hébergé par le ministère de l'Économie, mais accessible à tous les organismes centraux et les régulateurs. Cela permet à l'OPE Défenseur des droits français d'avoir accès à des connaissances et à des compétences techniques de pointe. Il est important, dans le cas d'une telle coopération, d'instaurer des mécanismes de collaboration clairs afin de s'assurer que les besoins des OPE seront satisfaits et de mettre en place des garanties pour l'indépendance des OPE.

Coopération avec la société civile et d'autres parties prenantes externes

Mettre en place dans chaque État membre un forum consultatif composé de membres de la société civile, d'organisations de terrain et d'universitaires qui, ensemble, combinent l'expérience du terrain et l'expertise en matière de législation sur l'égalité et la non-discrimination, mais aussi en matière de droits numériques. Ce forum consultatif peut signaler les utilisations préjudiciables de l'IA sur lesquelles les organismes de promotion de l'égalité e leurs États membres peuvent enquêter en priorité. Le forum devrait se réunir régulièrement et échanger avec les organismes de promotion de l'égalité.

Poursuivre et approfondir les interactions existantes avec les organisations de la société civile

Equinet et les organismes de promotion de l'égalité de certains États membres (par exemple, l'OPE belge Unia) collaborent régulièrement avec les organisations de la société civile en les invitant à des événements et à des débats. Ces interactions offrent la possibilité d'échanger des informations qui peuvent bénéficier aux deux parties. Il est possible d'approfondir ces interactions autour de l'IA en créant des liens entre les organisations de la société civile et les organismes de promotion de l'égalité dans d'autres États membres. Toutefois, cela nécessiterait probablement davantage de ressources humaines. Une solution consisterait à développer les structures de coopération formelles qui existent dans d'autres contextes réglementaires.

Améliorer la gestion des personnes plaignantes

Établir une procédure de traitement des plaintes claire et simple pour les préjudices liés à la discrimination causée par les systèmes d'IA[310]. Partager la procédure adoptée par les organismes de promotion de l'égalité/SNDH pour recevoir les plaintes dans le cadre de son mandat et protéger les droits de la personne plaignante tout au long de la procédure. Sensibiliser la société civile au processus, avec

309. Informations issues de l'entretien avec le Défenseur des droits.

310. Voir également Equinet (2023), Minimal guidelines on improving complaints data collection by equality bodies, disponible en anglais à l'adresse https://equineteurope.org/wp-content/uploads/2024/01/Minimal-Guidelines-on-Improving-Complaints-Data-Collection-by-Equality-Bodies-1.pdf, consulté le 11 novembre 2025.

des exemples de cas concrets qui clarifient les types de plaintes que les OPE et les SNDH peuvent traiter dans leur périmètre d'action. Il peut être utile de mettre également en évidence les cas qui relèvent de la compétence des régulateurs en vertu du règlement sur l'IA et/ou du RGPD, ainsi que les droits des personnes plaignantes en vertu de ces lois. Encourager les organisations de la société civile à partager ces informations avec les personnes lésées par les systèmes d'IA afin que les plaignant.e.s potentiel.le.s aient les moyens d'agir.

Des expert.es externes peuvent également apporter une expertise technique

Le projet AI Equality by Design a permis aux professeurs de donner des conseils académiques sur des cas d'IA et de discrimination[311]. Il peut être nécessaire de chercher à obtenir des financements auprès du gouvernement, via des initiatives de recherche universitaire, au niveau des procédures judiciaires ou ailleurs pour obtenir des preuves ou des conseils d'expert.es.

311. Equinet (2025), Embedding equality safeguards into technical standards for the EU AIA and empowering equality defenders: Equinet's participation in the project "Equality by Design, Deliberation and Oversight", disponible en anglais sur https://equineteurope.org/latest-developments-in-ai-equality/, consulté le 11 novembre 2025.

PARTIE II

6. Directives sur les normes

6.1. Contexte général

Le 7 mai 2024, deux nouvelles directives, surnommées « directives sur les normes », ont été adoptées pour garantir l'efficacité et l'indépendance des organismes de promotion de l'égalité et, ainsi, renforcer « l'application du principe de l'égalité de traitement »[312]. Ces directives font suite aux efforts de normalisation déployés par Equinet[313], la Commission européenne contre le racisme et l'intolérance (ECRI[314]) et la Commission européenne[315].

Elles établissent des normes minimales harmonisées :

- Directive 2024/1500 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;

312. Directives sur les normes, art. 1.
313. Equinet (2016), Developing Standards for Equality Bodies: An Equinet Working Paper.
314. ECRI, Recommandation de politique générale n° 2 (révisée) : Organismes de promotion de l'égalité chargés de lutter contre le racisme et l'intolérance au niveau national, décembre 2017.
315. Recommandation de la Commission (UE) 2018/951 du 22 juin 2018 relative aux normes applicables aux organismes pour l'égalité de traitement

- Directive 2024/1499 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.

Les directives sur les normes[316] peuvent aider les OPE à prévenir et à mettre au jour les discriminations dans les systèmes d'IA. Elles mentionnent explicitement l'intelligence artificielle et les systèmes automatisés dans leurs considérants :

> « 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 des 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 mandats multiples, il convient de garantir les ressources nécessaires à l'accomplissement de son mandat ayant trait à l'égalité »[317].

Les directives sur les normes doivent être transposées en droit national d'ici juin 2026.

6.2. Modifications du mandat et des ressources

6.2.1. Élargissement de la portée du mandat

Les directives sur les normes élargissent la portée du mandat des organismes de promotion de l'égalité. En particulier, les États qui ne l'ont pas encore fait devraient élargir le mandat des organismes de promotion de l'égalité aux domaines de la vie couverts par la directive sur la sécurité sociale égalitaire (79/7/CEE) et la directive-cadre sur l'emploi (2000/78/CE), en leur conférant davantage de compétences pour couvrir les motifs de la religion ou des convictions, de l'âge, du handicap et de l'orientation sexuelle dans le domaine de l'emploi, et du sexe et du genre dans le domaine de la sécurité sociale[318]. Ce point est important car les systèmes de prise de décision automatisée (PDA) dans le domaine de la sécurité sociale se sont souvent

316. Equinet (2024 et le MOOC créé par le Conseil de l'Europe, qui offrent une vue d'ensemble de la réglementation
317. Directive 2024/1500, considérant 21 et directive 2024/1499, considérant 22.
318. Directive 2024/1499, Art. 1 (2). Dans la pratique, le champ d'action de nombreux organismes de promotion de l'égalité dépasse déjà les exigences minimales de la législation européenne en matière de discrimination. Voir Equinet 2024 :24.

révélés biaisés à l'encontre des femmes, comme cela a été observé en Autriche, en France et en Suède[319].

Il est important de noter que le champ d'application des directives sur les normes reste limité aux motifs et domaines de la vie couverts par la législation européenne en matière de non-discrimination[320]. Toutefois, les directives sur les normes fixent des exigences minimales et le mandat des OPE peut être plus large au niveau national.

En outre, les directives sur les normes font explicitement référence au concept de discrimination intersectionnelle dans les activités de promotion et de prévention des OPE. Les directives sur les normes prévoient que, dans ce contexte, les OPE « peuvent prendre en considération des situations spécifiques de désavantage résultant d'une discrimination intersectionnelle »[321]. Dans les considérants des directives, il est également établi que « lorsqu'ils promeuvent l'égalité de traitement, préviennent la discrimination, recueillent des données sur la discrimination et aident les victimes conformément à la présente directive, il importe que les organismes pour l'égalité de traitement accordent une attention particulière à la discrimination intersectionnelle »[322].

Les systèmes d'IA, et tout particulièrement les systèmes d'analytique prédictive, étant susceptibles de donner lieu à des formes de discrimination intersectionnelles, il sera d'autant plus important que les OPE s'emparent de ces nouveaux pouvoirs, notamment sous la forme d'activités de promotion et de prévention (Xenidis, 2020).

6.2.2. Obligation de fournir des ressources suffisantes

L'article 4 prévoit que les États doivent veiller à ce que « chaque organisme pour l'égalité de traitement dispose des ressources humaines, techniques et financières dont il a besoin pour accomplir toutes ses missions et exercer toutes ses compétences de manière efficace ». Les considérants susmentionnés établissent un lien explicite entre la question des ressources et le traitement des systèmes d'IA et de PDA[323], qui peuvent nécessiter des ressources techniques et humaines nouvelles et supplémentaires.

---

319 Allhutter, D. et al. (2020), Algorithmic profiling of job seekers in Austria: How austerity politics are made effective, Frontiers in Big Data, 3, disponible sur https://doi.org/10.3389/fdata.2020.00005; Romain, M. et al. (2023), Is data neutral? How an algorithm decides which French households to audit for welfare fraud, Le Monde, disponible à l'adresse https://www.lemonde.fr/en/les-decodeurs/visuel/2023/12/05/how-an-algorithm-decides-which-french-households-to-audit-for-benefit-fraud_6313254_8.html; Amnesty International (2024), Sweden: Authorities must discontinue discriminatory AI systems used by welfare agency, disponible à l'adresse : https://www.amnesty.org/en/latest/news/2024/11/sweden-authorities-must-discontinue-discriminatory-ai-systems-used-by-welfare-agency/, consultés le 11 novembre 2025.

320 Directives 2024/1499 et 2024/1500, considérant 15.

321 Directive 2024/1500, Art. 5 (2).

322 Directive 2024/1499, considérant 16; 2024/1500, considérant 15; et également directive (UE) 2023/970 (égalité des rémunérations), considérant 15.

323 Directive 2024/1499, considérant 22; Directive 2024/1500, considérant 21.

Qu'est-ce que cela signifie en pratique ?

► Ressources humaines

Les organismes de promotion de l'égalité doivent disposer d'un « personnel qualifié [...] pour accomplir chacune de leurs missions de manière effective »[324], « embaucher suffisamment de personnes possédant des compétences diverses et complémentaires pour accomplir toutes ses missions de manière effective » et « proposer des salaires et des conditions de travail compétitifs »[325]. Dans le contexte des questions liées aux systèmes d'IA et de PDA, cela signifie qu'il faut :

- Embaucher des technologues, notamment des scientifiques des données, et leur proposer des salaires compétitifs afin de renforcer l'expertise technique interne des organismes de promotion de l'égalité. Une autre solution consisterait à s'assurer (par l'embauche ou la formation continue) que les organismes de promotion de l'égalité sont équipés en interne pour sélectionner des expert. es techniques externes et assurer la liaison avec eux et elles, par exemple, des chercheur.ses ou des expert.es au sein des ASM ou d'autres organismes relevant de l'article 77;
- S'assurer (par l'embauche ou la formation continue) qu'un nombre suffisant d'expert.es juridiques et politiques au sein des organismes de promotion de l'égalité connaissent le fonctionnement des systèmes de PDA/IA (et possèdent a minima une compréhension de base des questions techniques) et la législation relative à l'IA et aux PDA, en particulier le règlement sur l'IA; et
- Former le personnel chargé de recevoir les plaintes à la détection du recours potentielle à des systèmes de PDA/IA dans le cadre des plaintes déposées auprès des organismes de promotion de l'égalité, et disposer d'un personnel adapté, capable d'étudier les plaintes et de poser des questions de suivi, le cas échéant.

L'Institut néerlandais des droits humains, après avoir mené avec succès une enquête ex officio sur une éventuelle discrimination systémique concernant le fonctionnement du système de prestations de garde d'enfants, prévoit d'institutionnaliser un cadre pour traiter les cas de données fondés sur les technologies. Compte tenu de l'ampleur des ressources nécessaires pour ce projet, l'Institut estime que, dans l'idéal, cela devrait être accompli par la formation de juristes salarié.es et l'embauche de scientifiques des données (Ilieva 2024: 73).

De nombreuses recommandations relatives à la formation initiale et continue en la matière concordent avec ce que prévoit l'article 20 de la Convention-cadre du Conseil de l'Europe, qui porte sur la maîtrise du numérique et les compétences en la matière pour « toutes les catégories de la population, notamment les compétences spécifiques de pointe pour les personnes chargées de l'identification, de l'évaluation, de la prévention et de l'atténuation des risques posés par les systèmes d'intelligence artificielle ».

---

324. Directive 2024/1499, considérant 21; Directive 2024/1500, considérant 20.
325. Equinet (2023), Measuring Standards for Equality Bodies: Indicators for Self-Assessment, Ressources, Indicateurs 2.2.1.1 et 2.2.1.2.

► Ressources techniques

Les ressources techniques peuvent être considérées comme le troisième volet des ressources après le personnel et le financement des organismes de promotion de l'égalité (Equinet, 2024). Cette catégorie comprend les locaux, les infrastructures et les ressources suffisantes pour répondre aux besoins informatiques, ce qui inclut, entre autres, les systèmes de collecte de données et les outils de consultation en ligne[326]. Plus précisément, dans le contexte de la lutte contre l'IA et les systèmes de PDA, il peut s'agir de :

- veiller à ce que le système informatique de traitement des plaintes permette de signaler et de repérer facilement les plaintes pour lesquelles on soupçonne que des systèmes d'IA et de PDA ont été utilisés. Des constantes devraient pouvoir être décelées parmi les différentes plaintes qui pourraient être liées à des systèmes d'IA et de PDA. Cela permettrait de remédier à l'un des obstacles pratiques à l'identification des discriminations causées par des systèmes d'IA et de PDA, qui avait été mis en exergue par Unia : « la charge de travail qui pèse sur le traitement des plaintes est telle que les plaintes déposées dans le système technique ne permettent pas toujours de détecter facilement l'arrivée de plusieurs plaintes concernant une même organisation. Un autre point à améliorer dans le système technique est que, bien qu'il y ait un espace pour indiquer si un système d'IA ou de PDA est impliqué dans la plainte, cet espace est enfoui dans les multiples formulaires que les utilisateurs doivent remplir pour déposer une plainte et est donc facilement négligé (Xenidis, 2025) ».
- disposer d'infrastructures informatiques suffisamment sécurisées pour effectuer des analyses de données dans le cadre de l'étude des systèmes d'IA et de PDA. Compte tenu de la nature sensible des plaintes et des données à caractère personnel des personnes, l'utilisation de services de cloud fournis par des entreprises technologiques des États-Unis ne peut pas être considérée comme une infrastructure informatique sûre[327].

En ce qui concerne le volume de ressources nécessaires, Equinet souligne que « les organismes de promotion de l'égalité sont les mieux placés pour déterminer l'ampleur des ressources nécessaires pour « accomplir chacune de leurs missions de manière efficace, dans un délai raisonnable », à partir de leur évaluation des problèmes à résoudre et des activités et mesures à mettre en œuvre » (Equinet 2024). Les questions liées aux outils pourraient également être abordées à un niveau collectif, par exemple via Equinet, qui dispose d'un forum pour ces échanges via le Standards Project pluriannuel, lequel consiste en des rencontres annuelles des membres intéressé.es pour échanger sur tous les aspects liés aux ressources, à l'indépendance et aux pouvoirs des OPE.

---

326. Equinet (2023), Measuring Standards for Equality Bodies: Indicators for Self-Assessment.
327. Voir les lois Cloud Act et Foreign Surveillance Intelligence Act aux États-Unis. Voir également l'article récent annonçant que Microsoft aurait résilié le compte e-mail du procureur de la Cour pénale internationale en réaction à une injonction du Président américain Donald Trump. Molly Quell (2025), Trump's sanctions on ICC prosecutor have halted tribunal's work, disponible en anglais sur https://apnews.com/article/icc-trump-sanctions-karim-khan-court-a4b4c02751ab84c09718b1b95cbd5db3, consulté le 11 novembre 2025.

Les directives sur les normes indiquent que les organismes de promotion de l'égalité peuvent utiliser des outils d'IA et de PDA pour accomplir leur mandat. Si ces outils peuvent permettre aux organismes de promotion de l'égalité d'être plus efficaces (par exemple, en détectant des constantes entre différentes plaintes), ils doivent être achetés, évalués et déployés avec soin, au regard de leurs instructions d'emploi.

6.2.3. Rôle des organismes de promotion de l'égalité vis-à-vis de leur mandat et de leurs ressources dans le contexte des systèmes d'IA et de PDA

Concernant leur mandat, et en particulier le concept de discrimination intersectionnelle

- Tirer parti des références à l'intersectionnalité pour s'attaquer aux discriminations souvent intersectionnelles potentiellement intégrées dans les systèmes d'IA et de PDA, dans le cadre de la recherche et des communications.
- Plaider en faveur de l'élargissement du mandat des organismes de promotion de l'égalité pour couvrir des motifs allant au-delà de la législation européenne en matière de discrimination.
- Prôner une application élargie de l'intersectionnalité, les directives actuelles confinant le travail sur l'intersectionnalité aux activités de prévention et de promotion.

Concernant les ressources

- Déterminer les ressources humaines et techniques nécessaires pour s'attaquer aux questions liées aux systèmes d'IA/PDA.
- Plaider pour la mise à disposition de ces ressources humaines et techniques, y compris dans le cadre de la transposition des directives sur les normes.
- Identifier les types d'outils techniques qui seront utiles et collaborer à l'échelle internationale pour les acquérir ou les développer. Par exemple, dans le contexte des logiciels de gestion des plaintes, prioriser les logiciels libres de droits qui peuvent être utilisés par différents OPE et être modifiés si nécessaire en fonction des besoins. Pour citer un autre exemple, on peut envisager le développement d'outils pour les enquêtes qui peuvent être utilisés par plusieurs OPE au lieu que chacun développe le sien de son côté[328].

6.3. Modifications des pouvoirs

Les directives sur les normes établissent des normes minimales harmonisées en ce qui concerne les pouvoirs des organismes de promotion de l'égalité, et leur donnent de nouveaux outils pour lutter contre la discrimination algorithmique, sur deux fronts :

328. Des initiatives similaires existent dans d'autres domaines, tels que la protection des données. Voir l'outil d'audit de sites web du Comité européen pour la protection des données, disponible en anglais sur www.edpb.europa.eu/our-work-tools/our-documents/support-pool-experts-projects/edpb-website-auditing-tool_en, consulté le 11 novembre 2025

- Pouvoirs de promotion de l'égalité de traitement et de prévention contre la discrimination
- Pouvoirs d'enquête et de réparation

La partie 6.3 s'intéresse particulièrement à ce que les dispositions des directives signifient dans le contexte de l'IA et des systèmes de PDA.

Les pouvoirs établis par les directives sur les normes sont liés aux obligations introduites par la Convention-cadre du Conseil de l'Europe vis-à-vis de la création ou du maintien à jour de voies de recours[329], de garanties procédurales[330] et de mécanismes de contrôle[331]. Il s'agit notamment « de contrôle continu des capacités actuelles de développement et d'audit, de consultations et de participation du public, de cadres de gestion des risques et de l'impact et les cadres d'évaluation de l'impact sur les droits de l'homme, les normes techniques ainsi que les programmes d'éducation et de sensibilisation »[332].

6.3.1. Promotion de l'égalité de traitement et prévention contre la discrimination

Les directives sur les normes introduisant la « promotion de l'égalité de traitement » comme compétence de l'OPE « exigent des États qu'ils permettent aux organismes de promotion de l'égalité de passer d'une approche réactive et corrective de la discrimination à une approche proactive, préventive et promotionnelle » (Equinet, 2024). Ce changement est essentiel dans le contexte des systèmes de prise de décision automatisée et d'IA, où le manque de sensibilisation et le manque d'informations constituent deux écueils rencontrés dans la prévention et la médiation de la discrimination.

Sensibilisation

L'article 5 (1) des directives sur les normes introduit un nouveau pouvoir général de sensibilisation pour les organismes de promotion de l'égalité. Dans le contexte des systèmes d'IA et de PDA, l'idée est de sensibiliser aux risques de discrimination causée par les systèmes d'IA et de PDA, aux droits à la non-discrimination existants et aux voies de recours et de réparation potentielles, y compris l'existence d'organismes de promotion de l'égalité. Ces démarches de sensibilisation devraient être accessibles[333] et une attention particulière devrait être accordée aux personnes en situation de vulnérabilité qui sont souvent la cible des systèmes d'IA et de PDA. Ces démarches de sensibilisation s'inscrivent dans la lignée des obligations de l'ASM visant à prêter une attention particulière aux systèmes d'IA qui présentent un risque pour des groupes vulnérables, et à informer l'OPE dès qu'un risque pour les droits fondamentaux est identifié[334].

329. Convention-cadre du Conseil de l'Europe, art. 14.
330. Convention-cadre du Conseil de l'Europe, art. 15.
331. Convention-cadre du Conseil de l'Europe, art. 9.
332. Conseil de l'Europe (2024), Rapport explicatif de 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, paragraphe 63.
333. Directives sur les normes, art. 12.
334. Règlement européen sur l'IA, art. 79 (2).

Prévention proactive

L'article 5, paragraphe 2, des directives sur les normes marque un autre changement par rapport aux directives sur l'égalité[335], en habilitant les organismes de promotion de l'égalité à jouer un rôle proactif dans la prévention de la discrimination. Les mesures prises à cet effet comprennent, entre autres, « promouvoir des mesures positives et l'intégration de la dimension de genre au sein des entités publiques et privées, à fournir à ces entités des formations, des conseils et un appui dans ce domaine, à participer au débat public, à communiquer avec les parties prenantes concernées, y compris les partenaires sociaux, et à promouvoir l'échange de bonnes pratiques[336] ».

Si l'on tient compte de l'article 4 des directives sur les normes concernant les ressources, cela signifie que les organismes de promotion de l'égalité doivent disposer de ressources suffisantes pour mener à bien ce type d'activités.

Consultation

L'article 15 des directives sur les normes exige que les États consultent les organismes de promotion de l'égalité dans le cadre de l'élaboration et de la mise en œuvre des lois, politiques, procédures et programmes relatifs aux droits et obligations découlant des directives sur l'égalité. Cet article pourrait être interprété de manière restrictive, comme se limitant à la législation et aux politiques en matière d'égalité. Toutefois, l'intention de la consultation, telle qu'elle est exposée dans les considérants, est de permettre aux organismes de promotion de l'égalité de contribuer à la « prise en considération systématique des questions d'égalité »[337]. À ce titre, les organismes de promotion de l'égalité devraient être impliqués dans la législation et les politiques numériques, notamment dans la mise en œuvre du règlement sur l'IA ou de la législation nationale sur les utilisations spécifiques de l'IA et des systèmes de PDA.

Cette disposition fait également écho à l'article 19 de la Convention-cadre du Conseil de l'Europe, qui prévoit la tenue de consultations publiques et multipartites sur les « questions importantes soulevées par les systèmes d'intelligence artificielle ».

---

335. Aux fins de ce rapport, l'expressive 'directives sur l'égalité' désigne les directives clés adoptées dans la législation de l'UE : la directive 2000/43/CE du 29 juin 2000 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 du 27 novembre 2000 portant création d'un cadre général en faveur de l'égalité de traitement en matière d'emploi et de travail; la directive 2006/54/CE du Parlement européen et du Conseil du 5 juillet 2006 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 2004/113/CE du 13 décembre 2004 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 2010/41/CE du Parlement européen et du Conseil du 7 juillet 2010 relative à l'application du principe de l'égalité de traitement entre hommes et femmes exerçant une activité indépendante et abrogeant la directive 86/613/CEE du Conseil; la directive 2023/970 du Parlement européen et du Conseil du 10 mai 2023 visant à renforcer l'application du principe de l'égalité des rémunérations entre les femmes et les hommes pour un même travail ou un travail de même valeur par la transparence des rémunérations et les mécanismes d'application du droit; la directive 79/7/CEE du Conseil du 19 décembre 1978 relative à la mise en œuvre progressive du principe de l'égalité de traitement entre hommes et femmes en matière de sécurité sociale.
336. Directive sur les normes, art. 5 (2).
337. Voir Equinet (2024), Understanding the New EU Directives on Standards for Equality Bodies: Legal Digest on Standards for Equality Bodies, pour une analyse similaire de la disposition.

Collecte de données, accès aux données relatives à l'égalité et élaboration de rapports

L'article 16 des directives sur les normes confère aux organismes de promotion de l'égalité le pouvoir de collecter des données (en mettant l'accent sur leur ventilation[338]), de mener des enquêtes[339], d'accéder aux statistiques collectées par d'autres[340], et de formuler des recommandations sur les données à collecter en matière d'égalité[341]. Ce pouvoir est explicitement lié aux obligations de rapport des organismes de promotion de l'égalité, qui comprennent la production de rapports sur « la situation en matière d'égalité de traitement [...] dans leur État membre »[342]. Les considérants précisent également les objectifs de la collecte de données, qui consistent à « quantifier la discrimination », « évaluer la mise en œuvre de la législation en matière d'égalité » et « contribuer à l'élaboration de politiques fondées sur des données probantes ».

Ces pouvoirs peuvent être appliqués au domaine des systèmes d'IA/PDA, notamment pour avoir une meilleure visibilité sur le déploiement des systèmes d'IA/PDA par les agences publiques.

La législation relative à la protection des données ne doit pas être considérée comme un obstacle à la collecte de données sur l'égalité (Agence européenne des droits fondamentaux, 2021). Le RGPD autorise la collecte et le traitement de catégories particulières de données à caractère personnel dans certaines conditions, notamment à des fins de statistique et de recherche, dans son article 9, paragraphe 2, points a), g) et j). Le Sous-groupe consacré aux données sur l'égalité du High Level Group on Non-discrimination, Equality and Diversity de la Commission européenne a adopté différentes séries de lignes directrices et publié des notes d'orientation au sujet de la collecte et de l'exploitation des données sur l'égalité. En mars 2025, ce sous-groupe a publié un document intitulé « Collecting and using equality data in full compliance with EU General Data Protection Regulation and national data protection rules », dans lequel il mettait en lumière les bonnes pratiques de plusieurs États membres de l'Union européenne, ainsi que des expériences confiées par l'Agence des droits fondamentaux et des membres du Sous-groupe (Subgroup on Equality Data, 2025; voir également Ilieva, 2024).

338. Directives sur les normes, art. 16 (1).
339. Directives sur les normes, art. 16 (2).
340. Directives sur les normes, art. 16 (3).
341. Directives sur les normes, art. 16 (4).
342. Directives sur les normes, art. 17 (c).

6.3.2. Rôle des organismes de promotion de l'égalité concernant la promotion de l'égalité de traitement et la prévention des discriminations

Sensibilisation

- Sensibiliser aux nouveaux droits des personnes concernées : dans le cadre de leurs missions de sensibilisation, les organismes de promotion de l'égalité devraient informer les personnes concernées des droits qui leur sont reconnus en lien avec l'utilisation des systèmes d'intelligence artificielle, tels que le droit d'introduire une réclamation, ainsi que le droit à l'explication prévu par le règlement sur l'intelligence artificielle et par le droit de l'Union en matière de protection des données.
- Sensibiliser aux risques pesant sur les droits fondamentaux, en s'appuyant sur les éléments de preuve existants relatifs aux systèmes d'intelligence artificielle aux niveaux national et européen.
- Engager un dialogue avec les médias et les journalistes, afin d'encourager de nouvelles investigations et une couverture médiatique accrue de ces enjeux.
- Collaborer avec les initiatives existantes en matière de sensibilisation : les actions de sensibilisation destinées au grand public peuvent notamment consister à intervenir dans des espaces déjà investis par celui-ci. À titre d'exemple, le projet européen AlgoLit, consacré à la médiation algorithmique, vise à travailler avec des médiateur·rices et des travailleur·euses sociaux·ales afin de sensibiliser aux enjeux liés à l'intelligence artificielle et aux systèmes de prise de décision automatisée dans le secteur public. Les organismes de promotion de l'égalité pourraient s'associer à ce type d'initiatives afin de garantir que les risques de discrimination y soient suffisamment discutés et de promouvoir le rôle de ces organismes dans la mise en œuvre effective du droit de la non-discrimination. Cette collaboration apparaît d'autant plus nécessaire que les formations portant sur les défis liés à l'intelligence artificielle adoptent fréquemment une approche fondée sur l'éthique plutôt que sur les droits fondamentaux (Xenidis, 2025).
- Contrôler les stratégies des fournisseurs et déployeurs de systèmes d'IA, notamment les stratégies de « dérisquage » (Xenidis, 2025) et d'« éthique-washing » (Equinet, 2025) employées par les fournisseurs et les déployeurs de systèmes d'IA/PDA (par exemple : auto-exclusion de la catégorie de systèmes à haut risque, ou non-prise en considération de l'interdiction de certains systèmes en vertu de l'article 5).

Prévention proactive

- Former les déployeurs privés et publics de systèmes d'IA/PDA (y compris les agences gouvernementales), ainsi que les institutions législatives et judiciaires. Par exemple, le Conseil de l'Europe propose un cours consacré à l'IA et la discrimination dans plusieurs contextes nationaux, en coopération avec des organismes de promotion de l'égalité (dont le Médiateur de la non-discrimination en Finlande, le Défenseur des droits en France, la Commission pour la citoyenneté et l'égalité de genre au Portugal, et l'Unia en Belgique).
- Promouvoir l'inclusion et l'égalité dans les disciplines de sciences, technologie, ingénierie et mathématiques.

Collecte de données, accès aux données sur l'égalité et production de rapports

- Identifier les discriminations liées à l'IA/PDA et la mise en œuvre du règlement sur l'IA comme des domaines prioritaires pour la production de rapports publics.
- Rédiger des rapports fondés sur des études nationales concernant le recours aux systèmes d'IA/PDA dans un domaine particulier, en se concentrant sur les domaines considérés comme à haut risque en vertu de l'annexe III du règlement sur l'IA. Ces études documentaires peuvent être réalisées à partir des informations disponibles dans la base de données de l'UE sur les systèmes d'IA à haut risque, mais vont au-delà et incluent une analyse des systèmes qui ne sont pas considérés comme à haut risque ou qui ne figurent pas dans la base de données publique (tels que les systèmes répressifs, de migration et de contrôle des frontières), afin de démontrer qu'il est important de ne pas se limiter aux systèmes d'IA à haut risque. Ces rapports peuvent également être réalisés sur plusieurs pays afin d'identifier des constantes et des tendances communes.
- Mener des enquêtes sur les perceptions du public à l'égard des systèmes d'IA et de PDA et sur leur compréhension des risques liés à la discrimination. Idéalement, ces enquêtes pourraient être conduites dans plusieurs pays afin de comparer les contextes nationaux et les réponses. En Suède, le médiateur pour l'égalité a été chargé par le gouvernement suédois d'améliorer l'état des connaissances concernant les risques de discrimination associés au recours à l'intelligence artificielle et à d'autres types de prise de décision automatisée dans des contextes professionnels (ministère de l'Emploi, décision du gouvernement 07/06/2022, A2022/00877), ce qui a donné lieu à la publication d'un rapport en novembre 2023. Ce rapport évaluait le niveau de connaissance des employeurs concernant leur utilisation de l'intelligence artificielle, et démontrait que les employeurs ne savent pas qu'ils ont déjà recours à l'IA et à d'autres types de prise de décision automatisée[343].
- Travailler en étroite collaboration avec les agences de statistiques pour s'assurer que des données ventilées sont collectées et disponibles afin d'analyser techniquement les biais dans les systèmes d'IA/PDA.

6.3.3. Accès à la justice et aux voies de recours

Pour répondre à la problématique des disparités actuelles qui existent dans les différentes manières dont les organismes de promotion de l'égalité promeuvent l'accès à la justice, les directives sur les normes confèrent des pouvoirs explicites aux organismes de promotion de l'égalité pour agir en faveur des victimes et faciliter l'accès à la justice. Ces pouvoirs comprennent l'assistance aux victimes[344], les modes alternatifs de règlement des litiges[345], la conduite d'enquêtes[346], la publication d'avis

343. Diskriminerings ombudsmannen (2023), « AI och risker för disrkiminering I arbetslivet », disponible à l'adresse : www.do.se/rattsfall-beslut-lagar-stodmaterial/publikationer/2023/ai-och-risker-for-diskriminering-i-arbetslivet [en suédois], consulté le 11 novembre 2025.
344. Directives sur les normes, art. 6
345. Directives sur les normes, art. 7.
346. Directives sur les normes, art. 8.

et de décisions[347], et les actions en justice[348]. Cette partie du document aborde des dispositions spécifiques.

Assistance aux victimes et capacités à recevoir des plaintes

L'article 6 des directives sur les normes prévoit que les organismes de promotion de l'égalité donnent aux victimes « des informations sur ce qui suit : le cadre juridique, y compris des conseils adaptés à leur situation spécifique; les services offerts par [organismes de promotion de l'égalité] et les aspects procéduraux connexes; les voies de recours disponibles, y compris la possibilité d'intenter une action en justice; les règles de confidentialité applicables et la protection des données à caractère personnel; et la possibilité d'obtenir un soutien psychologique ou autre de la part d'autres organismes ou organisations ».

Dans le contexte des systèmes d'IA/PDA, cela implique d'informer les victimes de leurs droits en vertu du règlement sur l'IA et de la législation européenne sur la protection des données.

Enquêtes

L'article 8 des directives sur les normes autorise les organismes de promotion de l'égalité à mener des enquêtes, soit à la suite d'une plainte, soit de leur propre initiative pour enquêter sur les violations potentielles du droit à l'égalité de traitement). Comme le souligne l'ECRI, « [c]es enquêtes sont importantes pour rassembler les éléments qui, enfin, permettent de remédier à ces situations »[349]. Ce dernier mode est particulièrement important dans le contexte des systèmes d'IA/PDA, qui sont toujours opaques et peu connus.

L'article 8 (2) des directives sur les normes donne aux organismes de promotion de l'égalité un accès effectif aux informations et aux documents et est pertinent en ce qui concerne le règlement sur l'IA, pour deux raisons :

- Pour les organismes de promotion de l'égalité qui ne sont pas (encore) des organismes relevant de l'article 77, cette disposition pourrait être envisagée comme un moyen d'accéder aux informations et aux documents couverts par les règles de confidentialité, y compris la documentation technique ou les versions complètes des analyses d'impact sur les droits fondamentaux.
- Le droit d'accès à l'information et aux documents est souvent entravé par la difficulté d'accéder à des informations compréhensibles. Comme le règlement sur l'IA prévoit que les informations créées en vertu du règlement doivent être accessibles, les organismes de promotion de l'égalité peuvent tirer parti de cette disposition et se concentrer sur les demandes d'accès aux informations et aux documents produits dans le cadre du règlement, et faire ainsi en sorte qu'ils obtiennent des informations compréhensibles et utiles (à condition que l'obligation d'accessibilité soit clarifiée et respectée).

347. Directives sur les normes, art. 9.
348. Directives sur les normes, art. 10.
349. ECRI, Recommandation de politique générale n° 2r

L'article 8, paragraphe 3, des directives sur les normes permet aux organismes de promotion de l'égalité de confier à un autre organisme compétent, conformément à la législation et aux pratiques nationales, les pouvoirs susmentionnés. Cet article ouvre la possibilité d'une coopération, mais pas d'une supplantation des organismes de promotion de l'égalité, car le libellé de l'article suggère que ce pouvoir sera exercé par un autre organisme « en plus de" et non « à la place de »[350]. Cette disposition peut être envisagée comme un moyen d'instaurer des coopérations formelles avec d'autres organismes.

Avis et décisions

L'article 9 des directives sur les normes permet également d'améliorer la compréhension et la connaissance des systèmes d'IA/PDA dans le secteur public. Il prévoit que « le cas échéant, les avis non contraignants et les décisions contraignantes comprennent des mesures spécifiques visant à remédier à toute violation du principe de l'égalité de traitement constatée et à empêcher qu'une telle situation ne se reproduise », et que les organismes de promotion de l'égalité « publient au moins un résumé de ceux de leurs avis et décisions qu'ils considèrent comme particulièrement pertinents ».

Litiges

L'article 10 de la directive relative aux normes confère aux organismes de promotion de l'égalité le pouvoir non transférable d'agir dans le cadre de procédures juridictionnelles, lequel doit comprendre « au moins l'un des éléments suivants : le droit d'engager une procédure juridictionnelle au nom d'une ou de plusieurs victimes; le droit de participer à une procédure juridictionnelle à l'appui d'une ou de plusieurs victimes; ou le droit d'engager une procédure juridictionnelle en son nom propre, afin de défendre l'intérêt public ».

Les litiges relatifs aux systèmes d'IA et de PDA sont encore rares, et des initiatives ou un soutien des organismes de promotion de l'égalité dans ce domaine sont nécessaires.

350. Voir Equinet (2024), Understanding the New EU Directives on Standards for Equality Bodies: Legal Digest on Standards for Equality Bodies, p.84 pour une analyse similaire de la disposition.

6.3.4. Rôle des organismes de promotion de l'égalité vis-à-vis de la justice et des voies de recours

Assistance aux victimes et capacité à recevoir des plaintes

- Veiller à ce que le personnel des organismes de promotion de l'égalité chargé de l'assistance aux victimes disposent de connaissances actualisées en matière de réglementation relative à l'IA, afin d'informer correctement les victimes de leurs droits.
- Se coordonner avec les autorités de surveillance du marché (ASM) pour veiller à ce que les victimes qui déposent une plainte auprès d'ASM soient encouragées à déposer une plainte auprès d'OPE.

Enquêtes

- Mener des enquêtes de leur propre initiative sur des systèmes d'IA/PDA spécifiques.
- Étudier la transposition nationale de l'article 8 des directives sur les normes comme potentiel moyen d'accéder aux informations produites dans le cadre du règlement sur l'IA.

Avis et décisions

- Dans les avis et les décisions, proposer des mesures visant à faire connaître les systèmes. Par exemple, dans l'arrêt n° 2949/2020 du 31 décembre 2020, après avoir condamné Deliveroo à payer 50 000 euros en faveur du plaignant, le tribunal de Bologne a ordonné la publication du texte de la sentence dans un journal national dans le but de garantir une visibilité maximale[351].
- Publier systématiquement les avis et les décisions concernant les systèmes d'IA/PDA afin de rendre publiques les connaissances relatives à leurs utilisations et aux risques qu'ils présentent.

Litiges

- Intervenir dans les procédures judiciaires relatives aux systèmes d'IA/PDA.

351. Fernandez Sánchez S. F. (2021), « Frank, el algoritmo consciente de Deliveroo. Comentario a la Sentencia del Tribunal de Bolonia 2949/2020, de 31 de diciembre", Revista De Trabajo Y Seguridad Social CEF, pp. 179-93, disponible en espagnol à l'adresse https://doi.org/10.51302/rtss.2021.2374, consulté le 11 novembre 2025.




7. Thématique centrale

Cette partie concerne le recours à l'IA dans cinq domaines thématiques en fonction des secteurs les plus couverts par les travaux des organismes de promotion de l'égalité et des SNDH : application de la loi, migration, asile et contrôle des frontières; éducation; emploi, et sécurité sociale et services d'aide à l'emploi. Chaque domaine thématique présente des exemples d'utilisation de l'IA et recense les articles du règlement européen sur l'IA à prendre en considération pour le domaine en question.

7.1. Thématique centrale : Activités répressives, migration, asile et contrôle des frontières

7.1.1. Contexte

Utilisation des systèmes d'IA dans les domaines de la répression, de la migration, de l'asile et du contrôle des frontières

L'IA et la PDA sont utilisées dans toute une série de contextes au sein des domaines des activités répressives, de la migration, de l'asile et du contrôle des frontières.

Dans le domaine répressif, des recherches antérieures ont mis en évidence des utilisations allant de la cartographie des schémas récurrents de criminalité à partir de données antérieures, à la détection d'objets illicites à partir d'images satellite, en passant par la détection de discours de haine en ligne[352], la prise de décision concernant les libérations temporaires dans les prisons[353] ou encore l'évaluation des risques de violences sexistes[354].

La reconnaissance d'images et la biométrie sont également très répandues. En Finlande, la police peut recourir à la reconnaissance d'images pour identifier des caractéristiques non biométriques (ex. : vêtements, plaques d'immatriculation, etc.) (Xenidis, 2025). La France a eu recours à la télésurveillance lors des Jeux olympiques et paralympiques d'été de Paris 2024[355]. L'application de reconnaissance faciale commercialisée par la société américaine Clearview AI, qui repose sur le moissonnage d'images de visages sur les réseaux sociaux, aurait été utilisée par plusieurs services répressifs dans toute l'Europe (Veld et al. 2020).

La recherche démontre que les technologies de reconnaissance faciale peuvent présenter des biais discriminatoires problématiques[356]. 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. Cela peut se produire lorsque la surveillance excessive de lieux ou de populations spécifiques se répercute sur les outils de prévision et renforce la surveillance de masse ou l'évaluation des risques associés à ces communautés. Cela peut également se produire lorsque ces systèmes sont utilisés de manière différente en fonction de la couleur de peau de la personne : par exemple, des policiers qui interprètent les résultats d'un système différemment en fonction de la couleur de peau du suspect.

Les systèmes d'IA et de PDA sont également utilisés dans les domaines de la migration, de l'asile et du contrôle des frontières, dans différents secteurs (Dumbrava 2025, Jones et al. 2023, McGregor 2023). En 2023, un rapport produit dans le cadre du projet

352. Voir Agence des droits fondamentaux de l'Union européenne (2022), Bias in Algorithms: Artificial Intelligence and Discrimination, Luxembourg : Office des publications de l'Union européenne, pp.28-49.
353. García, T., Torrecillas, C., Maqueda, A., Cabo, D., Laursen, L. (2025), Spanish prisons use a 30-year-old algorithm to decide on temporary releases, Civio, disponible sur https://civio.es/justicia/2025/03/12/spanish-prisons-use-a-30-year-old-algorithm-to-decide-on-temporary-releases/.
354. Public Sector Tech Watch (2025), Viogen 5.0 : discovering Spain's risk assessment system of gender-based violence, disponible à l'adresse https://interoperable-europe.ec.europa.eu/collection/public-sector-tech-watch/viogen-50-discovering-spains-risk-assessment-system-gender-based-violence.
355. Ministère de l'Intérieur (2025), Expérimentation, en temps réel, de caméras « augmentées », disponible sur https://www.interieur.gouv.fr/actualites/actualites-du-ministere/experimentation-en-temps-reel-de-cameras-augmentees
356. L'OSC Liberty a attaqué une application de reconnaissance faciale utilisée par la police du Pays de Galles du Sud, entre autres pour discrimination fondée sur le sexe et/ou la race, parce qu'elle 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) v. 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. » Voir Xenidis 2025.

Algorithmic fairness for asylum seekers and refugees (AFAR) recensait les utilisations de technologies d'IA dans les domaines de la migration et de l'asile en Europe. Ce rapport avait identifié les utilisations suivantes : prévisions des mouvements d'immigration et de déplacement futurs; analyses de risques et triage; traitement des visas, autorisations de séjour et demandes de nationalité; contrôle de documents à des fins de vérification d'identité et de détection des fraudes : reconnaissance vocale (pour déterminer le pays d'origine des demandeurs d'asile ou évaluer le niveau de maîtrise de la langue dans les demandes de nationalité); surveillance électronique (ex. : bracelets électroniques); attribution de prestations sociales; outils d'attribution (par exemple pour l'attribution de centres d'accueil); et extraction de données de téléphones portables pour vérifier l'identité de la personne et son récit (Ozkul, 2023).

Ces utilisations présentent des risques pour les droits fondamentaux, notamment en raison des problématiques d'inexactitude, de biais et de stéréotypes qui peuvent être intégrées dans ces outils, mais également au vu des incidences en matière d'égalité et de procédure, et des problèmes impactant le respect de la vie privée et la protection des données (Dumbrava 2025, McGregor 2023). En 2020, suivant un recours judiciaire engagé par Foxglove et le Joint Council for the Welfare of Immigrants (conseil conjoint pour le bien-être des immigrés), le ministère de l'intérieur britannique a cessé d'utiliser un algorithme de fouille de flots de données qu'il employait dans le cadre des demandes de visa et qui attribuait un score de risque rouge aux demandeurs possédant certaines nationalités (BBC, 2020). Dans un cas étudié par Equinet, le Royaume-Uni utilisait un système de PDA pour déterminer l'éligibilité des personnes au statut de résident permanent (« Settled Status »), qui avait été mis en place pour régulariser le statut d'immigration des ressortissants de l'UE, de l'Espace économique européen et de la Suisse, ainsi que de leurs familles vivant au Royaume-Uni après le Brexit. Le travail d'investigation d'Equinet a mis en lumière plusieurs problèmes au niveau du système, dont son opacité, un certain niveau d'incertitude quant à la discrétion humaine et l'absence d'interrogation de bases de données spécifiques, le tout entraînant ainsi des effets préjudiciables pour les femmes (Allen et Masters 2020). En avril 2024, l'autorité grecque de protection des données a infligé une amende de 175 000 € au ministère chargé de la migration et de l'asile pour violations commises dans la mise au point et le déploiement des programmes Centaur et Hyperion au sein des structures d'accueil et d'hébergement de demandeurs d'asile. Centaur était un système de surveillance partiellement automatisé qui visait à prédire et signaler les « menaces » en se basant, entre autres, sur des images de vidéosurveillance et des drones; et Hyperion était un système de contrôle des entrées/sorties (Hellenic Data Protection Authority, 2024).

Un large éventail d'exemptions

L'une des caractéristiques du règlement sur l'IA est qu'elle accorde un large éventail d'exemptions aux systèmes d'IA utilisés dans le cadre des domaines de la répression, de la migration, de l'asile et du contrôle aux frontières. Par exemple, certains systèmes d'IA interdits pour certaines utilisations ne sont considérés comme à haut risque que dans le contexte d'activités répressives, comme c'est le cas de l'interdiction de la catégorisation biométrique.[357] En outre, ils sont soumis à des obligations d'enregistrement et de

357. Règlement sur l'IA, art. 5 (1) (g) et annexe III (1) (b).

publicité moindres par rapport aux autres systèmes d'IA à haut risque (voir article 49), ce qui suscite des préoccupations constantes en matière de transparence.

Les équipes des organismes de promotion de l'égalité chargées du domaine répressif devraient porter un regard particulier sur les dispositions suivantes du règlement sur l'IA.

7.1.2. Systèmes d'IA interdits

Article 5 : Systèmes d'IA interdits

- Article 5, paragraphe 1, point c), sur la notation sociale concernant les systèmes de classification et d'évaluation. Par exemple, des activités d'évaluation et de classification pourraient avoir lieu dans le contexte d'une incarcération ou d'un camp de réfugiés, sur la base d'images de vidéosurveillance. Ces pratiques pourraient être considérées comme de la notation sociale dans certaines circonstances.
- Article 5, paragraphe 1, point d) sur l'évaluation du risque de commettre une infraction pénale dans certaines circonstances, qui couvre certaines utilisations de la prévision policière.
- Article 5, paragraphe 1, point e) sur le moissonnage destiné à constituer ou enrichir des bases de données de reconnaissance faciale.
- Article 5, paragraphe 1, point h), sur l'identification biométrique à distance pour les services répressifs.

Plusieurs interdictions n'englobent pas les utilisations dans le domaine répressif, qui sont toutefois considérées comme à haut risque en vertu de l'annexe III :

- Article 5, paragraphe 1, point g) : l'interdiction de la catégorisation biométrique ne concerne pas les utilisations à des fins répressives, qui sont toutefois considérées comme à haut risque en vertu de l'annexe III, paragraphe 1, point b).
- Article 5, paragraphe 1, point f), sur la reconnaissance des émotions : cet article ne concerne pas les utilisations dans les domaines de la répression, des migrations et du contrôle des frontières. Toutefois, il est considéré comme présentant à haut risque au titre de l'annexe III, paragraphe 1, point c). Par exemple, la reconnaissance des émotions peut être utilisée dans le contexte de « détecteurs de mensonges » lors de l'interrogatoire de suspects.

7.1.3. Systèmes d'IA à haut risque relevant de l'annexe III

- Annexe III (1) sur la biométrie :
  - Systèmes d'identification biométrique à distance, à l'exception de ceux dont la seule finalité « est de confirmer qu'une personne physique spécifique est la personne qu'elle prétend être »[358] ;
  - Pour la catégorisation biométrique, en fonction des attributs ou des caractéristiques sensibles ou protégés, sur la base de la déduction de ces attributs ou caractéristiques[359] ;

358. Règlement sur l'IA, annexe III (1) (a).
359. Règlement sur l'IA, annexe III (1) (b).

- Pour la reconnaissance des émotions[360];

- Annexe III (1) (6) sur les utilisations « par les autorités répressives ou par les institutions, organes et organismes de l'Union, ou en leur nom, en soutien aux autorités répressives ou en leur nom ». Ces utilisations comprennent :
  - le fait d'évaluer « le risque qu'une personne physique soit victime d'infractions pénales. » Un exemple pourrait être le système VioGen mentionné plus haut, utilisé en Espagne dans le contexte des violences sexistes[361];
  - les « polygraphes ou outils similaires[362] »;
  - le fait d'» évaluer la fiabilité des preuves au cours d'enquêtes ou de poursuites pénales[363] »;
  - les évaluations du risque d'infraction ou de récidive qui ne sont pas fondées « uniquement sur le profilage des personnes physiques » (rappel : les évaluations des risques fondées uniquement sur le profilage sont interdites en vertu de l'article 5, paragraphe 1, point d)), ou qui visent à « évaluer les traits de personnalité, les caractéristiques ou les antécédents judiciaires de personnes physiques ou de groupes »[364];
  - « pour le profilage de personnes physiques [...] dans le cadre de la détection d'infractions pénales, d'enquêtes ou de poursuites en la matière »[365].
- Annexe III (1) (7) sur la migration, l'asile et la gestion des contrôles aux frontières :
  - les « polygraphes ou outils similaires »[366];
  - le fait d'« évaluer un risque, y compris un risque pour la sécurité, un risque de migration irrégulière ou un risque pour la santé, posé par une personne physique qui a l'intention d'entrer ou qui est entrée sur le territoire d'un État membre »[367];
  - « l'examen des demandes d'asile, de visas et de titres de séjour et des plaintes connexes au regard de l'objectif visant à établir l'éligibilité des personnes physiques demandant un statut, y compris les évaluations connexes de la fiabilité des éléments de preuve »[368];
  - « la détection, de la reconnaissance ou de l'identification des personnes physiques, à l'exception de la vérification des documents de voyage »[369].

Rappel : toutes ces utilisations sont autorisées « dans la mesure où leur utilisation est autorisée par le droit de l'Union ou le droit national applicable ». Lorsqu'une utilisation porte atteinte aux droits humains et/ou est discriminatoire, il faut déterminer si les mesures prises sont adaptées et nécessaires dans une société démocratique. Par le passé, la Cour européenne des droits de l'homme a estimé que le recours

360. Règlement sur l'IA, annexe III (1) (c).
361. Règlement sur l'IA, annexe III (6) (a).
362. Règlement sur l'IA, annexe III (6) (b).
363. Règlement sur l'IA, annexe III (6) (c).
364. Règlement sur l'IA, annexe III (6) (d).
365. Règlement sur l'IA, annexe III (6) (e).
366. Règlement sur l'IA, annexe III (7) (a).
367. Règlement sur l'IA, annexe III (7) (b).
368. Règlement sur l'IA, annexe III (7) (c).
369. Règlement sur l'IA, annexe III (7) (d).

aux nouvelles technologies dans la police ne pouvait être considéré comme nécessaire dans une société démocratique, dans le contexte du profilage d'ADN[370] et de la reconnaissance faciale en direct[371]. Dans l'affaire S. et Marper, le tribunal notait :

> [L]a protection offerte par l'article 8 de la Convention serait affaiblie de manière inacceptable si l'usage des techniques scientifiques modernes dans le système de la justice pénale était autorisé à n'importe quel prix et sans une mise en balance attentive des avantages pouvant résulter d'un large recours à ces techniques, d'une part, et des intérêts essentiels s'attachant à la protection de la vie privée, d'autre part. [...] La Cour considère que tout État qui revendique un rôle de pionnier dans l'évolution de nouvelles technologies porte la responsabilité particulière de trouver le juste équilibre en la matière.

Pour certains systèmes d'IA, les fournisseurs ont la possibilité de s'exclure eux-mêmes du régime à haut risque conformément à l'article 6, paragraphe 3. Par exemple, ce pourrait être le cas pour les outils de traduction par IA utilisés pour étudier les demandes d'asile. Cet outil pourrait être considéré comme une « tâche préparatoire » à une évaluation et, par conséquent, ne pas entrer dans le périmètre des systèmes d'IA à haut risque. Les performances limitées de ces outils dans certaines langues[372] les rend susceptibles d'avoir des effets discriminatoires sur les locuteurs de certaines langues.

7.1.4. Obligations d'enregistrement et base de données du règlement européen sur l'IA

Les systèmes d'IA à haut risque utilisés dans les domaines répressif, de la migration, de l'asile et du contrôle des frontières sont soumis à des obligations d'enregistrement moins strictes en vertu du règlement sur l'IA. Concrètement, leurs utilisations seront enregistrées dans la version non publique de la base de données de l'UE, et les synthèses des AIDF ne seront pas enregistrées dans la base de données.

Il est recommandé aux organismes de promotion de l'égalité et aux SNDH d'accorder la priorité à la mise en place de mécanismes de contrôle permanent de ces utilisations, au titre des pouvoirs qui leur sont conférés en tant qu'organismes relevant de l'article 77 (le cas échéant), ou en coopérant avec les autorités de surveillance du marché.

7.2. Thématique centrale : Éducation

7.2.1. Contexte

Les systèmes d'IA et de PDA peuvent être utilisés dans divers domaines de l'éducation, allant des admissions (voir le système Parcoursup utilisé en France pour classer les candidats et faire correspondre l'offre et la demande dans l'enseignement

370. S. et Marper c. Royaume-Uni, Requêtes n° 30562/04 et 30566/04 (2008), voir https://hudoc.echr.coe.int/fre?i=001-90052, consulté le 11 novembre 2025.
371. Glukhin c. Russia, Requête n° 11519/20 (2023), voir https://hudoc.echr.coe.int/eng?i=001-225817, consulté le 11 novembre 2025.
372. Bhuiyan, J. (2023), Lost in AI translation : growing reliance on language apps jeopardizes some asylum applications, The Guardian, disponible à l'adresse suivante https://www.theguardian.com/us-news/2023/sep/07/asylum-seekers-ai-translation-apps, consulté le 11 novembre 2025.

supérieur[373]), à l'évaluation des résultats d'apprentissage (par exemple dans le contexte des examens finaux de l'enseignement secondaire, au Royaume-Uni en 2020[374]), aux activités administratives, à l'apprentissage en classe et au suivi des étudiants. On note un intérêt croissant porté au développement de technologies pour les élèves ayant des besoins éducatifs particuliers et de systèmes de management conçus pour détecter de nombreux risques, allant de risques suicidaires aux affinités avec des groupes terroristes. En 2023, l'Institut néerlandais des droits humains a jugé qu'une université néerlandaise n'avait pas fait preuve de discrimination raciale à l'encontre d'un étudiant en utilisant un logiciel anti-triche[375]. L'étudiant avait avancé que le logiciel était discriminatoire parce que la reconnaissance faciale était moins performante pour les étudiants et étudiantes à la peau foncée, en raison de limitations techniques.

D'autres technologies sont utilisées dans les contextes éducatifs. En France, l'autorité de protection des données s'est opposée à l'expérimentation de la reconnaissance faciale à l'entrée de deux établissements d'enseignement secondaire, au motif que le recours à la reconnaissance faciale n'était ni proportionné ni nécessaire au regard de la législation sur la protection des données[376]. Le recours à un système de reconnaissance faciale dans des écoles en Suède a été jugé préjudiciable pour les droits de protection des données et les droits plus généraux à la vie privée et à l'intégrité[377]. De même, une obligation de fournir des éléments biométriques à des fins d'identification et de règlement de repas a été jugée illégale en Pologne au motif qu'il n'y avait pas de fondement juridique pour ces mesures[378].

7.2.2. Utilisations interdites

Les utilisations interdites suivantes présentent un intérêt particulier dans le contexte de l'éducation :

- L'article 5, paragraphe 1, points a) et b), sur les techniques trompeuses,
- L'article 5, paragraphe 1, point c) sur la notation sociale,
- L'article 5, paragraphe 1, point f), qui interdit la reconnaissance des émotions dans les établissements d'enseignement,

373. AI Law Hub (2020), French Parcoursup decision, disponible su rhttps://ai-lawhub.com/2020/04/16/french-parcoursup-decision/, consulté le 11 novembre 2025.
374. Office for Statistics Regulation Authority (2021), Ensuring statistical models command public confidence : Learning lessons from the approach to developing models for awarding grades in the UK in 2020, disponible sur https://osr.statisticsauthority.gov.uk/publication/ensuring-statistical-models-command-public-confidence/, consulté le 11 novembre 2025.
375. Racism and Technology Center (2023), Judgement of the Dutch Institute for Human Rights shows how difficult it is to legally prove algorithmic discrimination, disponible sur https://racismandtechnology.center/2023/10/17/judgement-of-the-dutch-institute-for-human-rights-shows-how-difficult-it-is-to-legally-prove-algorithmic-discrimination/, consulté le 11 novembre 2025.
376. CNIL (2022), Expérimentation de la reconnaissance faciale dans deux lycées : la CNIL précise sa position, disponible sur https://www.cnil.fr/fr/experimentation-de-la-reconnaissance-faciale-dans-deux-lycees-la-cnil-precise-sa-position, consulté le 11 novembre 2025
377. KamR Stockholm, affaire n° 5888-20, disponible en anglais sur : https://gdprhub.eu/index.php?title=KamR_Stockholm_-_Case_No._5888-20#, consulté le 11 novembre 2025.
378. EDPB (5 mars 2020), 'Fine for processing student's fingerprints imposed on a school'. Voir https://edpb.europa.eu/news/national-news/2020/fine-processing-students-fingerprints-imposed-school_en [en anglais], consulté le 11 novembre 2025.

- L'article 5, paragraphe 1, point g), qui interdit la catégorisation biométrique « 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 ».

7.2.3. Systèmes d'IA à haut risque

- L'annexe III (1) sur la biométrie, et en particulier :
  - L'annexe III, paragraphe 1, point a), sur les systèmes d'identification biométrique à distance[379]
  - L'annexe III, paragraphe 1, point b), qui concerne la catégorisation biométrique en fonction d'attributs ou de caractéristiques sensibles ou protégés, sur la base d'une déduction de ces attributs ou caractéristiques (il s'agit d'une catégorie plus large que celle des pratiques interdites par l'article 5, paragraphe 1, point g))
- L'annexe III (3) sur l'éducation et la formation professionnelle, qui couvre :
  - les « systèmes d'IA destinés à être utilisés pour déterminer l'accès, l'admission ou l'affectation de personnes physiques à des établissements d'enseignement et de formation professionnelle, à tous les niveaux »;
  - les « systèmes d'IA destinés à être utilisés pour évaluer les acquis d'apprentissage »;
  - les « systèmes d'IA destinés à être utilisés pour évaluer le niveau d'enseignement approprié qu'une personne recevra ou sera en mesure d'atteindre »;
  - les » systèmes d'intelligence artificielle destinés à être utilisés pour surveiller et détecter des comportements interdits chez les étudiants lors d'examens ».

7.2.4. Exigences de transparence

- L'article 50 sur les exigences de transparence est particulièrement pertinent dans le contexte de l'éducation, où l'IA générative et les technologies éducatives sont de plus en plus utilisées en classe pour interagir directement avec les étudiant.es.

7.2.5. Exigences en matière d'enregistrement

Les systèmes à haut risque dans le domaine de l'éducation seront enregistrés dans la version publique de la base de données par les fournisseurs et les déployeurs. Les informations enregistrées seront notamment les synthèses des AIDF menées par les déployeurs.

7.3. Thématique centrale : Emploi

7.3.1. Contexte

Les systèmes d'IA et de PDA peuvent être utilisés à différents stades du processus de recrutement et d'emploi, de la rédaction des offres d'emploi et du ciblage des destinataires des offres jusqu'au processus de recrutement (par exemple, pour traiter les CV, mener des entretiens ou mettre en relation des candidat.es avec des emplois),

379. Règlement sur l'IA, annexe III (1).

en passant par la gestion et l'évaluation des performances[380]. En 2024, 98,4 % des entreprises du classement Fortune 500 utilisaient l'IA ou des systèmes pilotés par les données lors de la phase de recrutement[381].

Souvent, les systèmes d'IA sont utilisés par les employeur.ses sous supervision humaine. Cette confirmation doit être examinée de près, car le contrôle humain doit remplir certaines conditions pour être efficace, et n'est jamais, à lui seul, une garantie appropriée pour les droits fondamentaux[382].

7.3.2. Pratiques interdites

- Article 5, paragraphe 1, point f) sur la reconnaissance des émotions sur le lieu de travail, à laquelle on pourrait avoir recours lors du processus de recrutement ou pour surveiller les émotions des employés.
- Article 5, paragraphe 1, point g), sur la catégorisation biométrique « 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 », qui pourraient intéresser les recruteur.ices ou les employeur.ses.

7.3.3. Pratiques à haut risque

- Annexe III (1) sur la biométrie, et en particulier :
  - Annexe III, paragraphe 1, point a), sur les systèmes d'identification biométrique à distance
  - Annexe III, paragraphe 1, point b), qui concerne la catégorisation biométrique en fonction d'attributs ou de caractéristiques sensibles ou protégés, sur la base de la déduction de ces attributs ou caractéristiques (catégorie plus large que celle des pratiques interdites par l'article 5, paragraphe 1, point g))
- Annexe III (4) sur les zones d'emploi, la gestion de la main d'œuvre et l'accès à l'emploi indépendant. Cela comprend les systèmes ayant les finalités suivantes :
  - « le recrutement ou la sélection de personnes physiques, notamment pour diffuser des offres d'emploi ciblées, analyser et filtrer les candidatures et évaluer les candidats »[383];
  - « la prise de décisions affectant les conditions des relations professionnelles, ainsi que la promotion et la résiliation des relations professionnelles contractuelles, [et les décisions] pour l'attribution de tâches fondée sur le comportement individuel, les traits de personnalité ou les caractéristiques

380. Simons J. (2020), Machine Learning at Work : Case Studies, Institute for the Future of Work, disponible à l'adresse : https://www.ifow.org/publications/2020/2/24/machine-learning-case-studies, consulté le 11 novembre 2025.
381. Jobscan (2024), 2024 Applicant Tracking System (ATS) Usage Report : Key Shifts and Strategies for Job Seekers, disponible sur https://www.jobscan.co/blog/fortune-500-use-applicant-tracking-systems, consulté le 11 novembre 2025.
382. Green, B. (2022), The flaws of policies requiring human oversight of government algorithms, Computer Law & Security Review, 45, p. 105681. https://doi.org/10.1016/j.clsr.2022.105681, consulté le 11 novembre 2025.
383. Règlement sur l'IA, annexe III (4) (a).

personnelles et pour le suivi ou l'évaluation des personnes dans le cadre de relations professionnelles contractuelles ».

Il est probable que certains des systèmes utilisés dans ces domaines ne seront pas considérés comme étant à haut risque par les fournisseurs au titre de l'article 6, paragraphe 3. Par exemple, un système d'IA utilisé pour évaluer les CV et les lettres de motivation pourrait être considéré comme une « tâche préparatoire à une évaluation ».

7.3.4. Exigences de transparence

- L'article 50 sur les exigences de transparence est particulièrement pertinent dans le contexte de l'emploi, notamment au stade du recrutement, où de plus en plus d'outils d'IA sont utilisés dans le processus d'embauche.

7.3.5. Obligations d'enregistrement et d'information

L'article 26, paragraphe 7, prévoit une obligation d'information pour les déployeurs qui sont des employeurs. Avant de mettre en service ou d'utiliser un système d'IA à haut risque sur le lieu de travail, ils doivent informer « les représentants des travailleurs et les travailleurs concernés qu'ils seront soumis à l'utilisation du système d'IA à haut risque »[384].

Les fournisseurs de systèmes d'IA relatifs à l'emploi, à la gestion de main d'œuvre et à l'accès à l'emploi indépendant (annexe III (4)) ont l'obligation d'enregistrer leurs systèmes dans la base de données de l'UE. Toutefois, cette obligation ne concerne pas les déployeurs qui sont des entités privées. En outre, les déployeurs n'ont pas l'obligation de procéder à une analyse d'impact sur les droits fondamentaux. Cette absence d'obligation risque de créer un manque de visibilité sur les systèmes utilisés et leurs impacts, y compris dans les grandes entreprises.

7.4.Thématique centrale : Sécurité sociale et services d'aide à l'emploi

7.4.1. Contexte

De nombreux organismes publics utilisent des systèmes d'IA et de PDA dans les domaines de la sécurité sociale et des services d'aide à l'emploi pour déterminer l'éligibilité des personnes, calculer les prestations, cibler les contrôles dans le cadre de la détection des fraudes et des erreurs, et distribuer les ressources aux bénéficiaires. Des systèmes d'IA sont proposés aux agents de la sécurité sociale et aux professionnels de l'aide à l'emploi, et les chatbots basés sur l'IA sont de plus en plus utilisés pour interagir avec les bénéficiaires. En France, dans le cadre de son programme Intelligence Emploi, l'agence pour l'emploi France Travail a mis au point plusieurs systèmes d'IA, dont MatchFT, un chatbot conçu pour envoyer des annonces de postes à des candidats potentiels et pour jauger leur degré d'intérêt et leur éligibilité, mais aussi ChatFT, un chatbot utilisé par les agents pour récupérer des informations dans

384. Règlement sur l'IA, art. 26(7).

les bases de données de l'agence (info.gouv.fr, 2025). Le projet IA-NAVIGATE, collaboration entre FARI (institut de recherche universitaire en intelligence artificielle) et l'agence publique pour l'emploi de Bruxelles, Actiris, s'est intéressé au recours à des outils d'IA par les professionnels de l'accompagnement des demandeurs d'emploi à Bruxelles, et a constaté qu'environ 60 % d'entre eux avaient déjà utilisé des outils d'IA dans le cadre de leur travail, principalement pour rédiger des rapports, concevoir des ateliers, éditer des CV et préparer des entretiens d'embauche. Ce projet a également révélé que plus de 80 % d'entre eux n'avaient reçu aucune directive de la part des entités pour lesquelles ils travaillaient concernant le recours à des outils d'IA (Xenidis, 2025).

Certains de ces systèmes ont déjà été jugés discriminatoires. Par exemple, aux Pays-Bas, un système utilisé pour anticiper les cas de fraude s'est avéré discriminatoire à l'égard des bénéficiaires en raison de leur race, de leur origine ethnique et de leur nationalité[385]. L'agence autrichienne pour l'emploi a développé l'algorithme AMS, qui permet de prédire les chances d'emploi, afin d'allouer des ressources d'aide aux demandeurs d'emploi. Le prototype s'est révélé discriminatoire à l'égard des femmes (en particulier les mères célibataires) et des demandeurs d'emploi issus de l'immigration[386]. En Pologne, un système utilisé par l'agence pour l'emploi a finalement été abandonné après avoir été jugé inconstitutionnel[387]. Des systèmes de PDA et d'IA sont ou ont également été utilisés pour contrôler les bénéficiaires de l'aide sociale dans de nombreux pays (dont la France, les Pays-Bas, le Danemark et la Belgique), avec des résultats discriminatoires[388].

Des OSC ont saisi le Conseil d'État français d'un recours contre un système de notation des risques utilisé par l'organisme de sécurité sociale français pour prévoir les risques de fraude et d'erreurs et cibler les contrôles[389].

---

385. De Rechtspraak (13 février 2019) « SyRI legislation in breach of European Convention on Human Rights », De Rechtspraak, disponible en anglais sur www.rechtspraak.nl/Organisatie-en-contact/Organisatie/Rechtbanken/Rechtbank-Den-Haag/Nieuws/Paginas/SyRI-legislation-in-breach-of-European-Convention-on-Human-Rights.aspx, consulté le 11 novembre 2025. 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.
386. Allhutter D et al. (2020) « Algorithmic profiling of job seekers in Austria: How austerity politics are made effective », Frontiers in Big Data, 3, disponible à l'adresse https://doi.org/10.3389/fdata.2020.00005, consulté le 11 novembre 2025.
387. Szymielewicz K. et al. (2015), « Profiling the unemployed in Poland: social and political implications of algorithmic decision making », Fundacja Panoptykon, 2015, disponible en anglais sur https://en.panoptykon.org/profiling-unemployed-poland-report, consulté le 11 novembre 2025.
388. Romain et al. (2023); Mehrotra D. et al. (2023), « Inside the suspicion machine », WIRED, disponible en anglais sur www.wired.com/story/welfare-state-algorithms/, consulté le 11 novembre 2025; Geiger G. (2023), « How Denmark's welfare state became a surveillance nightmare », WIRED, 7 mars, disponible en anglais sur www.wired.com/story/algorithms-welfare-state-politics/, consulté le 11 novembre 2025; Amnesty International (2024), « Denmark: AI-powered welfare system fuels mass surveillance and risks discriminating against marginalized groups – report », disponible en anglais sur www.amnesty.org/en/latest/news/2024/11/denmark-ai-powered-welfare-system-fuels-mass-surveillance-and-risks-discriminating-against-marginalized-groups-report/, consulté le 11 novembre 2025; Degrave E. (2020), « The use of secret algorithms to combat social fraud in Belgium », European Review of Digital Administration & Law 1-2: 167-78.
389. Amnesty International (2024), France : CNAF State Council Complaint, disponible en anglais sur www.amnesty.org/fr/documents/eur21/8795/2024/en/, consulté le 11 novembre 2025.

Les OSC ont mis en lumière, entre autres, des discriminations fondées sur le sexe, la situation familiale, l'âge et le handicap. Les requérant.es ont demandé au Conseil d'État de soumettre les questions à la CJUE, notamment s'agissant des discriminations algorithmiques indirectes[390]. La question de la discrimination indirecte et de l'égalité substantielle est particulièrement présente dans les systèmes de sécurité sociale[391].

Un autre point clé consistera à évaluer quels systèmes relèvent de la définition d'un système d'IA dans le règlement sur l'IA, étant donné que de nombreux systèmes de sécurité sociale restent relativement simples sur le plan technique, bien qu'ils présentent un risque important pour les droits fondamentaux.

7.4.2. Systèmes d'IA interdits

- L'article 5, paragraphe 1, point c), sur la notation sociale est une disposition cruciale à prendre en considération dans le domaine de la sécurité sociale et des services d'aide à l'emploi, où de nombreux systèmes de contrôle ou d'orientation des bénéficiaires sont déployés (voir ci-dessus).
- Article 5 (1) (f) concernant la reconnaissance des émotions sur le lieu de travail, à laquelle on pourrait avoir recours lors du processus de recrutement.

7.4.3. Systèmes d'IA à haut risque

L'annexe III (5), relative à l'accès et au droit aux services privés essentiels et aux services publics et prestations sociales essentiels, couvre les « systèmes d'IA destinés à être utilisés par les autorités publiques ou en leur nom pour évaluer l'éligibilité des personnes physiques aux prestations et services d'aide sociale essentiels, y compris les services de soins de santé, ainsi que pour octroyer, réduire, révoquer ou récupérer ces prestations et services »[392].

Les organismes publics s'appuient sur des systèmes de calcul pour calculer les prestations des bénéficiaires, souvent à l'aide de systèmes basés sur des règles. Ici, la question est de savoir quels sont les systèmes de calcul qui relèvent de la définition des systèmes d'IA dans le règlement sur l'IA.

Conformément à l'article 6 du règlement sur l'IA, les fournisseurs peuvent s'exclure eux-mêmes du régime à haut risque selon certaines conditions. Certaines de ces conditions sont susceptibles d'être appliquées pour les systèmes utilisés dans le

---

390. La question est la suivante : « Le traitement de données à caractère personnel, dont le responsable du traitement est une administration publique sociale chargée d'une mission de service public, qui a pour finalité d'établir un score de risque pour chaque usager du service public afin de cibler les contrôles de l'administration, ne constitue-t-il pas une discrimination indirecte au sens de la directive 2000/43/CE du 29 juin 2000 et de la directive 79/7/CEE du Conseil du 19 décembre 1978 ? lue à la lumière des articles 20 et 21 de la Charte, dans la mesure où la mise en place de ce traitement a entraîné une augmentation significative des contrôles auprès des personnes de moins de 30 ans, des étudiants, des personnes à faibles revenus, des inactifs, des familles monoparentales (groupe composé à 95 % de femmes), ou des personnes bénéficiant d'aides sociales ? Amnesty International (2024), France : CNAF State Council Complaint, paragraphe 71, disponible en anglais sur www.amnesty.org/fr/documents/eur21/8795/2024/en/, consulté le 11 novembre 2025.
391. Wachter, S., Mittelstadt, B., Russell, C. (2020), « Bias preservation in machine learning : the legality of fairness metrics under EU non-discrimination law, West Virginia Law Review, 123(123, 735.
392. Règlement sur l'IA, annexe III (5) (a).

cadre de l'aide sociale et de la sécurité sociale. Par exemple, les utilisations réelles comprennent le traitement automatique du langage naturel pour analyser le contenu des courriels des demandeurs d'emploi[393], ce qui peut être fait « en préparation d'une évaluation ». On peut également citer l'utilisation de la reconnaissance optique de caractères (OCR) pour aider à recadrer, faire pivoter et nettoyer les documents de demande[394], ce qui pourrait être considéré comme une « tâche procédurale étroite ». Toutefois, comme indiqué ailleurs dans les présentes lignes directrices, ces définitions restent vagues et, en l'absence de jurisprudence, une surveillance étroite est nécessaire.

7.4.4. Obligations d'enregistrement

Dans le domaine de la sécurité sociale, les systèmes à haut risque seront enregistrés dans la version publique de la base de données par les fournisseurs et les déployeurs. Les informations enregistrées seront notamment les synthèses des AIDF menées par les déployeurs.

393. Agence des droits fondamentaux de l'Union européenne (2020), Bien préparer l'avenir : l'intelligence artificielle et les droits fondamentaux, Luxembourg : Office des publications de l'Union européenne, p. 32.
394. Ibid. p. 33.

Références

Allen R. et Masters D. (2020), Regulating for an equal AI: A new role for equality bodies: Meeting the new challenges to equality and non-discrimination from increased digitisation and the use of Artificial Intelligence, Equinet.

BBC (2020), « Home Office drops 'racist' algorithm from visa decisions », BBC, disponible en anglais sur : www.bbc.com/news/technology-53650758, consulté le 7 novembre 2025.

Direction générale des Entreprises (2025), Les autorités compétentes pour la mise en œuvre du règlement européen sur l'intelligence artificielle, Ministère de l'Économie, des Finances et de la Souveraineté industrielle, disponible sur www.entreprises.gouv.fr/priorites-et-actions/transition-numerique/soutenir-le-developpement-de-lia-au-service-de-0, consulté le 7 novembre 2025.

Dumbrava C. (2025), « Briefing: Artificial intelligence in asylum procedures in the EU », Service de recherche du Parlement européen, disponible en anglais sur www.europarl.europa.eu/RegData/etudes/BRIE/2025/775861/EPRS_BRI(2025)775861_EN.pdf, consulté le 7 novembre 2025.

Equinet (2024), Understanding the new EU Directives on standards for equality bodies: legal digest on standards for equality bodies, disponible en anglais sur https://equineteurope.org/publications/understanding-the-new-eu-directives-on-standards-for-equality-bodies-legal-digest-on-standards-for-equality-bodies/, consulté le 11 novembre 2025.

Equinet (2025), How to use the Artificial Intelligence Act to investigate AI bias and discrimination: A guide for equality bodies, European Network of Equality Bodies, disponible sur https://equineteurope.org/publications/how-to-use-the-artificial-intelligence-act-to-investigate-ai-bias-and-discrimination-a-guide-for-equality-bodies/, consulté le 10 novembre 2025.

Agence des droits fondamentaux de l'Union européenne (2021), Equality in the EU: 20 years on from the initial implementation of the Equality Directives, disponible en anglais sur : https://fra.europa.eu/sites/default/files/fra_uploads/fra-2021-opinion-equality-directives-01-2021_en_0.pdf, consulté le 7 novembre 2025.

Hellenic Data Protection Authority (2024), Ministry of Migration and Asylum receives administrative fine and GDPR compliance order following an own-initiative investigation by the Hellenic Data Protection Authority, disponible en anglais sur : www.dpa.gr/en/enimerwtiko/press-releases/ministry-migration-and-asylum-receives-administrative-fine-and-gdpr, consulté le 7 novembre 2025.

Ilieva M. (2024), Handbook on identifying and using equality data in legal casework, Equinet, disponible en anglais sur : https://equineteurope.org/publications/handbook-on-identifying-and-using-equality-data-in-legal-casework/, consulté le 7 novembre 2025.

info.gouv.fr (2025), France Travail : comment ça marche?, disponible sur : www.info.gouv.fr/actualite/france-travail-comment-ca-marche, consulté le 7 novembre 2025.

Jones C., Lanneau R., Maccanico Y. (2023), Europe's techno borders, EuroMed Rights and Statewatch, disponible en anglais sur https://www.statewatch.org/publications/reports-and-books/europe-s-techno-borders/, consulté le 7 novembre 2025.

McGregor L. (2023), Digital border governance: a human rights based approach, University of Essex et HCDH, disponible sur www.ohchr.org/en/documents/tools-and-resources/digital-border-governance-human-rights-based-approach, consulté le 7 novembre 2025.

Ozkul D. (2023), Automating immigration and asylum: the uses of new technologies in migration and asylum governance in Europe, Oxford: Refugee Studies Centre, University of Oxford, disponible en anglais sur : www.rsc.ox.ac.uk/publications/automating-immigration-and-asylum-the-uses-of-new-technologies-in-migration-and-asylum-governance-in-europe, consulté le 7 novembre 2025.

Romain M. et al. (2023), « Is data neutral? How an algorithm decides which French households to audit for welfare fraud », Le Monde, disponible en anglais sur www.lemonde.fr/en/les-decodeurs/visuel/2023/12/05/how-an-algorithm-decides-which-french-households-to-audit-for-benefit-fraud_6313254_8.html, consulté le 11 novembre 2025.

Subgroup on Equality Data du High Level Group on Non-discrimination, Equality and Diversity de la Commission européenne (2025), Collecting and using equality data in full compliance with EU General Data Protection Regulation and national data protection rules, Publications Office of the European Union, disponible en anglais sur : https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combatting-discrimination/equality-data-collection_en, consulté le 7 novembre 2025.

Veld S. (in 't) et al. (2020), Letter to Andrea Jelinek, Chair of the European Data Protection Board on the possible use of the Clearview AI application by law enforcement authorities in the EU, disponible en anglais sur www.edpb.europa.eu/sites/default/files/files/file1/edpb_letter_out_2020-0052_facialrecognition.pdf, consulté le 7 novembre 2025.

Xenidis R. (2020), Tuning EU equality law to algorithmic discrimination: Three pathways to resilience, Maastricht Journal of European and Comparative Law, disponible en anglais sur : https://journals.sagepub.com/doi/10.1177/1023263X20982173, consulté le 7 novembre 2025.

Xenidis R. (2025), La protection juridique contre la discrimination algorithmique en Europe : cadres actuels et lacunes subsistantes, Conseil de l'Europe, Strasbourg

Les administrations publiques à travers l'Europe utilisent l'intelligence artificielle (IA) et/ou les systèmes de prise de décision automatisée (PDA) dans un large éventail de domaines politiques, notamment la migration, la protection sociale, la justice, l'éducation, l'emploi, la fiscalité, l'application de la loi ou les soins de santé. Ces systèmes sont également déployés dans des domaines sensibles du secteur privé, tels que la banque et l'assurance. Bien que les systèmes d'IA et de PDA présentent des risques importants de discrimination, des défis subsistent pour identifier et atténuer ces risques. Les organismes de promotion de l'égalité et autres structures nationales de défense des droits humains ont donc un rôle clé à jouer dans la promotion d'un déploiement, par les organisations du secteur public, de systèmes d'IA/PDA conforme aux droits fondamentaux. Les lignes directrices visent à donner aux organismes de promotion de l'égalité et autres structures nationales de défense des droits humains, en particulier dans l'Union européenne, les moyens de lutter contre la discrimination dans les systèmes d'IA/PDA. Elles les informent de leurs responsabilités dans un environnement réglementaire en mutation, notamment en ce qui concerne le règlement européen sur l'IA 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. Elles proposent des recommandations et des exemples pour l'application des nouvelles réglementations et servent de ressource pour aider et conseiller les acteurs·rices nationaux·ales, tels que les décideurs·euses politiques et les régulateurs·rices, en matière de droits humains, d'égalité et de non-discrimination.

YHDENVERTAISUUSVALTUUTETTU
NON-DISCRIMINATION OMBUDSMAN

UNIA

CIG
COMISSÃO PARA A CIDADANIA E A IGUALDADE DE GÉNERO

PREMS 163925

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