MÉTHODOLOGIE EUROPÉENNE D'ÉVALUATION DE LA DISCRIMINATION ALGORITHMIQUE

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

Louise Hooper
Ylja Remmits
Ioanna Papageorgiou

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

Cette publication a été réalisée avec le soutien financier de l'Union européenne et du Conseil de l'Europe. Son contenu relève de la responsabilité exclusive de ses autrices. Les opinions exprimées ici ne peuvent en aucun cas être interprétées comme reflétant le point de vue officiel de l'Union européenne ou du Conseil de l'Europe.

Toute autre correspondance concernant ce document doit être adressée au Coordinateur IA de la Division de l'inclusion et de la lutte contre la discrimination du Conseil de l'Europe, F-67075 Strasbourg Cedex, E-mail : anti-discrimination@coe.int

Couverture et mise en page : Division publications et identité visuelle, Conseil de l'Europe
Mise en page de la version anglaise : Luminess
Mise en page de la version française : ELASTIK LAB SRL
Traduction : LingvoHouse Translation Services Ltd
Révision juridique de la traduction : Diana Nunes
Photos : Shutterstock, Unsplash and Better Images of AI

Table des matières

Abréviations

AIDF	Analyse d'impact sur les droits fondamentaux
AIDH	Analyse d'impact sur les droits humains
AIPD	Analyse d'impact sur la protection des données
APD	Autorité de protection des données
ASM	Autorité de surveillance du marché
CJUE	Cour de justice de l'Union européenne
DAPC	Demande d'accès de la personne concernée (Data subject access request)
DPJ	Directive « Police-Justice »
EDPB	Comité européen de la protection des données (European Data Protection Board)
IA	Intelligence artificielle
IBD	Identification biométrique à distance
MARD	Mode alternatif de règlement des différends
OPE	Organisme de promotion de l'égalité
OSC	Organisation de la société civile
PDA	Prise de décision automatisée
RGPD	Règlement général sur la protection des données

Terminologie

Aux fins de la présente méthodologie, les choix terminologiques pertinents suivants ont été faits ou les termes ont la signification indiquée ci-dessous.

Autorité de surveillance du marché : l'autorité ou les autorités nationale(s) exerçant des activités et prenant des mesures en vertu du règlement (UE) 2019/1020 (surveillance du marché et conformité des produits).

Directives sur la non-discrimination : le terme fait référence aux directives européennes énumérées à l'Annexe G.

Données d'entrée : toute information, toute donnée brute ou tout signal fourni au système algorithmique, sur le fondement duquel le système parvient à son résultat. Par exemple, une application de dépistage du cancer fondée sur l'IA, destinée à détecter les grains de beauté suspects, peut utiliser l'âge, le genre et une photo du grain de beauté comme données d'entrée. Dans un système LLM (grand modèle de langage) comme ChatGPT, l'entrée est le texte de l'interaction de chat et potentiellement le texte des interactions de chat précédentes également.

Données de sortie : résultat du système algorithmique après traitement des données d'entrée. Par exemple, une caractéristique prédite concernant un individu, un score, une recommandation, une décision ou un contenu comme des images, des vidéos ou du texte.

Expert·e : une personne possédant des connaissances spécialisées ou une expertise technique, consultée ou impliquée par l'organisme de promotion de l'égalité en vue de soutenir la personne en charge du dossier à n'importe quel stade de cette méthodologie. L'expert·e apporte une contribution spécialisée mais n'est pas la principale personne en charge du dossier.

Fournisseur/déployeur : les termes sont utilisés comme défini à l'article 3 du règlement sur l'IA. Bien que définis en référence aux systèmes d'IA, les rôles respectifs peuvent également s'appliquer, par analogie, aux systèmes non IA.

Matrice de confusion : tableau utilisé pour évaluer les performances d'un système algorithmique en comparant ses prévisions aux résultats réels. Les lignes du tableau de la matrice de confusion indiquent les résultats réels, tandis que les colonnes indiquent les prévisions du système. Le tableau montre combien de prédictions étaient correctes ou incorrectes.

Non-discrimination : le terme « non-discrimination » est préféré à « égalité » ou à des termes connexes.

Organisation intimée : l'organisation ou l'entité qui fait l'objet d'une enquête menée par l'organisme de promotion de l'égalité pour discrimination interdite suite à une plainte, une procédure ex officio ou une notification de l'autorité de surveillance du marché (ASM). Dans le contexte du Règlement général sur la protection des donnée (RGPD) et du règlement sur l'IA, les organisations intimées peuvent être appelé·es

respectivement « contrôleurs de données » et « fournisseurs d'IA » ou « déployeurs », tels que ces termes sont définis dans ces règlements.

Organisme notifié : organisme d'évaluation de la conformité qui effectue des évaluations de conformité de tierces parties, y compris des tests, des certifications et des inspections. Par exemple, l'organisme notifié évalue la conformité des systèmes d'IA à haut risque conformément aux procédures d'évaluation de la conformité énoncées dans le règlement européen sur l'IA.

Personne en charge de dossier : membre d'un organisme de promotion de l'égalité qui procède à l'évaluation à n'importe quelle étape de cette méthodologie dans le cadre de son travail d'investigation de cas de discrimination[1].

Plaignant·e(s) : personne ou groupe de personnes alléguant ou subissant une discrimination (illégale). Par souci de cohérence avec la formulation du RGPD et du règlement sur l'IA, les termes « personne(s) concernée(s) » et « personne(s) affectée(s) » utilisés respectivement dans ces deux textes, correspondent dans le présent document à la notion de "plaignant·e(s)".

Système algorithmique : terme utilisé dans le cadre de cette méthodologie pour couvrir à la fois les systèmes d'IA et les systèmes non-IA, y compris les systèmes de prise de décision automatisée (ADM). Il englobe également les systèmes d'IA qui ne sont pas utilisés dans les processus décisionnels.

Système de prise de décision automatisée (Automated decision-making system ou ADM) : ce terme couvre un large éventail de systèmes automatisés, qui sont utilisés dans le contexte ou aux fins d'une prise de décision. Ces systèmes peuvent être fondés sur l'IA mais peuvent également reposer sur d'autres technologies, même aussi simples que des règles de base de type « si-alors ». Ils vont de systèmes entièrement automatisés, qui ne nécessitent aucune contribution humaine, à des systèmes semi-automatisés ou d'aide à la décision, qui combinent des traitements effectués par la machine et des interventions humaines, et qui sont très largement utilisés dans la pratique.

Système de prise de décision automatisée (ADM) admissible : aux fins de la présente méthodologie, le terme « ADM admissible » est utilisé pour désigner une sous-catégorie de systèmes ADM répondant aux critères définis dans l'article 22 du RGPD, à savoir une décision automatisée « uniquement » (règlement (UE) 2016/679). Ces systèmes déclenchent des obligations accrues de transparence et d'information en vertu du RGPD.

Système d'IA : « un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions,

1. Bien que cette méthodologie soit principalement destinée aux OPE, elle peut également être utile aux institutions nationales des droits humains (INDH) et aux institutions de médiation ou autres parties prenantes. Lorsque cette méthodologie est appliquée par des INDH et des institutions de médiation, le terme « personne en charge du dossier » pourrait également désigner les personnes en charge des dossiers au sein de ces institutions.

du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels » (règlement sur l'IA, article 3(1)).

Systèmes d'IA à haut risque ou « IA à haut risque » : désigne un système d'IA tel que défini de manière exhaustive à l'article 6 du règlement sur l'IA.

Systèmes d'IA hors haut risque : un système d'IA qui n'est pas inclus dans la définition exhaustive de l'IA à haut risque en vertu du règlement sur l'IA. Veuillez noter que cette classification signifie uniquement que ces systèmes ne sont pas classés comme présentant un risque élevé selon le règlement sur l'IA, et non pas que ces systèmes d'IA sont sans risque.

Pour la terminologie relative au règlement sur l'IA, veuillez-vous référer à l'article 3 du règlement sur l'IA.

Remerciements

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

Le rapport a été rédigé par Louise Hooper (consultante internationale, Conseil de l'Europe), Ylja Remmits et Ioanna Papageorgiou (audit en matière d'algorithmes/consultantes internationales, 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, Sophie Vincent, Thelma Raes et Laetitia Parmentier (Unia, Belgique) ; Tiina Valonen et Ville Rantala (YVV, Finlande) ; Carla Peixe, Lúcia Pestana et Alexandra Andrade (CIG, Portugal) ; Massimiliano Santini (Task Force Réformes et Investissements, Secrétariat Général, Commission européenne) ; et Menno Ettema, Sara Haapalainen, Ayca Dibekoğlu et Delfine Gaillard (Service de l'Anti-discrimination et de l'inclusion, Conseil de l'Europe).

Remerciements particuliers à Raphaële Xenidis, Kris Shrishak, Brent Mittelstadt, Chris Russell, Ydo Bergstra et Jurriaan Parie pour leurs contributions en matière de révision, ainsi qu'à Laurens Naudts et Plixavra Vogiatzoglou pour leur expertise. Le rapport a également bénéficié des discussions et des contributions d'expert·es d'organismes nationaux de promotion de l'égalité et d'organisations de la société civile qui ont participé à un séminaire et atelier ayant réuni plusieurs pays à Bruxelles le 15 avril 2026. Toutes ces perspectives ont enrichi les sections analytiques et opérationnelles de ces lignes directrices.

UNIA

YHDENVERTAISUUSVALTUUTETTU
NON-DISCRIMINATION OMBUDSMAN

CIG
COMISSÃO PARA A CIDADANIA
E A IGUALDADE DE GÉNERO

Synthèse

Les systèmes algorithmiques peuvent renforcer ou provoquer des discriminations dans la mesure où ils sont souvent conçus pour opérer des distinctions entre les individus, notamment en les classant, en les hiérarchisant ou en les ciblant. Combinés aux inégalités historiques et aux biais humains intégrés dans les données, les règles et les choix de conception des systèmes algorithmiques, ils sont susceptibles de reproduire et d'amplifier des schémas discriminatoires, à moins qu'ils ne soient identifiés et atténués de manière proactive.

Cette méthodologie propose un guide étape par étape aux organismes de promotion de l'égalité (OPE) et aux autres organisations de défense des droits humains en Europe pour évaluer les cas de discrimination liées à l'intelligence artificielle (IA) ou aux systèmes algorithmiques. Elle est destinée aux personnes en charge de dossier, aux conseils et aux avocat·es de ces entités qui enquêtent sur des plaintes pour discrimination impliquant ou soupçonnant l'implication de l'IA ou la prise de décision automatisée (ADM). Elle vise à servir de liste de référence des problèmatiques et des sources à examiner à chaque étape d'une évaluation. La méthodologie est divisée en quatre étapes itératives :

Étape 1 : l'identification des cas liés à l'IA ou aux algorithmes

Étape 2 : la collecte d'informations sur l'IA ou le système algorithmique pour identifier la discrimination

Étape 3 : la méthodologie d'évaluation des informations reçues

Étape 4 : les solutions

Ces étapes s'appuient sur le cadre réglementaire européen concernant l'IA, en particulier sur le règlement européen sur l'IA et sur la législation relative à la non-discrimination, et font référence à d'autres législations, telles que celle relative à la protection des données, des exemples d'affaires et la jurisprudence, le cas échéant. Les différentes étapes établissent une distinction entre les systèmes d'IA à haut risque et les autres systèmes, conformément au règlement européen sur l'IA, qui prévoit des mécanismes permettant d'accéder à la documentation technique, aux données et aux tests des systèmes d'IA à haut risque, notamment pour les organismes désignés en tant que des « autorités chargées de la protection des droits fondamentaux » en vertu de l'article 77 du règlement européen sur l'IA.

Cette méthodologie ne se concentre pas sur la législation nationale et, dans certains pays, le droit national peut offrir des moyens plus efficaces d'accéder à des informations ou des pouvoirs d'enquête sur la discrimination algorithmique. Les personnes chargées du traitement des dossiers doivent garder ces considérations à l'esprit lorsqu'elles utilisent cet outil.

Pour soutenir le suivi des quatre étapes, la méthodologie comporte une longue liste d'annexes qui présentent des modèles de lettres types pour demander des informations, qui aident à la compréhension des risques de discrimination algorithmique, par type de système d'IA ou par secteur où un système est utilisé, entre

autres, et qui contribuent à l'évaluation d'informations, par exemple en proposant une check-list/un document de suivi pour l'étape 3.

Pour faciliter l'utilisation de cette méthodologie, une compréhension de base de l'IA, des biais algorithmiques et de la discrimination, ainsi que de la terminologie associée est utile, tandis que l'outil fournit également un glossaire spécifique de la terminologie. D'autres publications produites 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 » complètent cette méthodologie et peuvent contribuer à renforcer cette compréhension de base.

- La protection juridique contre la discrimination algorithmique en Europe : cadres actuels et lacunes subsistantes, décembre 2025 (également disponible en anglais, finlandais, néerlandais, portugais, suédois).
- Les 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, décembre 2025 (également disponible en anglais, finlandais, néerlandais, portugais, suédois).

Introduction

De nouveaux cadres réglementaires concernant l'IA, comme le règlement européen sur l'IA[2], les directives européennes sur les normes applicables aux organismes de promotion de l'égalité (ci-après « les directives relatives aux normes ») 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 (ci-après[3] la « Convention-cadre du Conseil de l'Europe sur l'IA »), fournissent de nouveaux outils et pouvoirs aux organismes de promotion de l'égalité et de défense des droits humains pour enquêter sur la discrimination algorithmique.

Ces outils présentent certaines limites et défis. Par exemple, certains systèmes algorithmiques discriminatoires peuvent ne pas être catégorisés en tant qu'IA en vertu du règlement sur l'IA[4]. En outre, de nombreux domaines du droit relatifs à l'IA et à la discrimination ne sont pas encore parfaitement clairs. Enfin, pour de nombreux lecteurs, lectrices et personnes en charge de dossiers, il s'agira d'un domaine du droit nouveau et relativement complexe.

Malgré ces défis, cette méthodologie en quatre étapes vise à proposer un cadre structuré permettant d'analyser systématiquement l'impact discriminatoire d'une décision ou d'un traitement susceptible d'avoir été pris ou réalisé au moyen d'un système algorithmique. Les lecteurs et lectrices ne devraient pas considérer chaque étape de la méthodologie comme devant être accomplie. Les étapes fournissent plutôt un guide de référence contenant un large éventail d'outils permettant d'identifier les systèmes, de recueillir et d'évaluer les preuves pour décider des actions à entreprendre. En suivant les étapes de la méthodologie et en sollicitant l'assistance d'expert·es (techniques) si nécessaire, il devrait être possible de disposer d'éléments permettant d'établir la présomption de discrimination et de rechercher des voies de recours dans les cas de discrimination algorithmique.

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 (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 et 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.

4 Certains exemples de discrimination algorithmique sont fondés sur des systèmes simples définis par des règles qui ne sont pas qualifiés d'IA, comme les exemples néerlandais d'algorithmes discriminatoires dans le scandale des allocations familiales (Amnesty International 2021) et l'évaluation du risque de fraude par l'Agence exécutive néerlandaise pour l'éducation (Tribunal d'Overijssel 2024).

Les OPE devraient aborder les allégations de discrimination en relation avec l'IA et l'ADM comme tout autre cas de discrimination. Ils doivent toutefois également reconnaître que le manque de transparence ou l'incertitude quant au fonctionnement d'un système algorithmique peut être pris en compte pour apprécier l'existence d'un risque de discrimination ou pour déterminer si une présomption de discrimination (prima facie) est établie. Cette approche repose sur la reconnaissance croissante de l'existence d'un déséquilibre de pouvoir et d'information entre les entreprises qui développent des systèmes d'IA, les entités chargées de leur déploiement et les personnes à l'égard desquelles ces outils sont utilisés[5]. Dans ces circonstances, la charge de la preuve peut être renversée à l'organisation intimée, qui doit établir que le système algorithmique n'a pas été à l'origine d'une discrimination ou n'y a pas contribué de manière illicite. Pour que les personnes concernées aient effectivement un accès effectif à la justice, un degré plus élevé de transparence est indispensable.

L'omnibus numérique sur l'IA[6], un paquet législatif modifiant et rationalisant le règlement sur l'IA ainsi que d'autres textes législatifs, a été adopté par le Parlement européen. Lorsqu'elles sont pertinentes, les références aux modifications introduites par ce paquet législatif sont mentionnées dans le présent document.

Éléments clés du nouveau cadre juridique

Les directives européennes sur les normes applicables aux organismes de promotion de l'égalité

Les directives dites relatives aux normes, qui devaient être transposées d'ici le 19 juin 2026 (article 24 des directives relatives aux normes), fixent des exigences minimales pour le fonctionnement des organismes de promotion de l'égalité afin d'améliorer leur efficacité et de garantir leur indépendance. Ces compétences et pouvoirs doivent être lus parallèlement aux pouvoirs existants dans les directives européennes relatives à la non-discrimination.

La directive 2024/1500, considérant 21, et la directive 2024/1499, considérant 22, indiquent, en ce qui concerne l'allocation de ressources financières, que :

> 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

5 Ce point fait l'objet d'une discussion détaillée dans la sous-section 3.8.

6 Règlement (UE) 2026/1744 du Parlement européen et du Conseil du 8 juillet 2026 modifiant les règlements (UE) 2024/1689, (UE) 2018/1139 et (UE) 2023/1230 en ce qui concerne la simplification de la mise en œuvre des règles harmonisées concernant l'intelligence artificielle (train de mesures omnibus numérique sur l'IA), disponible : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=OJ:L_202601744 (consulté le 24 juillet 2026).

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[7].

Le règlement européen sur l'IA, article 77 : pouvoirs des autorités chargées de la protection des droits fondamentaux et coopération avec les autorités de surveillance du marché

De nouvelles compétences sont conférées aux organismes de promotion de l'égalité et aux institutions nationales des droits humains (INDH) qui répondent aux exigences de l'article 77 du règlement sur l'IA[8] en ce qui concerne les systèmes d'IA à haut risque afin de leur permettre d'exercer effectivement leur mandat[9]. L'Omnibus numérique sur l'IA vise à préciser que ces pouvoirs s'exercent sans porter préjudice aux pouvoirs existants et à l'indépendance de ces organismes[10].

La liste des autorités chargées de la protection des droits fondamentaux aurait dû être notifiée à la Commission européenne au 2 novembre 2024 et les États membres doivent tenir « la liste [des autorités] à jour »[11]. La Commission reconnaît que la liste des organismes visés à « l'article 77 » est continuellement mise à jour à la lumière de la mise en œuvre actuelle du règlement sur l'IA[12]. Les OPE ne figurant pas sur la liste établie par les États membres à la date limite de la première notification peuvent toujours y être ajoutés. Des OPE pourraient également être retirés de la liste.

En vertu de l'article 77(1), les autorités compétentes sont habilitées à demander et à consulter tout document créé ou conservé dans le cadre du règlement sur l'IA dans une langue et un format accessibles, si nécessaire afin qu'elles s'acquittent efficacement de leurs mandats dans les limites de leur compétence. Par conséquent, des demandes devraient pouvoir être introduites concernant la documentation créée ou conservée par d'autres autorités nationales compétentes, y compris les autorités

7. L'article 26 de la Convention-cadre sur l'IA exige également que les OPE 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..
8 Article 77(1) du règlement sur l'IA, incluant les amendements de l'Omnibus : « Les autorités ou organismes publics nationaux qui surveillent ou contrôlent le respect des obligations découlant du droit de l'Union protégeant les droits fondamentaux, y compris le droit à la non-discrimination, sont habilités à présenter une demande et à accéder à toute information ou documentation créée ou conservée par l'autorité de surveillance du marché concernée en vertu du présent règlement, dans une langue et un format accessibles, lorsque l'accès à ces informations ou à cette documentation est nécessaire à l'accomplissement effectif de leur mandat dans les limites de leurs compétences. Le présent Article s'entend sans préjudice des compétences, des tâches, des pouvoirs et de l'indépendance des autorités ou organismes publics nationaux compétents, dans le cadre de leur mandat. »
9 Tous les organismes de promotion de l'égalité ne relèvent pas de la définition donnée à l'article 77, bien que ce soit le cas de nombre d'entre eux. Le terme « autorité relevant de l'article 77 » est utilisé partout pour désigner tout OPE ou INDH répondant à la définition donnée à l'article 77 du règlement sur l'IA.
10 Règlement (UE) 2026/1744 du Parlement européen et du Conseil du 8 juillet 2026 modifiant les règlements (UE) 2024/1689, (UE) 2018/1139 et (UE) 2023/1230 en ce qui concerne la simplification de la mise en œuvre des règles harmonisées concernant l'intelligence artificielle (train de mesures omnibus numérique sur l'IA), disponible : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=OJ:L_202601744 (consulté le 24 juillet 2026).
11 Règlement sur l'IA, article 77(2).
12 Autorités de protection des droits fondamentaux dotées de pouvoirs spéciaux en vertu de la législation sur l'IA | Bâtir l'avenir numérique de l'Europe, consulté le 5 juin 2026.

de surveillance du marché et les organismes notifiés, ainsi que les fournisseurs, les déployeurs ou d'autres acteurs du cycle de vie d'un système d'IA. L'Omnibus numérique précise que les demandes d'accès aux informations et à la documentation doivent être adressées aux autorités de surveillance du marché, qui « accorderont l'accès à ces informations et à cette documentation, y compris en demandant ces informations et cette documentation au fournisseur ou au déployeur, si nécessaire et sans retard injustifié »[13], et que les autorités ou organismes concernés « coopéreront étroitement et se prêteront l'assistance mutuelle nécessaire à l'accomplissement de leurs mandats respectifs »[14]. Cette coopération et cette assistance comprennent l'échange d'informations, si nécessaire, en vue de la supervision ou de la mise en œuvre effectives du règlement sur l'IA et d'autres législations de l'Union, y compris les directives en matière de non-discrimination (voir Annexe G).

Si la documentation est insuffisante pour déterminer si une infraction aux obligations en vertu du droit de l'Union protégeant les droits fondamentaux a eu lieu, l'organisme de promotion de l'égalité peut adresser une demande motivée à l'autorité de surveillance du marché pour organiser des tests du système d'IA à haut risque par le biais de moyens techniques (règlement sur l'IA, article 77(3) ; voir étape 4).

Toute information ou documentation obtenue doit être traitée conformément aux exigences de confidentialité de l'article 78. L'article 78(2), précise que l'autorité au titre de l'article 77 ne devrait demander que les données strictement nécessaires à l'évaluation du risque présenté par les systèmes d'IA et à l'exercice de ses pouvoirs conformément à la réglementation et au droit de l'Union. Par conséquent, toute demande présentée dans le cadre du règlement sur l'IA devrait être claire et précise pour éviter d'être refusée.

Lorsque des données sont demandées et obtenues, l'organisme de promotion de l'égalité veille à ce que des mesures de cybersécurité appropriées et efficaces soient mises en œuvre afin de garantir la sécurité et la confidentialité des informations et des données obtenues. Il doit supprimer ces données dès qu'elles ne sont plus nécessaires aux fins auxquelles elles ont été obtenues, conformément au droit national ou de l'Union applicable.

De même, les normes du Conseil de l'Europe, comme la Convention-cadre sur l'IA et la Recommandation CM/Rec(2026)1 sur l'égalité et l'intelligence artificielle, définissent les démarches que les États membres et les autres parties prenantes doivent entreprendre pour promouvoir l'égalité et prévenir et combattre la discrimination dans toutes les activités du cycle de vie des systèmes d'IA.

13 Règlement (UE) 2026/1744 du Parlement européen et du Conseil du 8 juillet 2026 modifiant les règlements (UE) 2024/1689, (UE) 2018/1139 et (UE) 2023/1230 en ce qui concerne la simplification de la mise en œuvre des règles harmonisées concernant l'intelligence artificielle (train de mesures omnibus numérique sur l'IA), disponible : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=OJ:L_202601744 (consulté le 24 juillet 2026).

14 Ibid.

Confidentialité

Les exigences de confidentialité sont énoncées à l'article 78(1) du règlement sur l'IA pour la Commission européenne, les autorités de surveillance du marché et les organismes notifiés (article 70) ainsi que toute autre personne physique ou morale impliquée dans la mise en œuvre du règlement sur l'IA, afin de respecter la confidentialité des informations et des données obtenues de manière à protéger :

- les droits de propriété intellectuelle et les informations commerciales confidentielles ou les secrets d'affaires d'une personne physique ou morale, y compris le code source (directive (UE) 2016/943, article 5). En ce qui concerne les organismes de promotion de l'égalité, l'article 5 de la directive 2016/943 prévoit une exception aux fins de révéler une faute, un acte répréhensible ou une activité illégale, à condition que l'organisation intimée (l'OPE ou potentiellement l'ASM) ait agi dans l'objectif de protéger l'intérêt public général (article 5(b)) ou un intérêt légitime reconnu par le droit de l'Union ou le droit national (article 5(d)). La protection des droits fondamentaux et la prévention de la discrimination ou la lutte contre celle-ci sont toutes deux couvertes, au moins probablement, par cette exception ;
- la mise en œuvre effective de la réglementation, notamment aux fins d'inspections, enquêtes ou audits ;
- les intérêts de la sécurité publique et nationale ;
- a conduite de procédures pénales ou administratives ;
- les informations classifiées conformément au droit de l'Union ou au droit national.

Actuellement, la manière dont ces éléments peuvent être utilisés pour impacter la capacité de l'OPE à obtenir des informations de la part des parties concernées n'est pas claire. Cependant, lorsque la divulgation est refusée au motif du respect de la confidentialité, l'OPE peut avoir recours soit à une autorité compétente en vertu du règlement sur l'IA, soit à l'Autorité de protection des données (APD) en vertu du Règlement général sur la protection des donnée (RGPD). Lorsque le secret des affaires est invoqué en tant que motif de non-divulgation, l'autorité de contrôle compétente ou le tribunal est probablement tenu·e d'équilibrer les droits et les intérêts en cause en vue de déterminer ce qui devrait être divulgué (voir CK vs Dun & Bradstreet Austria GmbH et Magistradt der Stadt Wien (2025), examiné à l'étape 2.1).

Coopération intersectorielle

Pour que le nouveau cadre juridique soit pleinement effectif, il sera également nécessaire d'instaurer, de développer, de renforcer et de formaliser la coopération avec les autorités compétentes de surveillance du marché, les autorités de protection des données et toute autre autorité relevant de l'article 77, lorsqu'aucun mécanisme de coopération n'existe encore entre ces organismes. Cette obligation de coopération est formalisée dans l'Omnibus numérique par un amendement intégrant l'article

77(1)(b)[15]. À cette fin, il peut être nécessaire ou utile d'élaborer des protocoles de travail inter-agences[16]. Des exemples et des bonnes pratiques sont présentés dans l'encadré ci-dessous. Les nouvelles directives relatives aux normes prévoient la mise en place de cadres nationaux pour mener des enquêtes permettant aux organismes de promotion de l'égalité de réaliser des constats, y compris des mécanismes appropriés de coopération entre les organismes de promotion de l'égalité et les autres autorités publiques compétentes à cette fin (article 8 des directives relatives aux normes).

Les négociations entre les OPE et le gouvernement de leur État membre en vue de garantir une base juridique claire pour le partage d'informations entre les organismes visés à l'article 77, y compris les organismes de promotion de l'égalité et les autorités de protection des données, ainsi qu'entre les ASM et les organismes visés à l'article 77 (comme le prévoit leur coopération en vertu des Articles 73(7) et 79(2) du règlement sur l'IA) joueront un rôle clé à cet égard (Autoriteit Persoonsgegevens 2024, paragraphe 37). L'existence de ces bases juridiques constitue une condition préalable à la création d'accords de coopération bilatéraux et multilatéraux pour le partage d'informations. Ces accords devraient être rédigés de manière générale pour couvrir les compétences des organismes au titre de l'article 77 et des ASM et à permettre, le cas échéant, la conclusion d'accords bilatéraux supplémentaires spécifiques à des cas particuliers (Autoriteit Persoonsgegevens 2024, paragraphe 37).

Exemples et bonnes pratiques

Au Royaume-Uni, la Commission pour l'égalité et les droits humains (EHRC) a collaboré avec le Bureau de l'information dans le cadre d'une compétition sur l'équité et l'innovation, sur des cas pilotes dans l'enseignement supérieur, la finance, la santé et le recrutement, afin de trouver de nouvelles façons de traiter les discriminations et les biais statistiques, humains et structurels dans les systèmes d'IA. La compétition a été financée par le ministère des Sciences du Royaume-Uni (Department of Science, Technology and Innovation 2024).

En Norvège, l'organisme de promotion de l'égalité a coopéré avec l'autorité de protection des données dans le cadre d'un bac à sable réglementaire sur l'intelligence artificielle (Autorité de protection des données norvégienne 2023).

En France, la coopération entre le Défenseur des droits et l'autorité de protection des données, la Commission nationale de l'informatique et des libertés (CNIL), est facilitée par plusieurs mécanismes : les deux institutions disposent d'un mémorandum d'entente et le Défenseur des droits joue également un rôle consultatif auprès du conseil d'administration de la CNIL. En raison de ce lien

15 Règlement (UE) 2026/1744 du Parlement européen et du Conseil du 8 juillet 2026 modifiant les règlements (UE) 2024/1689, (UE) 2018/1139 et (UE) 2023/1230 en ce qui concerne la simplification de la mise en œuvre des règles harmonisées concernant l'intelligence artificielle (train de mesures omnibus numérique sur l'IA), disponible : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=OJ:L_202601744 (consulté le 24 juillet 2026).

16 Voir également l'article 26 de la Convention-cadre du Conseil de l'Europe sur l'IA (2024), « pour faciliter une coopération efficace entre eux », et l'article 14 des directives relatives aux normes exigeant des mécanismes de coopération appropriés.

structurel, le Défenseur des droits est informé de tous les projets soumis aux séances plénières du conseil d'administration de la CNIL (Conseil de l'Europe 2025b).

Au Royaume-Uni, un mémorandum d'entente entre le Bureau du Commissaire à l'information (APD) et l'EHRC (l'organisme de promotion de l'égalité) définit les principes de partage d'informations et de coopération, les bases juridiques sur lesquelles le partage d'informations peut avoir lieu, les modalités pratiques relatives au contrôle et au traitement des données et les rôles, fonctions et pouvoirs respectifs de chaque organisme. Voir : mémorandum d'entente

In the UK, a memorandum of understanding between the Information Commissioner's Office (the DPA) and the Equality and Human Rights Commission (the equality body) sets out principles of information sharing and co-operation, the legal bases on which information sharing may take place, practicalities relating to data control and processing and the respective roles, functions and powers of each body. See: Memorandum of Understanding

Catégories de risque en vertu du règlement sur l'IA

Le règlement européen sur l'IA instaure un cadre réglementaire qui classe les systèmes d'IA en fonction d'un niveau présumé de risque qu'ils représentent au regard des droits fondamentaux, de la santé et de la sécurité (règlement sur l'IA, article 1). Cette approche reflète l'idée selon laquelle tous les systèmes d'IA ne présentent pas le même niveau de préoccupation : certains présentent des risques inacceptables et doivent être totalement interdits, d'autres des risques élevés qui justifient une surveillance stricte, tandis que d'autres sont considérés comme ne nécessitant aucune réglementation spécifique à l'IA.

La catégorie de risque détermine les obligations applicables, qui vont de l'interdiction absolue à des exigences techniques complètes en passant par les obligations de transparence. Pour les organismes de promotion de l'égalité, cette structure fondée sur les risques est directement pertinente : elle détermine quelles informations et quels documents devraient exister et elle influence les recours disponibles en vertu du règlement sur l'IA.

IA interdite

Le règlement européen sur l'IA interdit certaines pratiques en matière d'IA qui présentent des risques inacceptables pour les droits fondamentaux, la sécurité ou les valeurs démocratiques (article 5 du règlement sur l'IA). Ces pratiques comprennent : Technique de manipulation qui altèrent de manière significative le comportement

- les techniques de manipulation qui altèrent de manière significative le comportement des personnes, y compris la génération de contenu sexuel et intime non consensuel ou de matériel représentant des abus sexuels sur des enfants ;
- l'exploitation de vulnérabilités ;
- les systèmes de notation sociale ;

- les évaluations des risques des personnes physiques visant à prédire la commission d'infractions pénales sur la base d'un profilage ;
- l'identification biométrique à distance en temps réel dans les espaces publics (hormis quelques rares exceptions) ;
- la reconnaissance des émotions dans des contextes professionnels ou éducatifs ;
- la catégorisation biométrique déduisant des caractéristiques protégées ;
- la collecte d'images faciales par moissonnage non ciblé.

Voir aussi Annexe J – Pratiques interdites en matière d'IA.

IA à haut risque

Les systèmes d'IA à haut risque comprennent des cas d'utilisation qui sont considérés comme présentant des risques considérables pour la santé, la sécurité ou les droits fondamentaux et qui sont utilisés dans des domaines spécifiques listés en Annexe III du règlement sur l'IA (article 6). Ils comprennent des systèmes utilisés dans les domaines suivants :

- la biométrie ;
- la gestion des infrastructures critiques ;
- l'éducation et formation professionnelle ;
- la gestion de l'emploi et des travailleurs et travailleuses ;
- l'accès à des services et avantages essentiels ;
- l'application de la loi ;
- la migration, l'asile et les contrôles aux frontières ;
- l'administration de la justice.

Remarque : un système d'IA visé en Annexe III sera toujours considéré comme étant à haut risque lorsqu'il effectue le profilage d'une personne physique.

Voir aussi Annexe I – Pratiques d'IA à haut risque

Si le système d'IA est destiné à effectuer l'une des opérations suivantes, le système ne sera pas considéré comme étant à haut risque (article 6(3)) :

- une tâche procédurale limitée ;
- améliorer les résultats d'une activité humaine achevée précédemment ;
- détecter les schémas décisionnels ou les écarts par rapport aux schémas décisionnels antérieurs ;
- une tâche préparatoire à une évaluation.

Les systèmes d'IA à haut risque sont assujettis à des exigences de documentation, à des obligations de transparence et à des procédures de gestion des risques en vertu du règlement sur l'IA.

IA à usage général (General-purpose AI - GPAI)

Les catégories à haut risque et interdites englobent des cas d'utilisation spécifiques et la catégorisation repose sur l'objectif explicite des systèmes. En revanche, les modèles d'IA à usage général, par exemple les grands modèles de langage (LLM), sont capables d'exécuter un large éventail de tâches distinctes et peuvent être intégrés dans divers systèmes ou applications en aval (règlement sur l'IA, article 3(63)). Par définition, la GPAI n'a pas d'objet spécifique. Toutefois, un modèle de GPAI peut être à la base d'applications spécifiques ayant des objectifs spécifiques qui pourraient relever de la catégorie à haut risque.

Il est peu probable que la documentation rédigée en conséquence des obligations relatives à la GPAI soit à la disposition des OPE en vertu du règlement sur l'IA, car les droits d'accès sont limités aux applications de systèmes d'IA à haut risque (du règlement sur l'IA, article 77). Les OPE pourraient néanmoins se prévaloir des compétences qui leur sont déjà reconnues ainsi que des dispositions du droit national ; toutefois, dans la plupart des cas, l'obtention d'une ordonnance judiciaire sera très probablement nécessaire. Les modèles de GPAI sont largement inclus dans la catégorie des systèmes algorithmiques aux fins de la présente méthodologie.

Comment la méthodologie applique-t-elle les catégories de risque ?

Comme indiqué précédemment, le règlement sur l'IA confère de nouvelles compétences permettant d'obtenir des informations concernant les systèmes à haut risque. La méthodologie est donc principalement structurée autour de la distinction entre les systèmes d'IA à haut risque et les autres systèmes algorithmiques. En ce qui concerne les systèmes d'IA ne présentant pas de haut risque, les OPE sont appelés à intervenir dans le même cadre informationnel que celui qui s'applique au reste des systèmes algorithmiques. L'étape 1 de la méthodologie aide les personnes chargées de l'instruction des dossiers à comprendre à quel type de système elles sont confrontées.

Le règlement sur l'IA impose des exigences supplémentaires aux différents acteurs impliqués dans le développement, la distribution et le déploiement de systèmes d'IA à haut risque. Pour ces systèmes, les exigences réglementaires prévues par le règlement sur l'IA fourniront des informations importantes et pertinentes sur les mesures de gestion des risques, les données utilisées et les mesures de gouvernance des données en place, une documentation technique expliquant les principaux choix de conception et de données et la finalité prévue du système, ainsi que des informations concernant les tests et le suivi, ainsi que les risques de discrimination identifiés. Cette documentation sera une ressource essentielle pour les organismes de promotion de l'égalité qui cherchent à identifier ou évaluer d'éventuels cas de discrimination algorithmique.

Ces informations ne seront pas disponibles dans tous les cas, ni même dans la plupart d'entre eux. Toutefois, les exigences réglementaires applicables aux systèmes d'IA à haut risque aident également les OPE à identifier et à interpréter les éléments de preuve concernant l'IA et d'autres systèmes algorithmiques, même lorsque aucune

obligation formelle comparable n'existe. En effet, ces exigences servent de référence pour identifier, réduire ou éliminer le risque de discrimination et la documentation fait souvent partie du développement logiciel/technique standard.

La méthodologie vise tout au long de son développement à distinguer dans l'ensemble les circonstances dans lesquelles un OPE est habilité, en vertu de l'article 77 du règlement sur l'IA, à demander et à consulter des documents créés ou conservés en vertu du règlement sur l'IA en ce qui concerne les systèmes à haut risque et les cas dans lesquels ce pouvoir n'existe pas.

Il est important de garder à l'esprit que les cadres juridiques nationaux confèrent souvent aux organismes de promotion de l'égalité des pouvoirs de collecte d'informations ou d'enquête. Les organismes de promotion de l'égalité qui ne sont pas des autorités au titre de l'article 77 ne disposent pas d'une base juridique directe en vertu du règlement sur l'IA pour accéder à la documentation pertinente concernant les systèmes d'IA à haut risque. Ils pourront toutefois s'appuyer sur les compétences de collecte d'informations et d'enquête qui leur sont reconnues par le droit national ainsi que par les autres dispositions pertinentes du droit de l'Union.

Caractéristiques protégées

Des caractéristiques protégées sont mentionnées dans l'ensemble de cette méthodologie. Ce sont des caractéristiques protégées par la Convention européenne des droits de l'homme (la Convention) et les directives européennes.

La Cour européenne des droits de l'homme a publié un Guide sur l'article 14 de la Convention européenne des droits de l'homme et l'article 1 du Protocole n° 12 à la Convention.

Tableau 1. Caractéristiques protégées couvertes par le droit européen
X dans le tableau signifie que la caractéristique est couverte par la réglementation. Une cellule vide signifie qu'elle n'est pas couverte.

	Article 14 de le Convention européenne des droits de l'homme	Charte des droits fondamentaux de l'UE	Directives européennes sur la non-discrimination
Sexe	X	X	X (2004/113/CE, 2000/78/CE, 2006/54/CE, 2023/970/UE, 2010/41/UE)
« Race »[17]	X	X	X (2000/43/CE)

17. Tous les êtres humains appartiennent à la même espèce, le Comité des Ministres du Conseil de l'Europe rejette les théories fondées sur l'existence de « races » différentes. Toutefois, dans le présent document, le terme « race » est utilisé pour éviter que les personnes qui sont généralement et de manière erronée perçues comme appartenant à une « autre race » ne soient exclues de la protection prévue par la législation et de la mise en œuvre des politiques.

	Article 14 de le Convention européenne des droits de l'homme	Charte des droits fondamentaux de l'UE	Directives européennes sur la non-discrimination
Couleur	X	X	
Langue	X	X	
Religion	X	X (ou croyance)	X (2000/78/CE)
Opinion politique ou autre	X	X	X (2000/78/CE)
Origine nationale ou sociale	X	X (origine ethnique ou sociale)	
Appartenance à une minorité nationale	X	X (appartenance)	
Fortune	X	X	
Naissance	X	X	
Autre statut	X		
Handicap	Autre statut	X	X (2000/78/CE)
Âge	Autre statut	X	X (2000/78/CE)
Orientation sexuelle	Autre statut	X	X (2000/78/CE)
Nationalité	Autre statut	X (qualifié)	
Caractéristiques génétiques	Autre statut	X	
État civil	Autre statut		
Statut de personne migrante ou réfugiée	Autre statut		

Comment utiliser la méthodologie

La méthodologie adopte une approche en quatre étapes (voir également Illustration 1 ci-dessous).

Étape 1 : l'identification des cas liés à l'IA ou aux algorithmes

Étape 2 : la collecte d'informations sur l'IA ou le système algorithmique pour identifier la discrimination

Étape 3 : la méthodologie d'évaluation des informations reçues

Étape 4 : les solutions

Le cas échéant, la méthodologie établit une distinction entre les différentes sources des cas :

A. une plainte (d'une personne ou d'un groupe)
B. une enquête *ex officio*
C. les informations fournies par une ASM

Remarque : les étapes sont présentées consécutivement pour que cette méthodologie soit structurellement claire. Elles ne sont pas réellement consécutives mais, dans la pratique, souvent circulaires et itératives.

Étape 1

Le signalement initial ou les informations relatives à un cas de discrimination dans lequel l'implication de systèmes algorithmiques peut être reconnue.

Questions-clés :

- S'agit-il d'un cas impliquant un système d'IA à haut risque en vertu du règlement sur l'IA ?
- S'agit-il d'un cas impliquant un système algorithmique ?

Étape 2

L'étape suivante consiste à rechercher des informations concernant le système identifié, ce qui peut contribuer à l'identification ultérieure de la discrimination. L'étape 2 définit les types d'informations pertinentes, de documentation et les autres types de preuves qui peuvent être disponibles ou accessibles à l'OPE et leurs sources potentielles.

Remarque : les informations identifiées à cette étape sont analysées en détail à l'étape 3. Par conséquent, les étapes 2 et 3 devraient être prises en compte conjointement. De même, lorsqu'il manque des informations et qu'il existe un pouvoir légal de les obtenir, ce recours est abordé à l'étape 4.

Question clé :

- Quelles sont les informations disponibles et comment y accéder ?

Étape 3

Une analyse complète des éléments de preuve pertinents est effectuée au moyen d'une série de questions suggérées à titre indicatif. Ces questions ne sont pas exhaustives.

Ensuite, le cadre juridique de la non-discrimination est appliqué à ces éléments de preuve.

Questions clés :

- Une preuve prima facie de discrimination peut-elle être établie ?
- Est-il probable que l'organisation intimée soit en mesure de s'acquitter de la charge de la preuve ?
- Compte tenu de l'ensemble des éléments de preuve, une discrimination interdite a-t-elle été établie ?

Étape 4

Envisagez et appliquez les solutions à disposition.

Question clé :

- Quelles solutions devraient être recherchées ou appliquées ?

À toutes les étapes, la personne en charge du dossier devra tenir compte de ce qui suit.

- En l'absence d'allégation de discrimination défendable, l'affaire ne devrait pas être poursuivie.

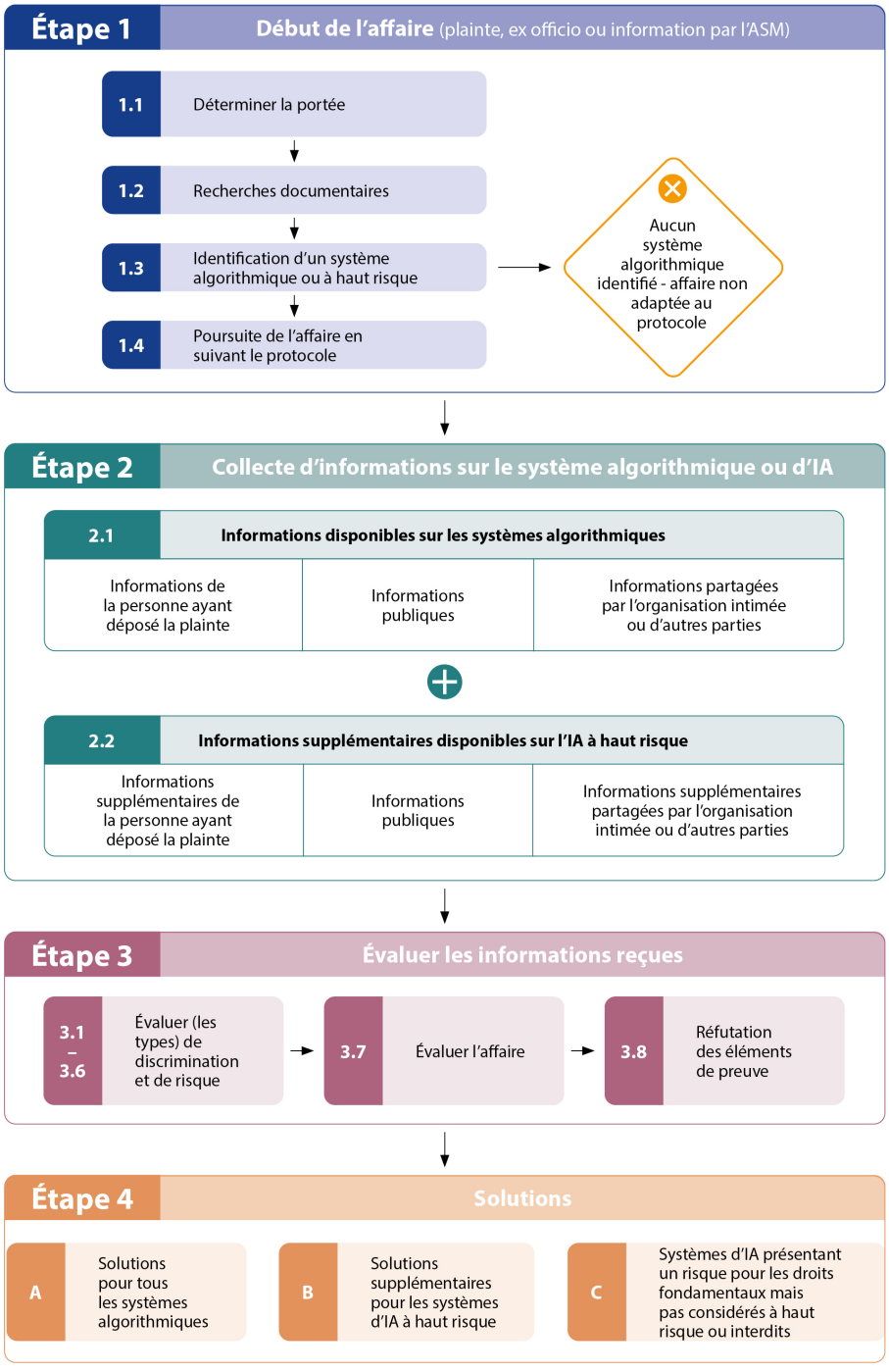
- Si aucun système algorithmique n'est impliqué, l'affaire ne devrait plus être classée selon la présente méthodologie et devrait être examinée dans le cadre de la procédure ordinaire de l'OPE en matière de discrimination.

Recours à des expert·es

Il peut être nécessaire de solliciter l'assistance d'expert·es pour certains éléments de cette méthodologie, par exemple lorsqu'une expertise en données ou en informatique est requise pour l'évaluation technique soit du modèle, soit des données. La méthodologie indique les types de questions auxquels un·e expert·e pourrait être invité·e à répondre et quand cela devrait être envisagé. Si les amendements à l'article 77 de l'Omnibus numérique sont adoptés[18], les organismes de promotion de l'égalité pourront demander l'assistance de l'ASM, d'autres autorités publiques ou d'organismes disposant d'une expertise pertinente, en s'appuyant sur la disposition d'assistance mutuelle prévue à l'article 77(1)(b) proposé.

18 Voir ci-dessus, sous la rubrique « Règlement sur l'IA, article 77 : pouvoirs des autorités chargées de la protection des droits fondamentaux ».

Illustration 1. Étapes 1 à 4 de la méthodologie



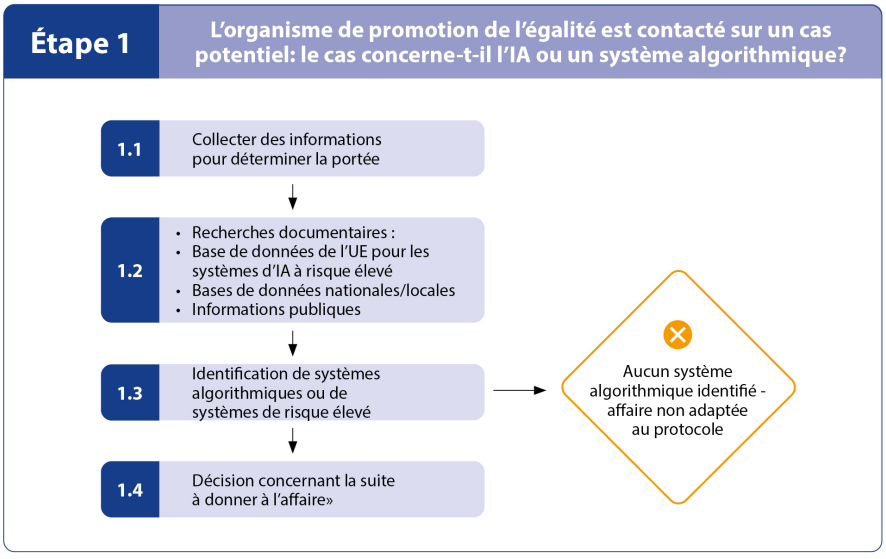
Étape 1

Identification des systèmes d'IA ou algorithmiques

L'objectif de cette première étape est uniquement de comprendre si un système d'ADM ou d'IA (à haut risque) est impliqué et de déterminer si les informations sont suffisantes pour passer à l'étape 2. Lors de cette étape, aucune investigation détaillée n'est effectuée.

L'étape 1 comprend les sous-étapes suivantes.

Illustration 2. Étape 1 – Le cas concerne-t-il l'IA ou un système algorithmique ?



Remarque : les étapes sont présentées consécutivement pour que ce protocole soit structurellement clair. Ces sous-étapes ne sont pas nécessairement consécutives mais, dans la pratique, elles sont souvent circulaires et itératives.

L'identification des cas d'IA et d'ADM diffère selon la source du cas :

A. suite à une plainte (d'un individu ou d'un groupe)
B. en conséquence d'une enquête *ex officio*
C. par le biais d'informations fournies par une ASM

Comment reconnaître les systèmes algorithmiques et systèmes d'IA

Les systèmes algorithmiques peuvent être aussi simples que l'automatisation d'un ensemble limité de règles humaines. En général, les systèmes algorithmiques (y compris l'IA) sont difficiles à reconnaître. Les définitions ont changé au fil des ans ou sont interprétées différemment.

Dans le règlement européen sur l'IA, la définition technique à l'article 3(1) stipule ce qui suit :

> « système d'IA », un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels[19].

Même lorsqu'il existe un consensus sur les définitions, les systèmes algorithmiques ne sont souvent pas reconnus par les organisations ou sont désignés par des termes différents.

Une évaluation visant à déterminer si un système algorithmique est utilisé devrait couvrir un large éventail de termes.

Tous les termes suivants pourraient indiquer l'utilisation de systèmes algorithmiques.

- Automatisation, automatisé, informatisé, basé sur un algorithme ;
- Profilage ou création de profils ;
- Prise de décision automatisée ou systèmes d'aide à la prise de décision ;
- Un système ou un ordinateur décide/analyse/compare/recommande/traite les données d'une personne ;
- Un système, un ordinateur ou une caméra voit/écoute/comprend ;
- Toute mention de calcul de scores, de probabilité, de vraisemblance ou d'estimation de comportement à venir ;
- Évaluation automatisée des risques, reconnaissance automatisée ;
- Toute mention de statistiques ;

19. Cette définition est complétée par les lignes directrices de la Commission C(2025) 5053 sur la définition d'un système d'intelligence artificielle établi par le règlement (UE) 2024/1689 (règlement sur l'IA).

- Toute mention de machine learning.

Les mots suivants pourraient indiquer l'utilisation de l'IA.

- Machine learning, apprentissage profond ;
- Un modèle/une machine/un ordinateur a été entraîné·e ;
- Des modèles ont été appris à partir de données ;
- Basé sur de grands modèles de langage (LLM) ;
- Modèles prédictifs.

Exemple : « Sur la base de ces données, notre fournisseur calcule des valeurs de probabilité au moyen de méthodes mathématico-statistiques afin d'évaluer le risque de défaut de paiement et de vérifier votre adresse. »

1.1. Collecte d'informations initiale

A. Plainte

Lorsque le cas trouve son origine dans une plainte, la première étape correspond au dépôt de celle-ci. À ce stade, des informations initiales peuvent être collectées quant à l'implication potentielle d'un système algorithmique ou d'IA. Cette collecte initiale d'informations est volontairement limitée, afin de s'aligner autant que possible sur les processus existants de réception de plaintes au sein des organismes de promotion de l'égalité.

1.1.1. Détermination initiale du champ d'application et de la compétence

- Source de la plainte : téléphone, formulaire en ligne, courrier, autre
- Nom et coordonnées de la personne ayant déposé la plainte
- Bref compte-rendu du problème
- Quelle est l'organisation intimée (organisation qui fait l'objet de la plainte) ?
  - Organisme public ou privé ?
- Évaluation des compétences
  - Motifs protégés ?
  - Secteur ?
- Solution recherchée/objet visé par le cas
  - Donner des conseils ?
  - Poursuivre un litige ?
  - Recommander une action ?

Ce cas relève-t-il de la compétence de l'organisme de promotion de l'égalité ? La plainte a-t-elle déjà été déposée auprès d'un autre organisme/d'une autre entité ou une autre entité est-elle déjà impliquée dans le cas ?

1.1.2. Collecte d'informations supplémentaires concernant la plainte

La personne ayant déposé la plainte peut potentiellement disposer de plus d'informations. Cette section définit des exemples de questions à poser à cette personne, soit par le biais d'un formulaire en ligne, soit lors d'une conversation (téléphonique), afin de déterminer si un système algorithmique ou d'IA a été impliqué.

Tant le règlement sur l'IA que le RGPD imposent des obligations de transparence à l'égard des individus (personnes concernées ou personnes affectées) quant à l'utilisation de systèmes algorithmiques ou d'IA : voir les encadrés contenant les dispositions pertinentes ci-dessous. Des informations pertinentes peuvent être fournies aux personnes concernées ou aux personnes affectées par le biais d'une correspondance personnelle ou par des accords contractuels, généralement inclus dans les avis de confidentialité ou les conditions générales, ou autrement communiquées publiquement par le biais du site Web ou des plateformes de l'organisation (voir 1.2.3 Site Web et communications externes de l'organisation intimée)

Par conséquent, il est utile, dès ce stade, de demander aux plaignant·es de communiquer tout document pertinent en leur possession et de procéder à un examen sommaire des informations accessibles au public, afin de faciliter les étapes ultérieures de la méthodologie (voir 1.2 Recherche documentaire complémentaire)

Remarque : ces dispositions ne couvrent pas tous les cas dans lesquels des systèmes d'IA ou algorithmiques sont utilisés et, par conséquent, l'absence d'informations peut simplement signifier que des enquêtes supplémentaires sont nécessaires.

Plusieurs dispositions prévues par le règlement sur l'IA exigent que les personnes soient informées de l'utilisation de l'IA :

- lorsqu'un système d'IA à haut risque est utilisé sur le lieu de travail, les représentant·es des travailleurs et travailleuses ainsi que les travailleurs et travailleuses concerné·es doivent être informé·es au préalable de l'utilisation d'un système d'IA à haut risque (règlement sur l'IA, article 26(7)) ;
- lorsque des systèmes d'IA à haut risque prennent des décisions ou aident à prendre des décisions concernant des personnes physiques, ces personnes doivent être informées au préalable par le développeur qu'elles font l'objet de l'utilisation du système d'IA à haut risque (règlement sur l'IA, article 26(11)). Cette information doit inclure l'objectif visé et les types de décisions prises. Le déployeur devrait également informer les personnes physiques de leur droit à une explication prévue dans le cadre de ce règlement (considérant 93 du règlement sur l'IA) ;
- lorsque des personnes physiques interagissent directement avec l'IA (par le biais de chatbots ou d'assistants vocaux, par exemple), elles devraient en être informées (règlement sur l'IA, article 50(1)). Cette information peut passer par des notifications claires ou des pop-ups comme : « Salut, je suis un chatbot »/« Bonjour, vous discutez actuellement avec un assistant IA »/« Je m'appelle X, je suis un assistant IA et je suis là pour vous aider » ;
- de même, les individus doivent être informés du fonctionnement des systèmes de reconnaissance des émotions (un système d'IA qui identifie la colère ou

la satisfaction des client·es par des caractéristiques faciales ou vocales, par exemple) ou des systèmes de catégorisation biométrique (un système d'IA qui effectue une identification du genre fondée sur la forme des cheveux ou du visage, par exemple) (règlement sur l'IA, article 50(3)).

RGPD

- Le RGPD impose également des exigences de transparence (RGPD, article 5(1)(a), 13, 14) qui requièrent que la personne concernée soit informée de l'utilisation de systèmes algorithmiques, ceci incluant des informations sur l'identité de la personne responsable du traitement, les finalités du traitement et l'existence de la prise de décision automatisée et du profilage (voir étape 2.1.1 Informations fournies par la personne ayant déposé la plainte).

Ces informations peuvent figurer dans les conditions générales ou dans l'avis ou la déclaration de respect de la vie privée. En ce qui concerne le règlement sur l'IA, le Bureau de l'IA a publié des lignes directrices relatives aux obligations de transparence en juillet 2026[20].

Questions à poser à la personne ayant déposé la plainte et/ou à prendre en considération pour la recherche documentaire

Remarque : une personne ayant déposé une plainte peut ne pas savoir de quelles informations elle dispose ou si ces informations sont pertinentes. Par conséquent, demander tout simplement « Quels sites/quelles applications utilisiez-vous ? » et « Vous a-t-on remis des papiers/documents ? » est un point de départ utile.

Les documents pertinents peuvent inclure, inter alia, une demande soumise par la personne ayant déposé la plainte, des lettres, des courriels, des factures ou des formulaires de consentement et des avis de confidentialité envoyés par l'organisation intimée, des contrats entre la personne ayant déposé la plainte et l'organisation, y compris les annexes et les mises à jour contractuelles fournies au format papier, par courriel ou par le biais de notifications d'application ou de portail. Si elle est disponible, la documentation pertinente devrait être collectée et consultée à l'étape 1.2 Recherche documentaire complémentaire pour répondre à certaines, voire à toutes les questions suivantes.

- La personne ayant déposé la plainte a-t-elle été informée qu'un système d'IA était ou est utilisé ?
  - Lorsque le cas concerne un lieu de travail : la personne ayant déposé la plainte a-t-elle consulté les représentant·es des travailleurs·ses ou de son syndicat ?
  - Si c'est le cas : que lui a-t-on dit ?
- La personne ayant déposé la plainte a-t-elle été informée que le traitement de données automatisé a été utilisé ?[21]

20. Pour en savoir plus, cliquez ici : Lignes directrices sur les obligations de transparence pour les fournisseurs et les déployeurs de systèmes d'IA.
21. Lorsque des données à caractère personnel sont utilisées pour une prise de décision entièrement automatisée, les utilisateurs et utilisatrices auraient dû être informé·es qu'ils et elles font l'objet d'une prise de décision entièrement automatisée (RGPD, article 13(2)(f)).).

- La personne ayant déposé la plainte a-t-elle été invitée à donner son autorisation au traitement spécifique des données ou à la confirmer en ayant lu une déclaration ou un avis de respect de la vie privée ?
  - Si c'est le cas : le texte de la demande d'autorisation, de la déclaration ou de l'avis de respect de la vie privée peuvent-ils être partagés ?
  - Si ce n'est pas le cas : y a-t-il eu une autre mention de « voici comment nous traitons vos informations » ou « voici comment vos informations peuvent être utilisées » ?
- A-t-il été fait mention de l'automatisation, d'un système informatique exécutant des tâches autonomes ou d'un système informatique apportant sa contribution à une décision ? Recherchez des phrases et des éléments de langage comme les suivants :
  - Un système ou un ordinateur décide/analyse/compare/recommande/traite les données d'une personne ;.
  - Un système, un ordinateur ou une caméra « voit »/« écoute »/« comprend » ;
  - Toute mention du calcul de scores, de probabilité, de vraisemblance ou d'estimation de comportement à venir ;
  - L'utilisation d'un système informatique pour soutenir un processus, potentiellement avec un nom de marque. Par exemple, « Nous utilisons Clippy/Frank/Maya pour traiter vos données et soutenir notre travail ».
- Un contrat entre la personne ayant déposé la plainte et l'organisation intimée contient-il une référence à l'automatisation, à des processus algorithmiques, à des systèmes spécifiquement nommés, à des statistiques, à des mesures de performances, à une évaluation ou à des scores (comme la désactivation automatisée ou les scores de fiabilité) ?
  - Si c'est le cas : le contrat ou les clauses contractuelles pertinentes peuvent-ils être partagés ??
- Où l'interaction ou la communication avec l'organisation (intimée) a-t-elle eu lieu ?
- En cas d'interaction en personne ou par téléphone :
  - en personne, l'autre personne a-t-elle consulté un ordinateur pour parvenir à des décisions ou à des conclusions ?
- En cas d'interaction par le biais d'un ordinateur ou d'un smartphone :
  - un formulaire en ligne a-t-il été rempli ?
  - combien de temps après l'interaction un résultat a-t-il été obtenu ou une décision a-t-elle été prise (instantanément, quelques minutes à une heure après, quelques jours plus tard, plus longtemps) ?
  - l'autorisation de traiter des données à caractère personnel a-t-elle été demandée, par exemple au moyen d'une case à cocher mentionnant les données à caractère personnel ou la vie privée ?
  - des informations sur un motif protégé (voir Caractéristiques protégées) ont-elles été demandées lors de l'interaction avec le système ?

1.1.3. Passage en revue rapide des informations publiques

Base de données de l'UE pour les systèmes d'IA à haut risque

À partir du 2 août 2026, une base de données de l'UE[22] devrait être disponible pour les systèmes à haut risque avant leur introduction sur le marché ou leur mise en service. Selon l'Omnibus numérique sur l'IA, les systèmes mis en service avant le 2 décembre 2027 pour les systèmes autonomes d'IA à haut risque et le 2 août 2028 pour les systèmes d'IA à haut risque intégrés dans des produits ne nécessiteront d'enregistrement que s'ils font l'objet d'un « changement significatif »[23]. Les parties publiques de la base de données seront accessibles à toutes et tous.

En cas de plaintes relatives aux systèmes utilisés dans les domaines de l'éducation, de l'emploi, du gouvernement, de la banque, des assurances, de l'administration de la justice et des élections, et en ce qui concerne les actions/activités comme l'automatisation de l'identification des individus, la reconnaissance faciale ou vocale, la reconnaissance des émotions, la prédiction des intentions ou des risques, une recherche rapide dans la base de données (à l'aide de mots clés pertinents qui pourraient être, par exemple, le nom de l'organisation intimée, le nom du système ou le secteur concerné) peut être un point de départ utile.

- Si l'organisation intimée est inscrite dans la base de données comme jouant un rôle quel qu'il soit dans un cas d'utilisation pertinente pour la plainte, il existe une certitude relative que l'IA (à haut risque) est impliquée[24]. Veuillez passer à l'étape 1.3.

Remarque : il existe certaines exceptions aux exigences d'enregistrement, où il se peut qu'un fournisseur ou un déployeur n'ait pas respecté ses obligations réglementaires. Vous trouverez plus informations sur les renseignements disponibles dans la base de données à l'étape 2.2.2 : Informations publiques..

Comparer à la liste des cas d'utilisation

- L'Annexe C contient une liste des utilisations courantes des systèmes algorithmiques et de l'IA dans différents secteurs. L'un de ces exemples semble-t-il pertinent pour la situation, pour le cas ou pour le traitement de la plainte ? Voir aussi l'Annexe F des bases de données et des cas d'utilisation de l'IA.

22 Bien que cette base de données n'existe pas encore et que l'on ne sache pas à quoi ressembleront sa conception et son interface, elle est mentionnée ici parce qu'elle devrait devenir une source d'information extrêmement pertinente à l'avenir.

23 Règlement (UE) 2026/1744 du Parlement européen et du Conseil du 8 juillet 2026 modifiant les règlements (UE) 2024/1689, (UE) 2018/1139 et (UE) 2023/1230 en ce qui concerne la simplification de la mise en œuvre des règles harmonisées concernant l'intelligence artificielle (train de mesures omnibus numérique sur l'IA), disponible : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=OJ:L_202601744 (consulté le 24 juillet 2026). Dans les autres cas où le système à haut risque est destiné à être utilisé par les autorités publiques, les démarches nécessaires doivent être entreprises pour se conformer au règlement d'ici août 2030, règlement sur l'IA, article 111(2).

24 Veuillez noter que lorsqu'un système n'est pas enregistré dans cette base de données, cela ne signifie pas qu'aucun système d'IA (à haut risque) n'est impliqué. Des exceptions existent.

Recherche générale sur internet

- Effectuez une recherche sur internet pour l'innovation/l'automatisation dans la situation, le cas où le traitement décrit par la plainte ou par l'organisation intimée.
  - Par exemple, « l'IA est-elle utilisée par la police [nationalité] ? ». À titre d'exemple, pour la Finlande, cette recherche aboutit à une liste de résultats, y compris : la Finlande prévoit une expansion majeure des pouvoirs de collecte de données biométriques de la police – ID Tech.
  - Par exemple, « l'IA est-elle utilisée dans le recrutement en [pays] ? ». Pour la Belgique, cette recherche aboutit au résultat suivant, suggérant qu'un tiers des entreprises belges utilisent l'IA dans les processus de recrutement : www.anews.com.tr/europe/2023/07/19/in-belgium-one-third-of-companies-use-artificial-intelligence-in-their-recruitment-processes.
  - Veillez à ce que les résultats soient des résultats de recherche et non des réponses générées par IA.

Remarque : certaines pratiques sont interdites en vertu de l'article 5 du règlement sur l'IA. Voir Annexe J – Pratiques interdites en matière d'IA pour accéder à la liste.

Ces étapes ne sont données qu'à titre indicatif. Si la réponse à ces questions démontre qu'il est probable qu'un système d'IA ou un système algorithmique ait été utilisé, veuillez passer à l'étape 1.3.

Si les informations ne sont pas concluantes, envisagez de collecter davantage d'informations (publiques) pendant la recherche documentaire, comme décrit à l'étape 1.2.

B. *Ex officio*

Certains organismes de promotion de l'égalité sont habilités, en vertu du droit national, à ouvrir des enquêtes *ex officio* ou *actio popularis*.

Pendant cette phase, il est présumé qu'aucune plainte émanant d'une personne ou d'un groupe de personnes n'est en cause. Au cours des étapes ultérieures de l'enquête, l'organisme de promotion de l'égalité pourra toutefois décider d'impliquer les personnes directement concernées.

Aux fins de la présente méthodologie d'évaluation, il est supposé qu'un domaine d'intérêt a déjà été identifié (un secteur, un domaine opérationnel au sein de ce secteur et potentiellement, des organisations spécifiques). Il peut s'agir, par exemple, d'algorithmes de tarification utilisés par les assurances ou des systèmes de notation du risque de fraude appliqués aux prestations sociales.

Informations à collecter

- La situation, l'affaire ou le processus faisant l'objet de l'examen, ainsi que les informations accessibles au public.
- Les plaintes pertinentes (non résolues) enregistrées par l'organisme de promotion de l'égalité ou d'une autre entité compétente. Il peut s'agir de

plaintes qui, prises isolément, ne présentaient pas un fondement suffisant ou qui ont été retirées par leur auteur·e.

- Toute indication permettant d'établir l'implication de l'IA ou des systèmes algorithmiques, si disponible. Voir la section Comment reconnaître les systèmes algorithmiques et l'IA ci-dessus.
- Les coordonnées de l'organisation qui ferait l'objet de cette enquête ex officio si elle devait être lancée officiellement (ci-après dénommée « l'organisation intimée ») ou, dans le cas d'un passage en revue du secteur, une liste d'organisations représentatives du secteur faisant l'objet du passage en revue[25]

Comparer à une liste d'affaires existantes

- L'Annexe C contient une liste des utilisations courantes des systèmes algorithmiques et de l'IA dans différents secteurs. L'un de ces exemples semble-t-il pertinent pour la situation, l'affaire ou le traitement de l'enquête potentielle ?

Lorsqu'une affaire est initiée *ex officio*, l'implication d'un système d'IA peut déjà être évident. Les étapes 1.2 et 1.3 ne sont pas pertinentes en l'espèce : vous pouvez passer à l'étape 1.4.

C. Information par l'ASM

L'organisme de promotion de l'égalité est censé recevoir des informations de l'autorité de surveillance des marchés (ASM) concernant les systèmes d'IA qui présentent un risque de discrimination, en particulier dans les cas suivants.

- **Incidents graves (règlement sur l'IA, article 73(7))**
  - Lorsqu'un fournisseur notifie à l'ASM un incident grave[26], cette dernière doit en informer les autorités relevant de l'article 77, y compris l'organisme de promotion de l'égalité, le cas échéant. La consultation de la Commission sur une assistance spécifique concernant les procédures de signalement et sur un projet de modèle de rapport d'incident s'est achevée le 7 novembre 2025.[27]
- **Produits présentant un risque (règlement sur l'IA, article 79 en conjonction avec l'article 3(19), du règlement (UE) 2019/1020)**
  - Lorsque l'ASM a des raisons suffisantes de croire qu'un système d'IA relève des produits présentant un risque pour le droit à la non-discrimination, elle a le

25. Par exemple, si un OPE ouvre une enquête ex officio sur les algorithmes de tarification utilisés par les assurances, une liste représentative d'assurances au lieu d'une organisation intimée spécifique pourrait être utilisée à ce stade du protocole.

26 Un incident grave désigne un incident ou un dysfonctionnement d'un système d'IA qui entraîne directement ou indirectement l'un des événements suivants :

a) le décès d'une personne ou une atteinte grave à la santé d'une personne ;

b) une perturbation grave et irréversible de la gestion ou du fonctionnement d'infrastructures critiques ;

c) la violation des obligations au titre du droit de l'Union visant à protéger les droits fondamentaux ;

d) un dommage grave à des biens ou à l'environnement (règlement sur l'IA, article 3(49)).

27. En savoir plus : Législation sur l'IA : la Commission publie un modèle de déclaration pour les incidents graves impliquant des modèles d'IA à usage général présentant un risque systémique, 4 novembre 2025.

devoir d'informer et de coopérer pleinement avec les autorités compétentes au titre de l'article 77, y compris l'organisme de promotion de l'égalité.

Ces informations peuvent inclure les éléments suivants.

- Organisation : organisme public ou privé.
- Système d'IA : identification du système et du domaine de classification à haut risque (annexe I ou annexe III du règlement sur l'IA, y compris le domaine et le cas d'utilisation).
- Rôle de l'organisation (déployeur, fournisseur, autre).
- Lorsque la notification est fondée sur l'article 73 du règlement sur l'IA (incident grave) :
  - les circonstances de l'incident ou du dysfonctionnement ;
  - le type de préjudice qui s'est produit ou qui a été suspecté ;
  - La relation entre le système d'IA et l'incident (comment l'IA a causé l'incident ou y a contribué) ;
  - toute information disponible sur l'évaluation des risques de l'enquête du fournisseur concernant l'incident et les mesures correctives proposées ou mises en œuvre.
- • Lorsque la notification est fondée sur l'article 79 du règlement sur l'IA (risque) :
  - la raison pour laquelle le système d'IA est considéré comme présentant un risque pour les droits fondamentaux ;
  - (facultatif, le cas échéant) des informations pertinentes sur les évaluations déjà réalisées ;
  - (facultatif, si disponible) mesures correctives proposées ou mises en œuvre.
- Dans les affaires impliquant un système d'IA à haut risque dans les domaines de la mise en application de la loi, de la migration, de l'asile et de la gestion des contrôles aux frontières (points 1, 6 et 7 de l'annexe III), l'ASM pourra accéder aux informations sur le système d'IA telles qu'enregistrées dans la base de données de l'UE pour les systèmes d'IA à haut risque.

Lorsqu'un organisme de promotion de l'égalité est informé par une ASM, il est certain qu'un système d'IA à haut risque est impliqué. Les étapes 1.2 et 1.3 ne sont pas pertinentes en l'espèce : vous pouvez passer à l'étape 1.4.

1.2. Recherche documentaire complémentaire

Si les ressources disponibles permettent d'effectuer des recherches documentaires, plusieurs sources pourraient être consultées pour déterminer si l'ADM ou l'IA est susceptible d'avoir été utilisée dans le cas concerné. Cette section détaille les sources et les éléments auxquels il convient de prêter attention.

Les étapes mentionnées dans cette section varient en termes d'étendue et de temps requis pour effectuer cette recherche documentaire :

- ▲ indique les étapes qui pourraient être prioritaires ;
- ▼ indique des étapes supplémentaires pour une recherche documentaire plus approfondie.

Par exemple, consulter la base de données de l'UE (ci-dessous) devrait être relativement facile une fois celle-ci opérationnelle et si l'organisation intimée est enregistrée dans la base de données. Cette consultation fournira une réponse rapide à la question « un système d'IA est-il utilisé ? ». Si le déployeur n'est pas enregistré, d'autres recherches seront nécessaires.

Cette section ne s'applique pas aux cas où la source est « C. Information par l'ASM ». Dans ce cas, l'ASM aura déjà déterminé qu'un système d'IA à haut risque est impliqué et, par conséquent, que la recherche documentaire décrite dans cette section serait redondante.

1.2.1. Base de données de l'UE pour les systèmes d'IA à haut risque

Recherchez l'organisation intimée ou tout déployeur ou fournisseur connu dans la base de données.

L'organisation intimée est-elle inscrite en tant que déployeur dans la partie C de la base de données ?[28]

- ▲ L'organisation intimée est-elle le déployeur dans un cas d'utilisation pertinent pour la plainte ? Quelle organisation est le fournisseur ?

Informations pertinentes disponibles dans la base de données à des fins d'identification initiale :

- ► les coordonnées du fournisseur et de son représentant autorisé ;
- ► le nom et les coordonnées du déployeur (uniquement les autorités publiques, les institutions, organes et agences de l'UE ou les personnes agissant en leur nom) ;
- ► l'URL de l'entrée du système d'IA dans la base de données de l'UE par son fournisseur.

L'organisation intimée figure-t-elle dans la base de données en tant que fournisseur dans la partie A ou B de la base de données ?

- ▲ l'organisation intimée est-elle le fournisseur dans un cas d'utilisation à haut risque qui est pertinent pour la plainte ?
- ▲ l'organisation intimée est-elle le fournisseur dans un cas d'utilisation destiné à un domaine à haut risque, mais qui a été auto-évalué comme relevant de

28. Veuillez noter que seules les autorités publiques sont tenues de s'enregistrer en tant que déployeurs. D'autres organisations n'ont pas cette obligation.

l'exception prévue à l'article 6(3) du règlement sur l'IA ? Cette évaluation peut être demandée par les autorités nationales compétentes.

Informations pertinentes disponibles dans la base de données à des fins d'identification initiale :

- les coordonnées du fournisseur et de son représentant autorisé ;
- la dénomination commerciale ou autre référence permettant l'identification et le traçage du système d'IA (ces informations pourraient être utiles, par exemple lorsqu'un système à haut risque est utilisé dans le cadre d'un autre système) ;
- l'usage prévu.

Si ces informations figurent dans cette base de données, alors un système d'IA est utilisé. Veuillez passer à l'étape 1.3 et/ou 1.4.

Si ce n'est pas le cas, recherchez dans la base de données des organisations similaires à l'organisation intimée :

▼ Des organisations similaires sont-elles répertoriées dans la base de données en tant que fournisseurs ou déployeurs dans un cas d'utilisation pertinent pour la plainte ? Par exemple, d'autres assurances similaires sont répertoriées dans la base de données comme utilisant des systèmes d'IA pour la tarification lorsque l'affaire concerne une tarification (in)juste, ou d'autres gouvernements similaires dans le même pays ont déclaré utiliser des systèmes d'IA pour l'évaluation de prestations lorsque l'affaire concerne le refus de prestations..

En vertu de l'article 49(4) du règlement sur l'IA, l'enregistrement de systèmes d'IA à haut risque dans les domaines de l'application de la loi, de la migration, de l'asile et de la gestion des contrôles aux frontières doit avoir lieu dans une section non publique sécurisée de la base de données de l'UE visée à l'article 71 et est limité à un ensemble spécifique d'informations, défini à l'article 49(4). L'accès à ces sections de la base de données sera limité.

Pour en savoir plus sur la section non publique de la base de données et sur les aspects de transparence dans le domaine de l'application de la loi, voir restrictions d'accès aux informations à l'étape 2.2.2 et l'encadré Domaine d'application de la loi à l'étape 2.2.4, respectivement.

1.2.2. Autres bases de données et inventaires pour les systèmes d'IA (à haut risque) et les systèmes algorithmiques

Certaines organisations et gouvernements nationaux tiennent à jour une base de données locale, un registre ou un inventaire des systèmes d'IA utilisés. Vous trouverez une liste de ces bases de données à l'annexe F. Toutes ces bases de données ne sont pas pertinentes. Vérifiez uniquement celles qui sont pertinentes dans votre contexte national.

Recherchez l'organisation intimée dans la base de données :

- ▲ Un système algorithmique ou d'IA est-il répertorié comme étant utilisé ou ayant été développé par l'organisation intimée et est-il pertinent dans la situation concernée, dans l'affaire ou le traitement visé par la plainte ?

Recherchez des organisations similaires à l'organisation intimée :

- ▼ Un système algorithmique ou d'IA est-il répertorié comme étant utilisé ou ayant été développé par une organisation comparable et aurait-il été pertinent dans la situation concernée, dans l'affaire ou le traitement couvert par la plainte, si cette plainte concernait cette organisation comparable ?

Comparez à la liste des cas d'utilisation existants :

- ▼ Annexe C pour une liste des utilisations courantes des systèmes algorithmiques et d'IA dans différents secteurs.

L'un de ces cas d'utilisation semble-t-il pertinent dans la situation concernée, dans l'affaire ou dans le traitement de la plainte ?

1.2.3. Site Web et communications externes de l'organisation intimée

Avis/déclaration de respect de la vie privée

Recherchez des formulations indiquant l'utilisation de systèmes algorithmiques, par exemple l'un des éléments suivants.

- ▲ Des objectifs qui suggèrent une automatisation, la création de profils ou de méthodes statistiques. Reportez-vous à la section Comment reconnaître des systèmes algorithmiques et d'IA et aux questions contenues dans le chapitre 1.1.2 Questions à poser à la personne ayant déposé la plainte et/ou à prendre en considération pour la recherche documentaire.
- ▲ Des tiers recevant des données à caractère personnel (destinataires ou catégories de destinataires). Vérifiez les indications d'implication de tiers, comme des sociétés de notation ou des agences de référence de crédit, qui peuvent impliquer un profilage ou des décisions automatisées..
- ▼ Coordonnées de la personne responsable du traitement des données : la personne responsable du traitement des données est-elle connue pour mettre en œuvre des solutions d'IA ou d'automatisation dans le secteur d'intérêt ? S'il existe un système d'IA ou algorithmique, la personne responsable du traitement des données sera généralement le déployeur du système et l'organisation intimée au sens de cette méthodologie..

Autres types de communication

Est-il fait mention d'un système algorithmique qui pourrait être pertinent dans l'affaire concernée ?

- ▼ Recherchez sur le site Web de l'organisation intimée des articles sur la mise en œuvre de l'IA, l'innovation et l'automatisation ou des FAQ sur le site Web/la plateforme de l'organisation intimée. La façon la plus simple de rechercher un site Web spécifique est d'ajouter « site : [site Web de l'organisation intimée] » à la requête de recherche. Par exemple, le site d'IA d'innovation : bigbank.com.

- Vérifiez les offres d'emploi techniques pour les spécialistes des données (data scientists), de l'IA ou de l'apprentissage automatique. Est-il fait mention d'un système algorithmique qui pourrait être pertinent dans l'affaire concernée ? Une description de poste pour un « data scientist senior » dans une banque contient ce qui suit : « Prévenir le blanchiment d'argent et le financement du terrorisme est l'un des aspects clés de notre société et de notre organisation… Vous serez responsable de l'optimisation de notre système actuel de surveillance des transactions (transaction monitoring ou TM), sur la base des règles et éventuellement de l'introduction de méthodologies d'analyse plus avancées telles que l'apprentissage automatique dans nos modèles de TM. » Voir data scientist senior surveillance des transactions – Rabobank.

Si applicable : passez en revue le site Web et les communications des déployeurs de cas d'utilisation pertinents de l'IA à haut risque.

Si des cas d'utilisation pertinents pour l'affaire sont répertoriés par des déployeurs tiers dans la base de données de l'UE pour les systèmes d'IA à haut risque, consultez le site Web du fournisseur.

- L'organisation intimée est-elle répertoriée sur le site Web en tant « qu'organisation avec laquelle nous travaillons » ou « à laquelle ces clients font confiance » ?
- L'organisation intimée est-elle mentionnée d'une autre manière sur le site Web du fournisseur ? Un moteur de recherche sur internet serait le moyen le plus simple de trouver ces informations. Par exemple, recherchez le site de [l'organisation intimée] : [fournisseur].com.

1.2.4. Passage en revue plus approfondi des informations publiques

Certaines des étapes mentionnées ici ont déjà été mentionnées à l'étape 1.1. Par souci d'exhaustivité, elles sont répétées ici.

Recherche générale sur internet

- Voir recherche internet à l'étape 1.1.3.
- Affiner la recherche effectuée à l'étape 1.1.3 en vérifiant spécifiquement si des organisations non gouvernementales (ONG), des organismes nationaux ou internationaux ont signalé l'utilisation d'IA ou de systèmes algorithmiques dans une situation, une affaire ou un traitement de plainte similaires, soit par l'organisation intimée, soit par des organisations similaires. Les ONG concernées sont Amnesty, AlgorithmWatch et des organisations de la société civile nationale. Aux Pays-Bas, par exemple, l'APD produit un rapport sur les risques liés à l'IA et aux algorithmes deux fois par an. La recherche pour le contexte national néerlandais pourrait être « [secteur de la plainte] Autoriteit Persoonsgegevens rapport de risques IA et algorithmes ».
- Option : affiner la recherche et inclure des stratégies nationales et internationales en matière d'IA.
- Les autres ressources pertinentes pourraient inclure le Système de surveillance des incidents et des risques liés à l'IA de l'OCDE (disponible en anglais : AI incidents and hazards monitor) et la base de données des incidents d'IA ou les rapports produits par le Conseil de l'Europe tels que l'étude intitulée

Discrimination, intelligence artificielle et décisions algorithmiques (Zuiderveen Borgesius, 2018).

1.3. Identification des systèmes algorithmiques ou d'IA à haut risque

Les informations recueillies aux étapes 1.1 et 1.2 permettent de déterminer raisonnablement si un système algorithmique ou un système d'IA à haut risque a été utilisé, ainsi que le degré de certitude de cette conclusion.

Rappel : cette méthodologie opère uniquement une distinction entre les systèmes d'IA à haut risque et les systèmes algorithmiques. Les systèmes algorithmiques comprennent les systèmes d'IA qui ne sont pas à haut risque et les systèmes d'ADM. Pour les systèmes d'IA à haut risque, le règlement sur l'IA impose des exigences étendues en matière de transparence ainsi que des obligations en matière de documentation, ce qui signifie que des documents supplémentaires devraient être disponibles. Si un système d'IA n'est pas à haut risque, les informations disponibles et les étapes à suivre sont les mêmes que pour les systèmes algorithmiques plus simples et par conséquent, elles sont regroupées en une seule catégorie.

Type de système

- **Système d'IA à haut risque**
- **Système algorithmique**

Certitude

- **Relativement certain**
- **Potentiellement**

Remarque : conclure si des systèmes d'IA ou des systèmes algorithmiques sont utilisés constitue une appréciation délicate et nuancée à ce stade, les informations disponibles étant très limitées. Les degrés de certitude associés aux conclusions énumérées ci-dessous sont données à titre indicatif et ont uniquement pour objet d'orienter l'analyse.

Base de données de l'UE pour les systèmes d'IA à haut risque

Le système est-il enregistré dans la base de données de l'UE ?

- OUI dans la section A ou C : un système d'IA à haut risque a été identifié pour cette affaire. Il est **relativement certain** qu'un **système d'IA à haut risque** est utilisé dans cette affaire.
- OUI, dans la section B[29] un système d'IA est impliqué mais aucune documentation requise pour les systèmes à haut risque ne devrait être disponible. Il est **relativement certain** qu'un **système algorithmique** ia été utilisé dans cette affaire.

29. En vertu de l'article 6(3) du règlement sur l'IA, le système d'IA est considéré comme ne présentant pas de haut risque.

- NON, mais des organisations similaires (comme une autre banque) ont listé des systèmes pertinents pour l'affaire (comme des systèmes de notation de crédit) dans la section A ou C. Par conséquent, il peut être raisonnable de supposer qu'un système similaire à haut risque est utilisé par l'organisation intimée. Pour cette étape de la méthodologie, supposons qu'il est **relativement certain** qu'un **système algorithmique** est utilisé dans cette affaire. **Potentiellement, un système d'IA à haut risque** a été utilisé dans cette affaire.

Informations fournies par l'organisation intimée (soit à la personne ayant déposé la plainte, soit dans des documents publics, soit sur demande

- L'organisation intimée a mentionné l'utilisation de l'automatisation, d'algorithmes ou de l'IA à des fins pertinentes dans cette affaire, soit directement à la personne ayant déposé la plainte, soit dans l'avis/la politique de respect de la vie privée. Il est **relativement certain** qu'un **système algorithmique** a été utilisé dans cette affaire.
- L'organisation intimée a déclaré utiliser l'automatisation, des algorithmes ou de l'IA à des fins pertinentes pour les catégories d'IA à haut risque (voir annexe I), soit directement à la personne ayant déposé la plainte, soit dans l'avis/la politique de respect de la vie privée. Il est **relativement certain** qu'un **système algorithmique** est utilisé dans cette affaire. **Potentiellement,** un **système d'IA à haut risque** a été utilisé dans cette affaire..
- L'organisation intimée a listé un système algorithmique ou d'IA dans une base de données publique (voir annexe F) à des fins pertinentes pour l'affaire. Il est **relativement certain** qu'un **système algorithmique** a été utilisé dans cette affaire.
- • L'organisation intimée a publiquement déclaré qu'elle utilisait des systèmes algorithmiques ou d'IA à des fins pertinentes pour l'affaire dans des articles de blog, des communiqués de presse ou d'autres sources. Notez que les références à l'utilisation de systèmes algorithmiques peuvent être vagues ou évasives (voir étape 1.1 pour les phrases potentielles utilisées). Il est **relativement certain** qu'un **système algorithmique** a été utilisé dans cette affaire.

Autres informations

- Des sources crédibles comme des ONG et des organisations médiatiques de confiance ont signalé que l'organisation intimée utilisait des systèmes algorithmiques ou d'IA à des fins pertinentes pour l'affaire. Aucune confirmation de la part de l'organisation n'a été trouvée. Il est **relativement certain** qu'un **système algorithmique** a été utilisé dans cette affaire.
- D'autres parties ont signalé l'utilisation de systèmes algorithmiques ou d'IA par l'organisation intimée à des fins pertinentes pour l'affaire. Aucune confirmation de la part de l'organisation intimée ne peut être trouvée. **Potentiellement,** un **système algorithmique** a été utilisé dans cette affaire.
- L'IA et les systèmes algorithmiques sont fréquemment utilisés à des fins pertinentes pour l'affaire, comme indiqué en annexe C. **Potentiellement,** un **système algorithmique** a été utilisé dans cette affaire.

- Des organisations similaires sont signalées comme utilisant des systèmes algorithmiques à des fins pertinentes pour l'affaire. **Potentiellement,** un **système algorithmique** a été utilisé dans cette affaire
- • Des indices contextuels indiquent que des systèmes algorithmiques peuvent être impliqués dans l'affaire. Par exemple, en raison de la rapidité de la prise de décision, lorsque des processus complexes ou des décisions qui nécessitent de la discrétion sont traités instantanément ou en quelques minutes. **Potentiellement,** un **système algorithmique** a été utilisé dans cette affaire
- Des indices contextuels indiquent que des systèmes algorithmiques peuvent être impliqués dans l'affaire. Par exemple, en raison de la rapidité de la prise de décision, lorsque des processus ou des décisions complexes nécessitant de la discrétion sont traités instantanément ou en quelques minutes et que soit l'IA et les systèmes algorithmiques sont fréquemment utilisés à des fins pertinentes pour l'affaire, comme indiqué en annexe C, soit des organisations similaires sont signalées comme utilisant des systèmes algorithmiques à des fins pertinentes pour l'affaire. Il est **relativement certain** qu'un **système algorithmique** est utilisé dans cette affaire.

1.4. Poursuite de l'affaire

Les étapes 1.1 à 1.3 portent sur l'évaluation de la probabilité qu'un système algorithmique soit impliqué. Afin de décider s'il y a lieu de poursuivre le traitement de l'affaire, il convient de procéder à une évaluation complémentaire relative à la probabilité que celle-ci soit suffisamment étayée sur le plan de la discrimination. Cette probabilité devrait être évaluée sur le fondement de critères établis au sein de l'organisme de promotion de l'égalité. Lorsqu'un système algorithmique est potentiellement impliqué, il peut être nécessaire d'appliquer un seuil de certitude moins élevé avant de décider de poursuivre l'examen de l'affaire.

***Manjang v. Uber Eats UK Ltd & Others* (2021)**

Cette affaire de discrimination raciale a été introduite sur le fondement du harcèlement, des mesures de rétorsion (victimisation) et de discrimination indirecte suite à l'introduction d'une technologie de reconnaissance faciale. Le requérant, un livreur, a reçu un courriel d'Uber Eats (l'organisation intimée) l'informant de sa suspension permanente de la plateforme au motif qu'il « partageait son compte partenaire de livraison Uber Eats » et qu'il avait échoué à un contrôle de reconnaissance faciale. Il a répondu en déclarant : « votre algorithme fondé sur l'apparence est raciste et ce problème doit être abordé car il ne peut pas reconnaître et vérifier mes photos, ce qui est probablement la raison pour laquelle on me demande de prendre des photos de moi-même plusieurs fois par jour ». Uber Eats a contesté l'allégation de discrimination et a demandé son exclusion à un stade précoce parce que, selon l'entreprise, elle était fondée sur des « erreurs factuelles ». La société a ensuite fourni des explications concernant leurs moyens de défense qu'elle invoquait concernant le fonctionnement de la technologie, à la suite de quoi le requérant a modifié sa plainte. Le tribunal a refusé de rejeter

la demande modifiée, considérant qu'au moment où le requérant avait plaidé sa cause, la raison pour laquelle son compte avait été désactivé n'était pas très claire. Quelques mois plus tard, au cours de la procédure, l'organisation intimée a déclaré que les motifs donnés au requérant, selon lesquels la décision était automatisée, étaient inexacts et qu'un examen humain avait eu lieu. Le requérant a soutenu, et le tribunal a convenu, qu'une certaine opacité entourait la raison pour laquelle les photos du requérant avaient été rejetées et son compte désactivé. L'affaire s'est finalement conclue par un règlement transactionnel.

Il s'agit d'un exemple d'affaire financée et soutenue par un OPE, en l'occurrence la Commission pour l'égalité et les droits humains - EHRC (Royaume-Uni).

Tableau 2. Certitude de système à haut risque ou algorithmique et probabilité de discrimination

	Faible probabilité de discrimination	Probabilité moyenne de discrimination	Forte probabilité de discrimination
IA à haut risque relativement certaine	Affaire hautement prioritaire, envisagez l'implication d'un spécialiste (interne) des systèmes algorithmiques	Affaire hautement prioritaire, envisagez l'implication d'un spécialiste (interne) des systèmes algorithmiques	Affaire hautement prioritaire, envisagez l'implication d'un spécialiste (interne) des systèmes algorithmiques
Système algorithmique relativement certain potentiellement IA à haut risque	Affaire modérément prioritaire, envisagez l'implication d'un spécialiste (interne) des systèmes algorithmiques	Affaire hautement prioritaire, envisagez l'implication d'un spécialiste (interne) des systèmes algorithmiques	Affaire hautement prioritaire, envisagez l'implication d'un spécialiste (interne) des systèmes algorithmiques
Système algorithmique relativement certain	Affaire peu prioritaire, en cas de poursuite envisagez l'implication d'un spécialiste (interne) des systèmes algorithmiques	Affaire modérément prioritaire, envisagez l'implication d'un spécialiste (interne) des systèmes algorithmiques	Affaire hautement prioritaire, envisagez l'implication d'un spécialiste (interne) des systèmes algorithmiques
Potentiellement système algorithmique	Se référer aux méthodologies existantes, au lieu de cette méthodologie	Affaire peu prioritaire dans le cadre de cette méthodologie, où la probabilité de discrimination est moyenne, mais la probabilité du système d'IA est faible ; se référer à des méthodologies existantes, au lieu de cette méthodologie	Affaire modérément prioritaire, envisagez l'implication d'un spécialiste (interne) des systèmes algorithmiques

Suspicion de classification incorrecte en vertu du règlement sur l'IA

Classification erronée comme étant hors haut risque

Il est possible que les fournisseurs de systèmes d'IA évitent le régime d'IA à haut risque dans les zones visées à l'annexe III en procédant à une auto-évaluation de l'une des manières suivantes.

- Pas un système d'IA. Dans ce cas, le règlement sur l'IA ne s'appliquerait pas.
- Pas un système à haut risque en raison d'exceptions (règlement sur l'IA, article 6(3)). Dans ce cas, le système devrait toujours être enregistré dans la base de données de l'UE conformément à la section B. Le fournisseur est alors tenu de documenter son évaluation pour choisir de ne pas participer avant l'introduction sur le marché ou la mise en service du système d'IA, et d'enregistrer le système dans la base de données de l'UE conformément à la section B (règlement sur l'IA, article 49(2)). La publication de l'évaluation n'est pas obligatoire mais le fournisseur est tenu de fournir la documentation de l'évaluation à la demande des autorités nationales compétentes (ASM ou organisme de notification).
- Il ne s'agit pas d'un système à haut risque car l'objectif poursuivi n'est pas visé par les cas d'utilisation énumérés en annexe III du règlement (règlement sur l'IA, article 6(3)).

Si un organisme de promotion de l'égalité soupçonne qu'un système a été considéré à tort comme n'entrant pas dans le champ d'application du régime d'IA à haut risque, il devrait en informer l'ASM concernée, qu'un cas de discrimination puisse être établi ou non (voir étape 4, procédure de l'ASM pour traiter les systèmes d'IA classés par le fournisseur comme ne présentant pas de haut risque, et étape 4.3).

Aux fins de la présente méthodologie, le système devrait être considéré comme un système algorithmique, étant donné que les informations qui devraient être disponibles pour les systèmes d'IA à haut risque ne seront probablement pas disponibles dans ce cas.

Questions à poser

- Le système est-il couvert par les cas d'utilisation de l'annexe III ?
- Le système est-il susceptible d'être un système d'IA ? Voir Comment reconnaître les systèmes algorithmiques et l'IA.
- Fait-il partie de la liste des exemples pratiques ? (Voir annexe C et/ou la liste établie par la Commission lorsqu'elle sera disponible, article 6(4)).
- Vérifiez la section B de la base de données. Le fournisseur du système a l'obligation de s'enregistrer lui-même et d'enregistrer le système en vertu de l'article 49

Suspicion de système interdit

Certaines pratiques sont interdites en vertu du règlement sur l'IA (voir annexe J – Pratiques interdites en matière d'IA).

Si l'OPE découvre une pratique ou un système interdit :

- étant donné que ces pratiques sont interdites, l'utilisation elle-même viole le droit de l'UE. Des mesures immédiates devraient être prises pour notifier l'ASM concernée, qu'un cas de discrimination puisse être établi ou non ;

- en outre, lorsqu'un OPE conclut qu'une pratique actuellement non incluse dans cette liste viole les droits fondamentaux, il doit rassembler les preuves et fournir des rapports aux autorités compétentes de l'État membre concerné, ainsi qu'à la Commission (voir étape 4.3).

Les affaires impliquant des pratiques interdites peuvent également être des systèmes à haut risque et/ou avoir causé de la discrimination, et l'OPE devra décider si des mesures supplémentaires sont nécessaires pour lutter contre la discrimination.

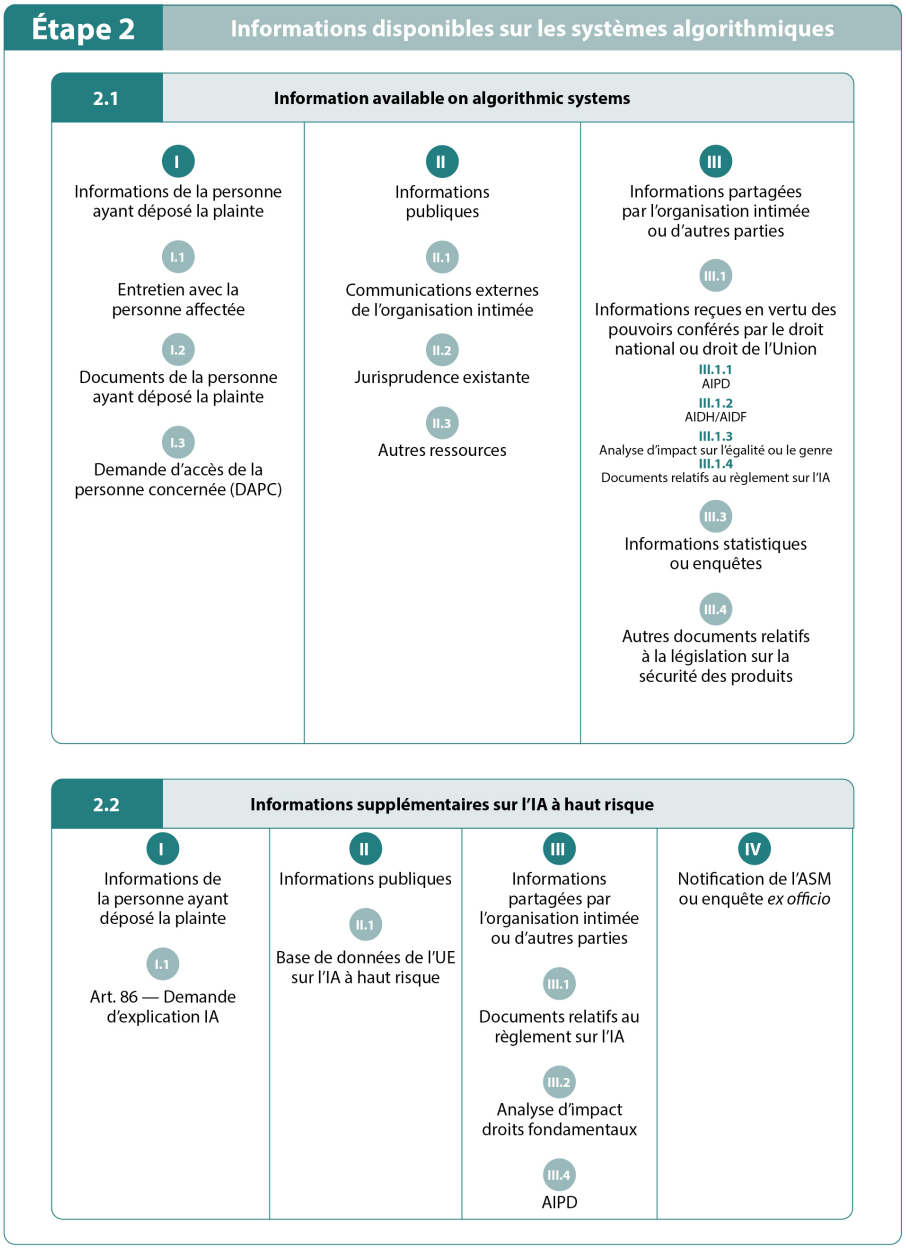
Étape 2

Collecte d'informations sur le système algorithmique ou d'IA afin d'identifier la discrimination

L'étape 2 identifie les types de renseignements qui peuvent être disponibles ou accessibles concernant le système et la probabilité d'établir la discrimination. Elle fournit des conseils sur les sources potentielles et sur la manière de recueillir les informations.

L'étape 2 comprend les sous-étapes suivantes.

Illustration 3. Étape 2 – Collecte d'informations afin d'identifier la discrimination



Si un OPE n'est pas une autorité relevant de l'article 77 et/ou si le système n'est pas identifié comme présentant un haut risque, l'étape 2.2 (par exemple, évaluation de l'impact sur l'égalité et informations statistiques ou issues d'enquêtes) peut toujours être utilisée comme un cadre conceptuel pour permettre à tout OPE de demander des informations comparables au déployeur ou au fournisseur, en s'appuyant sur ses pouvoirs d'enquête ou de collecte d'informations, et pour évaluer toute réponse ou

absence de réponse au titre de l'étape 3. Toutefois, veuillez noter que les organisations intimées ne sont pas légalement tenues, en vertu du règlement sur l'IA, de fournir ce type de documentation dans ces circonstances (bien qu'elles puissent l'être en vertu d'autres lois, par exemple la sécurité des produits ou le droit de la consommation).

L'analyse de ces éléments de preuve recueillis à l'étape 2 est effectuée à l'étape 3. Dans la pratique, ces étapes sont généralement réalisées ensemble.

À l'étape 2, l'on a identifié avec un certain degré de certitude qu'un système algorithmique ou un système d'IA est impliqué dans l'affaire examinée, et qu'il est potentiellement catégorisé en tant que système d'IA à haut risque[30]. L'étape suivante consiste à rechercher des informations sur ce système, ces informations pouvant contribuer à l'identification ultérieure de la discrimination à l'étape 3.

Les sources peuvent appartenir à trois catégories.

1. Les informations pouvant être fournies à l'OPE par la personne ayant déposé la plainte. Ces informations peuvent être disponibles ou accessibles aux personnes ayant déposé la plainte en raison de leur expérience personnelle du processus contesté, de leur correspondance avec l'organisation intimée et de l'exercice de leurs droits à l'information, principalement en vertu du RGPD et du règlement sur l'IA.
2. D'autres informations peuvent être rendues publiques.
3. Les informations que l'OPE peut demander à l'organisation intimée dans l'exercice de ses pouvoirs nationaux d'enquête et de collecte d'informations ou découlant du droit de l'Union, y compris ceux qui lui sont conférés par le règlement européen sur l'IA.

Veuillez garder à l'esprit que des informations peuvent également être divulguées ou demandées lors de la correspondance préalable à l'action ou dans le cadre d'un litige.

30 En cas d'impossibilité d'identifier un système d'IA ou algorithmique, cessez de suivre ce protocole et revenez à vos procédures habituelles.

Remarque : l'étape 2 ne cherche pas à être exhaustive. D'autres types d'informations ou sources d'information peuvent également être disponibles. Les cadres juridiques nationaux, y compris la législation sur l'accès à l'information ou les règles relatives à la transparence des documents, peuvent conférer aux OPE des pouvoirs de collecte d'informations ou d'enquête allant au-delà de ceux visés dans la méthodologie. L'accès à l'information peut également être soutenu par de nouveaux pouvoirs d'enquête découlant des directives relatives aux normes (voir Introduction et étape 4).

Obstacles connus à l'obtention d'informations

Les stratégies déployées pour obtenir des informations, telles que les demandes d'accès à l'information, peuvent se heurter à des difficultés et à des contraintes, notamment en ce qui concerne les arguments relatifs à la confidentialité et aux exigences opérationnelles. En France, par exemple, une demande d'accès à l'information formulée par les organisation de la société civile (OSC) pour clarifier un système de notation des risques employé par la Caisse nationale des allocations familiales n'a entraîné que la publication du code source des anciens modèles (par opposition au modèle actuel) et des listes de variables utilisées (Conseil de l'Europe 2025 ; La Quadrature du Net e.a contre Premier ministre et Ministère de la Culture 2024 ; voir aussi Big Brother Watch 2025).

Dans l'affaire Kelly v. National University of Ireland (C-104/10), dans laquelle le requérant alléguait une discrimination fondée sur le sexe après s'être vu refuser l'accès à un programme de perfectionnement professionnel et avoir demandé à l'établissement d'enseignement des informations sur les qualifications d'autres candidats de l'établissement d'enseignement, la Cour de justice de l'Union européenne (CJUE) a précisé que la législation dérivée de l'Union ne confère pas à un tel individu le droit d'accéder à ce type de données. Toutefois, la CJUE a estimé qu'il appartient aux juridictions nationales de déterminer si, dans les circonstances de l'affaire, la divulgation publique de ces données est nécessaire pour donner effet aux objectifs de la législation pertinente, à savoir la directive 97/80/CE.

De même, à l'avenir, la CJUE préciser davantage l'interprétation des dispositions pertinentes des directives en matière de non-discrimination afin d'assurer une protection judiciaire efficace contre la discrimination, d'une manière qui facilite l'accès à l'information des personnes ayant déposé une plainte dès la phase menant à l'affaire prima facie (Farkas et O'Farrell 2015).

Il peut être utile de conserver un registre, pouvant être géré et passé en revue, des sources consultées. Il pourrait aider à identifier les schémas de discrimination et faire gagner un temps précieux aux personnes en charge du dossier. Les rubriques suggérées pour les passages en revue de la littérature pourraient inclure :

- source (avec un lien) ;
- mots clés (discrimination par proxy, atténuation de biais, par exemple) ;
- résumé ;
- pertinence (discrimination constatée dans le système de reconnaissance faciale, par exemple).

2.1. Systèmes algorithmiques

2.1.1. Informations fournies par la personne ayant déposé une plainte

La personne ayant déposé une plainte aura souvent des informations très pertinentes à fournir pour identifier une discrimination potentielle. Elle pourra disposer de ces informations ou y avoir accès en raison de son expérience personnelle du processus de contestation ou de sa correspondance avec l'organisation intimée. En outre, ces informations peuvent être fondées sur les droits à l'information de la personne ayant déposé à la plainte, notamment ceux découlant du RGPD (voir encadré ci-dessous).

Lorsque l'affaire est traitée dans le cadre d'une enquête *ex officio* ou est identifiée par une ASM, aucune personne ayant déposé la plainte n'est directement impliquée. Aux fins de la collecte d'informations, veuillez tenir compte des éléments suivants :

- impliquez les personnes ayant déposé d'autres plaintes/des plaintes anciennes/incomplètes ou examinez si une plainte est déjà déposée auprès d'un autre organisme/d'une autre entité.
- le cas échéant, suggérez l'exercice collectif du droit d'accès au RGPD dans les affaires ayant un enjeu élevé[31]

31. Bien que le RGPD, le règlement sur l'IA et la Directive « Police-Justice » (DPJ) offrent des possibilités limitées de représentation collective, les chercheurs ayant traité ces questions ont soutenu que l'exercice collectif des droits individuels peut néanmoins être fructueux (Mahieu, Asghari et van Eeten, 2018). Par exemple, plusieurs personnes pourraient soumettre des demandes d'accès, et les informations qui en résulteraient pourraient être partagées avec un intermédiaire de confiance par le biais de dons de données. Ces efforts collectifs, par des actions individualisées, pourraient créer un réservoir de connaissances plus large concernant l'utilisation des données par des acteurs ou concernant des secteurs spécifiques. De même, l'exercice collectif (mais individuel) d'un droit d'accès en vertu des articles 15 et 22 du RGPD, ou du droit à une explication en vertu de l'article 86 du règlement sur l'IA, par les personnes concernées ou affectées, pourrait offrir des passerelles pour acquérir une connaissance plus étendue du fonctionnement des systèmes d'IA. Toutefois, étant donné les amendements proposés au droit d'accès en vertu du RGPD dans la proposition de règlement de l'Omnibus numérique, l'efficacité des demandes d'accès en tant qu'outils de transparence individuels et collectifs pourrait se voir considérablement réduite. Pour en savoir plus, veuillez consulter la section : Préparer et obtenir une réponse à une demande d'accès d'une personne concernée (DAPC) conformément à l'article 15 du RGPD.

Obligations de transparence en vertu du RGPD et des droits des données

Le RGPD inclut des obligations de transparence ex ante et des droits d'accès ex post, qui permettent aux personnes concernées d'être informées de la juste utilisation de leurs données à caractère personnel

Ex ante

- En vertu de l'article 5(1) et du considérant 60 du RGPD, les données à caractère personnel sont traitées de manière juste et transparente par rapport à la personne concernée, qui est informée de l'existence du traitement et de ses finalités. La personne responsable du traitement devrait fournir à la personne concernée toute autre information nécessaire pour assurer un traitement équitable et transparent, en tenant compte des circonstances et du contexte spécifiques dans lesquels les données à caractère personnel sont traitées.
- En vertu des articles 13 et 14 du RGPD, lorsque des données à caractère personnel sont traitées, les personnes concernées reçoivent des informations sur l'identité et le contact de la personne responsable du traitement des données, la finalité du traitement, la base juridique du traitement, les catégories de données à caractère personnel dans lesquelles les données sont obtenues auprès de tiers ou de sources publiques, et l'existence d'une prise de décision automatisée, y compris le profilage[32].

Ces informations *ex ante* sont de nature plus générales et souvent incluses dans les avis ou déclarations de respect de la vie privée.

ADM admissible

- En vertu des articles 13(2)(f) et 14(2)(g) du RGPD, les personnes concernées doivent être explicitement informées de l'existence de processus décisionnels et de profilage exclusivement automatisés visés à l'article 22 du RGPD, et recevoir des informations utiles concernant la logique sous-jacente ainsi que l'importance et les conséquences envisagées d'un tel traitement pour la personne concernée.

La décision exclusivement automatisée est autorisée en vertu de l'article 22(2) du RGPD uniquement si cette décision est :

- nécessaire à la conclusion ou à l'exécution d'un contrat ;
- autorisée par le droit de l'Union ou de l'État membre ;
- fondée sur le consentement explicite de la personne concernée. Les OPE devraient également pouvoir identifier les dispositions pertinentes du contrat, la loi autorisant le traitement ou le consentement et ce à quoi la personne concernée a donné son consentement, par exemple dans une déclaration écrite et signée, un formulaire électronique ou un courriel.

32. Il semble y avoir un consensus de plus en plus large quant au fait que ce partage d'informations s'applique également aux ADM non admissibles, en conséquence des deux articles exigeant des personnes responsables du traitement qu'elles divulguent les finalités du traitement, en combinaison avec le règlement (UE) 2016/679, considérant 60 (Barros Vale et Zanfir-Fortuna, 2022:16).

Ex post

- En vertu de l'article 15 du RGPD, les personnes concernées ont le droit de demander l'accès à leurs données à caractère personnel, y compris aux informations sur les catégories de données à caractère personnel, les finalités du traitement effectué, les destinataires ou catégories de destinataires des données à caractère personnel, et l'existence d'une prise de décision automatisée, incluant le profilage.

ADM admissible

- En vertu de l'article 15(h) du RGPD, dans les cas de décision exclusivement automatisée et de profilage visés à l'article 22, les personnes concernées ont également le droit de demander des informations utiles concernant la logique sous-jacente, ainsi que l'importance et les conséquences prévues de ce traitement pour la personne concernée.

En tant que « bonne pratique », le Contrôleur européen de la protection des données (CEPD) encourage les personnes responsables du traitement à fournir les informations utiles concernant la logique sous-jacente et à expliquer l'importance et les conséquences prévues de l'ADM et du profilage, même pour les ADM non admissibles (Barros Vale et Zanfir-Fortuna 2022:16).

Remarque : les juridictions nationales peuvent accorder des droits supplémentaires concernant la divulgation de documents. En Suède, par exemple, une personne qui candidate à un emploi, une promotion ou une formation mais qui n'est pas sélectionnée a le droit de demander et de recevoir de l'employeur des informations écrites concernant le parcours académique, l'expérience professionnelle et les autres qualifications de la personne sélectionnée afin de pouvoir les comparer aux siennes (loi suédoise sur la discrimination, section 4).

Pour les pays non couverts par le RGPD, les dispositions de la législation nationale doivent être prises en considération.

Affaires clés : SCHUFA Holding (2023) et CK c. Dun & Bradstreet Austria GmbH et Magistradt der Stadt Wien (2025)

(les chiffres entre crochets […] se rapportent aux paragraphes de la décision)

La question de savoir si un système automatisé a été utilisé et quelles informations doivent être fournies à une personne concernée a été soulevée dans le contexte du RGPD, où des dispositions spécifiques s'appliquent en ce qui concerne « exclusivement […] un traitement automatisé, y compris le profilage » (article 22). Trois conditions doivent être remplies pour que l'article 22 soit applicable.

- Une « décision » doit avoir été prise (voir considérant 71, qui confirme que la portée est vaste et donne des exemples concrets de refus automatique d'une demande de crédit en ligne ou de pratiques de recrutement en ligne sans intervention humaine).

- La décision doit être fondée « exclusivement sur un traitement automatisé, y compris le profilage ».
- Cette décision, fondée exclusivement sur un traitement automatisé, doit produire « des effets juridiques [concernant] et [affectant] de manière significative de façon similaire » l'individu.

Lorsque ces conditions sont remplies, les organisations qui utilisent un système de prise de décision automatisée doivent être en mesure de fournir à la personne concernée une « explication utile de la manière dont la décision a été prise ». C'est très similaire au langage utilisé à l'article 86 du règlement sur l'IA, « droit à une explication ».

SCHUFA Holding

Dans l'affaire C-634/21 SCHUFA Holding (Scoring), la CJUE a examiné un litige dans lequel une société privée fournissait des évaluations de crédit à ses partenaires, comme les banques. Pour parvenir à la décision sur la solvabilité, les caractéristiques d'un individu étaient comparées aux données sur le comportement d'autres personnes présentant des caractéristiques comparables. Un score était ensuite attribué en utilisant des procédures mathématiques et statistiques pour indiquer, par exemple, dans quelle mesure une personne était susceptible de rembourser un prêt [14].

La personne ayant déposé une plainte, OQ, s'est vu refuser un prêt par une banque (un tiers) en conséquence d'informations négatives transmises à la banque par SCHUFA. OQ a demandé des informations sur les données à caractère personnel invoquées et a demandé l'effacement de certaines données prétendument erronées [15].

- SCHUFA a répondu en informant OQ du niveau de son score (une note) et a exposé, dans les grandes lignes, les modalités de calcul des scores. Cependant, elle s'est refusée, en invoquant le secret d'affaires, à divulguer les différentes informations prises en compte aux fins de ce calcul ainsi que leur pondération. Enfin, SCHUFA a indiqué qu'elle se limitait à faire parvenir des informations à ses partenaires contractuels et que c'était ces derniers qui prenaient les décisions contractuelles proprement dites. En effet, la société a soutenu que la note n'était pas une « décision » [16].

En réponse aux trois questions posées, la CJUE a statué comme suit.

- Décision : la notion de « décision » était suffisamment large pour englober le résultat du calcul de la solvabilité d'une personne sous la forme d'une valeur de probabilité concernant la capacité de cette personne à honorer ses engagements de paiement à l'avenir [47].
- « fondée exclusivement sur un traitement automatisé, y compris le profilage » : il était admis que l'activité de SCHUFA répondait à la définition du profilage et que la condition était remplie.
- « des effets juridiques concernant affectant de manière significative de façon similaire » : une valeur de probabilité insuffisante conduit, dans quasiment tous les cas, au refus de la banque d'accorder le prêt demandé. Étant donné que le tiers

fonde « de manière déterminante » sa décision sur cette valeur de probabilité, cette condition était également remplie [48].

Ainsi, la CJUE a décidé que lorsque « la valeur de probabilité établie par une société fournissant des informations commerciales et communiquée à une banque joue un rôle déterminant dans l'octroi d'un crédit, l'établissement de cette valeur doit être qualifié en soi de décision produisant à l'égard d'une personne concernée « des effets juridiques la concernant ou l'affectant de manière significative de façon similaire », au sens de l'article 22, paragraphe 1, du RGPD » [50].

La CJUE a également réitéré la nécessité :

- de prévenir les risques discriminatoires découlant du profilage ;
- d'instaurer des procédures mathématiques ou statistiques appropriées pour le profilage ;
- de mettre en place des mesures techniques et organisationnelles, notamment pour garantir que les inexactitudes dans les données à caractère personnel soient corrigées et que le risque d'erreurs soit réduit au maximum ;
- de sécuriser les données.

Fait important, la CJUE a également jugé que les États membres ne peuvent adopter des règlements autorisant le profilage au mépris des exigences énoncées aux articles 5 (principes relatifs au traitement des données à caractère personnel) et 6 (licéité du traitement) du RGPD et ne peuvent prescrire de manière définitive le résultat de la pondération des droits et des intérêts en cause [70].

Dun & Bradstreet

Les explications doivent permettre à la personne concernée de comprendre et, si nécessaire, de contester la décision et d'exercer ses droits en matière de données. La simple communication d'une formule mathématique complexe, telle qu'un algorithme ou une description détaillée de toutes les étapes d'une prise de décision automatisée ne satisfera pas à ces exigences, car elle ne constituerait pas une explication suffisamment concise et compréhensible [58-59].

Les « informations utiles concernant la logique sous-jacente » doivent décrire la procédure et les principes concrètement appliqués de telle manière que la personne concernée puisse comprendre lesquelles de ses données à caractère personnel ont été utilisées de quelle manière dans la prise de décision automatisée en cause. La complexité des opérations n'est pas considérée comme étant susceptible de libérer la personne responsable du traitement de son devoir d'explication [61].

Dans le contexte du profilage de crédit, la CJUE a jugé qu'une juridiction de renvoi peut considérer qu'informer la personne concernée de la mesure dans laquelle une variation au niveau des données à caractère personnel prises en compte aurait conduit à un résultat différent est suffisamment transparent et intelligible [62]. Ce n'est peut-être pas le cas pour tous les systèmes algorithmiques.

Lorsque la personne responsable du traitement estime que ces informations ne peuvent être fournies sans divulguer des détails sur des tiers ou des secrets commerciaux, la personne responsable du traitement des données est tenue

de fournir les informations prétendument protégées à l'autorité de surveillance compétente ou au tribunal, qui doit établir un équilibre entre les droits et les intérêts en jeu en vue de déterminer l'étendue du droit d'accès de la personne concernée prévu à l'article 15 de ce règlement.

Ces deux affaires constituent un guide de l'approche que la CJUE peut adopter lorsqu'elle examine dans quelle mesure des informations devraient être mises à disposition et sous quelle forme en vertu du RGPD et par analogie, les explications requises en vertu du règlement sur l'IA.

Entretien avec la personne affectée

Les personnes qui déposent une plainte pour discrimination fondée sur un ou plusieurs motifs protégés disposeront souvent d'informations pertinentes sur les caractéristiques qu'ils ont fournies à l'organisation intimée, récemment ou dans le passé. Recueillir d'autres informations sur leurs caractéristiques protégées et les occasions où ces caractéristiques auraient pu être divulguées peut également s'avérer utile. Soyez attentif à vos propres obligations en matière de protection des données. Pensez à poser des questions telles que les suivantes.

- Quand avez-vous eu votre premier contact et les contacts suivants avec cette organisation/ce service public ? En cas de contact antérieur avec l'organisation/le service public avant l'événement faisant l'objet de la plainte, des informations sur vos caractéristiques protégées peuvent avoir été fournies à cette occasion.
- Vous a-t-on demandé de remplir des formulaires ?
- Si oui, quels formulaires vous a-t-on demandé de remplir ?
- Veuillez examiner la liste des caractéristiques protégées. Avez-vous dû fournir des informations ou cocher des cases relatives à l'un des éléments suivants ?
- Veuillez examiner la liste des proxys connus à l'Annexe E. Vous a-t-on demandé de fournir des renseignements sur l'un de ceux-ci ?

Documents de la personne ayant déposé la plainte

La documentation disponible ou accessible à la personne ayant déposé la plainte peut contenir des informations pertinentes. Ces informations peuvent inclure :

- les documents soumis par la personne ayant déposé la plainte (formulaires de candidature, correspondance, accès aux données ou autres demandes) ;
- les documents fournis par l'organisation intimée à la personne ayant déposé la plainte (tout contrat entre l'organisation intimée et la personne ayant déposé la plainte, ainsi que leurs annexes, informations relatives au traitement des données ou avis de respect de la vie privée et communications internes, y compris les mises à jour contractuelles).

Préparer et obtenir une réponse à une demande d'accès d'une personne concernée (DAPC) conformément à l'article 15 du RGPD

Les personnes concernées ont le droit de demander l'accès à leurs données à caractère personnel à la personne responsable du traitement (généralement le déployeur du

système) en vertu de l'article 15 du RGPD. Les réponses aux DAPC correctement traitées peuvent contenir des informations très pertinentes (voir étape 3 pour plus de détails), adaptées au cas particulier de l'individu, y compris :

- les catégories de données à caractère personnel traitées et une copie numérique de celles-ci ;
- les objectifs du traitement effectué ;
- les destinataires ou les catégories de destinataires des données à caractère personnel ;
- l'existence d'une prise de décision automatisée et/ou d'un profilage, ainsi que des informations utiles sur la logique sous-jacente et les conséquences envisagées d'un tel traitement pour la personne concernée..

Si une personne ayant déposé une plainte n'a pas encore soumis de DAPC, il est recommandé de l'encourager à le faire avant de poursuivre toute plainte ou enquête. Bien qu'il n'y ait pas d'exigence formelle quant à la manière de préparer une DAPC, il est conseillé de soumettre la demande par écrit et de l'enregistrer (conserver une copie de tout courriel ou courrier papier, conserver toute preuve de courrier ou de réception, prendre une capture d'écran si vous utilisez un formulaire en ligne, un portail ou un canal de service à la clientèle). Pour un modèle de lettre, voir Annexe A.

Le non-respect des obligations du RGPD concernant les DAPC est répandu. Par conséquent, il est courant de rencontrer soit une absence de réponse, soit un retard dans l'action, soit la transmission d'informations inadéquates ou incomplètes après la présentation d'une demande. Les personnes ayant déposé une plainte exerçant leurs droits en tant que personnes concernées devraient être préparés à une éventuelle communication de suivi, incluant des rappels et des demandes de clarification supplémentaires adressées à l'organisation. Dans les affaires entraînant des enjeux importants, envisagez de collaborer avec les APD nationales concernées (voir Coopération intersectorielle).

La proposition de règlement de l'Omnibus numérique introduit des modifications au droit d'accès de la personne concernée en vertu du RGPD. La proposition de modification de l'article 12(5) du RGPD[33] a été critiquée (Skiotytė et Sadauskaitė, 2026). Si celle-ci est adoptée, elle permettrait aux personnes responsables du traitement de refuser ou de facturer des frais pour les DAPC considérées comme excessives ou poursuivant des objectifs de « hors protection des données ». Elle ne devrait pas affecter les demandes valides destinées à faire respecter les droits des personnes concernées en matière de données.

33 Considérant 35, article 3, proposition de règlement du Parlement européen et du Conseil modifiant les règlements (UE) 2016/679, (UE) 2018/1724, (UE) 2018/1725 et (UE) 2023/2854 ainsi que les directives 2002/58/CE, (UE) 2022/2555 et (UE) 2022/2557 en ce qui concerne la simplification du cadre législatif numérique, et abrogeant les règlements (UE) 2018/1807, (UE) 2019/1150 et (UE) 2022/868 ainsi que la directive (UE) 2019/1024 (règlement omnibus numérique). Disponible en suivant ce lien : eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:52025PC0837.

2.1.2. Informations publiques

Communications externes de l'organisation intimée

Des informations pertinentes peuvent être communiquées par l'organisation intimée sur son site Web ou sa plateforme, y compris dans les pages FAQ, les conditions générales ou les avis de respect de la vie privée, conformément aux obligations de transparence du RGPD applicables (voir l'encadré sur les obligations de transparence du RGPD et les droits en matière de données ci-dessus).

Prenez note en particulier des références à l'objectif du système, aux finalités du traitement des données, aux types de données traitées, y compris toute référence à des données sensibles (également appelées « catégories particulières de données ») et toute référence à l'existence d'une ADM (admissible) ou de profilage, sa logique et ses conséquences. Celles-ci peuvent indiquer l'utilisation d'algorithmes ou de caractéristiques protégées (pour plus de détails, voir les étapes 3.2 et 3.3).

Dans certains pays, comme la Finlande, les administrations publiques sont soumises à des obligations d'information et doivent publier des informations concernant le recours à des systèmes d'ADM sur leur site Web et informer les personnes concernées de cette utilisation lors de la prise de décisions. Cependant, il n'existe généralement pas de registre national ou transnational systématique, malgré des initiatives locales visant à créer des bases de données d'applications d'IA et d'ADM (comme en Belgique et en Finlande). Les OPE font donc face à des difficultés dans l'exercice de leur mandat d'enquête, de supervision, d'identification et de contestation des affaires de discrimination algorithmique et de soutien aux victimes (Conseil de l'Europe, 2025a).

Jurisprudence existante

Recherchez la jurisprudence couvrant soit les systèmes similaires d'IA ou algorithmiques, soit les processus de prise de décision automatisée. Elle peut inclure la jurisprudence nationale, la jurisprudence européenne (y compris de la Cour européenne des droits de l'homme et de la CJUE) et la jurisprudence nationale étrangère, étant donné que les systèmes et plateformes algorithmiques fonctionnent souvent selon des logiques et des schémas récurrents similaires dans différents pays.

Étant donné que la jurisprudence de l'UE sur la réglementation en matière d'IA et d'algorithmes en est encore à ses balbutiements, envisagez également d'examiner la jurisprudence liée au RGPD, en particulier les décisions portant sur la prise de décision automatisée, étant donné que celles-ci peuvent être utiles. Envisagez de consulter ce qui suit :

- Des bases de données telles que Jurisprudence de l'Union – EUR-Lex ou la Base de données jurisprudentielle de l'Agence des droits fondamentaux de l'Union européenne pour les affaires de la CJUE ou les affaires judiciaires, en utilisant des mots-clés comme « intelligence artificielle », « décision automatisée », « profilage » et « notation de crédit » pour examiner des jugements existants ou suivre les demandes en cours des juridictions nationales adressées à la CJUE

pour des décisions préjudicielles relatives aux aspects pertinents du règlement sur l'IA ou du RGPD.

- Pour les bases de données adaptées à l'IA ou aux systèmes algorithmiques, tenez compte également de la Base de données jurisprudentielle – Tech Litigation (disponible en anglais), qui recueille les efforts de justice contre les systèmes automatisés, y compris les jugements, décisions et avis des tribunaux nationaux et internationaux ainsi que des autorités de protection des données. En outre, la DAIL – the Database of AI Litigation – Ethical Tech Initiative (disponible en anglais), qui compile des données sur les litiges aux États-Unis impliquant l'IA et les systèmes algorithmiques, peut fournir des informations précieuses sur les tendances émergentes en matière de litiges liés à la discrimination algorithmique, aux défauts systémiques et aux risques de discrimination, liés à des types spécifiques de technologies, y compris des informations pertinentes sur les entreprises basées aux États-Unis opérant au sein de l'Union européenne, soit en tant que fournisseurs de systèmes d'IA similaires, soit en tant qu'organismes responsables du traitement des données (comme Clearview AI).
- Certaines OSC commencent à compiler des bases de données spécifiquement axées sur les cas de discrimination algorithmique (voir le formulaire de signalement Algorithm Watch – disponible en anglais). Ces bases de données seront plus utiles si les OPE y contribuent en communiquant leurs propres résultats en matière de litiges ou d'enquêtes. En outre, elles doivent respecter les principes de protection des données et de privilège relatif aux litiges.
- L'outil de Suivi de l'application du RGPD listant les amendes (GDPR Enforcement Tracker – list of GDPR fines – disponible en anglais) et la page Amendes du Comité européen de la protection des données donnent une vue d'ensemble des amendes et des sanctions infligées par les autorités de protection des données au sein de l'UE en vertu du RGPD, en incluant le pays, l'autorité, la date, le montant de l'amende, la personne responsable/chargée du traitement, le secteur et le type d'amende.
- Le rapport du Future of Privacy Forum sur la prise de décision automatisée dans le cadre du RGPD (Automated decision making under the GDPR – disponible en anglais) présente un ensemble de jurisprudences sur l'ADM dans le cadre du RGPD depuis avril 2022.
- Bases de données nationales de jurisprudence.

Autres sources

Il existe un large éventail de sources de données pouvant s'avérer utiles pour identifier, enquêter ou prouver un cas de discrimination algorithmique.

- Les rapports nationaux du 7e cycle de suivi de la Commission européenne contre le racisme et l'intolérance (ECRI) tiennent compte de l'IA.
- Des rapports d'OSC portant soit sur des systèmes d'IA ou algorithmiques similaires, soit sur des processus décisionnels similaires.
- Les constats des organes de suivi des Nations Unies, notamment le Comité pour l'élimination de la discrimination à l'égard des femmes (CEDEF ou CEDAW), le

Comité pour l'élimination de la discrimination raciale (CERD), le Comité des droits de l'enfant (CDE), le Comité contre la torture (CCT), le Comité des droits économiques, sociaux et culturels (CDESC), et les rapporteurs spéciaux, ainsi que les constats d'autres agences onusiennes telles que l'Organisation mondiale de la santé (OMS).

- Les constats d'autres organismes internationaux, comme le Bureau des institutions démocratiques et des droits de l'homme de l'Organisation pour la sécurité et la coopération en Europe (OSCE/BIDDH-ODIHR) et l'Organisation de coopération et de développement économiques (OCDE). À cet égard, le registre de l'OCDE des incidents et dangers liés à l'IA (OECD AI Incidents Monitor, an evidence base for trustworthy AI – OECD.AI – disponible en anglais) peut s'avérer particulièrement pertinent.
- Les constats des organes nationaux spécialisés, y compris les institutions de médiation, les organismes de promotion de l'égalité et les OSC.
- Les registres administratifs, comme les registres centraux et locaux de la population.
- Les données relatives aux plaintes, comme les statistiques émanant de la police, des procureur·es et des dossiers et registres des affaires judiciaires.
- Les recherches académiques et ad hoc, y compris les constats d'agences européennes telles que l'Agence des droits fondamentaux de l'Union européenne (FRA) et l'Institut européen pour l'égalité entre les hommes et les femmes (EIGE). Une recherche par domaine dans le Répertoire des risques liés à l'IA du MIT (MIT AI Risk Repository – disponible en anglais) peut également s'avérer utile.
- Des informations comparatives sur la jurisprudence, les politiques et les pratiques dans les États membres (données du « consensus européen »).

2.1.3. Informations fournies par l'organisation intimée ou d'autres parties

Informations demandées à l'organisation intimée en faisant usage des compétences nationales ou d'autres compétences du droit de l'Union

Les organismes de promotion de l'égalité pourraient demander des informations aux organisations intimées et à d'autres parties, par exemple le déployeur ou le fournisseur du système (s'il s'agit d'une organisation différente de l'organisation intimée), en exerçant leurs pouvoirs d'enquête et de collecte d'informations en vertu de la législation pertinente, incluant les lois nationales transposant les directives relatives aux normes[34], la législation sur l'accès à l'information et les instruments

34 En vertu de l'article 16 des directives relatives aux normes, les organismes de promotion de l'égalité disposent de divers pouvoirs en matière de collecte de données relatives à l'égalité, y compris la collecte de données en vue de l'établissement de rapports visant à procéder à des études indépendantes (article 16(2)) et le droit d'accéder aux statistiques lorsque cela est nécessaire pour procéder à une évaluation globale de la situation en matière de discrimination dans l'État membre et pour établir des rapports (article 16(3)).

internationaux tels que la Convention-cadre du Conseil de l'Europe sur l'IA[35], ou en utilisant des canaux de collaboration avec les autorités nationales de protection des données.

Les informations peuvent être divulguées au cours des enquêtes, des négociations et de la correspondance préalable à l'action ou dans le cadre d'un litige.

Enquêtes : directives relatives aux normes (article 8)

L'article 8 des directives 2024/1499 et 2024/1500 pourrait renforcer les droits d'accès à l'information des organismes de promotion de l'égalité, notamment :

- en vertu de l'article 8(1) des directives relatives aux normes, les États membres doivent habiliter les organismes de promotion de l'égalité à mener une enquête pour déterminer s'il y a eu violation du principe de l'égalité de traitement ;
- en vertu de l'article 8(2), des directives relatives aux normes, les États membres doivent prévoir un cadre pour la réalisation d'enquêtes permettant aux organismes pour l'égalité d'établir les faits, et leur conférant :
  - des droits effectifs d'accès aux informations et aux documents nécessaires pour établir l'existence d'une discrimination ;
  - des mécanismes appropriés permettant aux organismes de promotion de l'égalité de coopérer avec les organismes publics compétents à cette fin.

Il est important de noter que l'article 8(3) prévoit que les États membres peuvent confier ces pouvoirs à un autre organisme compétent. Il pourrait s'agir, par exemple, de une ASM. Dans ce cas, il s'agirait de l'ASM exerçant les pouvoirs prévus à l'article 8(1) et (2). Lorsqu'un tel organisme compétent a terminé son enquête, il fournit à l'organisme de promotion de l'égalité, à sa demande, des informations sur les résultats de l'enquête.

Si ces pouvoirs ont été confiés à l'ASM, il est important de noter que, dans le règlement sur l'IA, cela nécessite une coopération avec l'OPE (article 79(2)).

- Les enquêtes pourraient être particulièrement pertinentes dans un cas mettant en cause des systèmes n'étant pas classés à haut risque lorsqu'il s'est avéré difficile d'obtenir des preuves documentaires ou lorsqu'il existe une suspicion de discrimination de groupe mais que les preuves sont insuffisantes dans les cas individuels.

35 La Convention-cadre du Conseil de l'Europe sur l'IA (2024) prévoit que les Parties adoptent ou maintiennent des mesures pour garantir des exigences adéquates en matière de transparence et de contrôle (article 8), et pour garantir l'obligation de rendre des comptes et d'assumer la responsabilité concernant les impacts négatifs sur les droits humains (article 9). L'article 14, notamment, exige des Parties qu'elles garantissent des voies de recours accessibles et effectives, y compris à travers la documentation et la transmission d'informations sur les systèmes d'IA ayant une incidence significative sur les droits humains aux organismes autorisés et, le cas échéant, aux personnes concernées, afin que celles-ci puissent contester la ou les décisions prises par le biais de l'utilisation de l'IA et former un recours auprès des autorités compétentes.

Remarque : aucune enquête ne peut être ouverte ou poursuivie tant qu'une procédure judiciaire est en cours dans la même affaire. Les enquêtes sont susceptibles d'être régies par des dispositions de droit national.

Analyse d'impact sur la protection des données (AIPD)

Tout traitement de données à caractère personnel susceptible d'entraîner un risque élevé pour les droits des personnes concernées en raison de sa nature, de sa portée, de son contexte ou de ses finalités nécessite que la personne responsable du traitement effectue une AIPD conformément à l'article 35(1), du RGPD. En conséquence, les AIPD présentent donc un intérêt particulier dans les affaires impliquant une prise de décision automatisée et un profilage par le biais de systèmes algorithmiques (article 35(3) du RGPD). Toutefois, à l'exception d'un nombre limité de déploiements de systèmes d'IA à haut risque ayant des obligations de publication en vertu du règlement sur l'IA (voir aussi AIPD à l'étape 2.2.3 ci-dessous), il n'existe aucune obligation légale pour le déployeur, en vertu du RGPD ou du règlement sur l'IA, de publier ou de fournir l'AIPD à un organisme de promotion de l'égalité. Tout contrôle concernant l'analyse d'impact est laissé à l'initiative des APD – les canaux de collaboration avec les APD pourraient s'avérer pertinents à cet égard. Cela ne devrait toutefois pas dissuader l'organisme de promotion de l'égalité de formuler une demande de ce type. Certaines organisations peuvent accepter de fournir l'AIPD si celle-ci est disponible (soit en vertu de l'article 35(3) du RGPD, soit volontairement) ou peuvent même choisir d'en publier une synthèse. L'AIPD peut être demandée lors de la correspondance préalable à l'action ou si un litige est engagé.

Analyses d'impact sur les droits humains ou les droits fondamentaux (AIDH/AIDF)

Le règlement sur l'IA prévoit l'obligation explicite de réaliser tout type d'AIDF uniquement pour certains déployeurs et fournisseurs de systèmes d'IA à haut risque (voir également l'analyse d'impact sur les droits fondamentaux à l'étape 2.2.3 ci-dessous). Pour les autres systèmes algorithmiques ou d'IA, aucune obligation de ce genre n'existe en vertu du règlement sur l'IA. Toutefois, les fournisseurs peuvent toujours adopter volontairement ce type d'instruments (European Center for Not-for-profit Law et Danish Institute for Human Rights 2025 ; FRA 2025), en raison de leur obligation générale de respecter et de protéger les droits fondamentaux établis aux niveaux de l'UE et des États membres, et en tant que moyen de renforcer la responsabilité des opérateurs d'IA à cet égard (Mantelero 2024).

En particulier, une AIDH est un terme général désignant le processus et l'outil d'identification, d'évaluation et de traitement des effets négatifs potentiels et réels de produits ou d'activités sur les droits humains, et peut être mise en œuvre sous diverses formes. La méthodologie et le modèle HUDERIA développés par le Conseil de l'Europe constituent un type d'évaluation des risques pour les droits humains (Conseil de l'Europe 2025c, 2026).

De même, une AIDH ou une AIDF peut avoir été réalisée par des développeurs sur une base volontaire pour des systèmes d'IA hors haut risque, y compris dans le cadre

de leurs propres engagements en matière de diligence raisonnable ou de droits humains ou conformément à des cadres tels que les Principes directeurs relatifs aux entreprises et aux droits de l'homme des Nations Unies. Les lois, politiques ou pratiques nationales ou régionales peuvent également exiger la réalisation et la publication d'études d'impact sur les droits humains ou les droits fondamentaux.

Dans tous ces cas, les organismes de promotion de l'égalité peuvent demander la divulgation de ces évaluations dans le cadre de leurs pouvoirs de collecte d'informations.

Égalité ou évaluation de l'impact sur le genre

Lorsque des évaluations de ce type sont requises en vertu du droit national, les organismes de promotion de l'égalité peuvent en faire la demande. Il peut n'y avoir aucune obligation légale de les réaliser ou de les fournir.

Documentation dans le cadre du règlement sur l'IA

Lorsqu'un système n'est pas classé comme étant à haut risque, les organismes de promotion de l'égalité peuvent toujours demander des informations comparables au déployeur ou au fournisseur, en s'appuyant sur leurs pouvoirs d'enquête ou de collecte d'informations. Bien que les organisations intimées ne soient pas légalement tenues d'avoir rédigé et/ou fourni cette documentation pour les systèmes qui ne sont pas à haut risque ou qui ne constituent pas des systèmes d'IA, les organismes de promotion de l'égalité peuvent être guidés par le cadre du règlement sur l'IA pour structurer leurs demandes et évaluer la réaction (ou l'inaction) de l'organisation intimée en conséquence, aux étapes 3 et 4.

Enquêtes et informations statistiques

Les organismes de promotion de l'égalité peuvent mener des enquêtes ou y avoir accès, y compris des enquêtes par sondage qui recensent les expériences des groupes exposés à la discrimination et des enquêtes d'opinion qui mesurent la prévalence des préjugés (directives relatives aux normes, article 16). Ils peuvent également avoir accès à des sources statistiques officielles pertinentes, comme celles compilées par les bureaux nationaux de statistiques dans le cadre des directives européennes en matière de non-discrimination, ainsi qu'à des données recueillies dans le cadre de recensements de la population et d'autres sources importantes, comme l'enquête sur les forces de travail (EFT) et l'enquête Statistiques sur les ressources et conditions de vie (SRCV).

Directives relatives aux normes

- En vertu de l'article 16(2) des directives 2024/1499 et 2024/1500, les organismes de promotion de l'égalité peuvent procéder à des études indépendantes concernant la discrimination.
- En vertu de l'article 16(3), des directives 2024/1499 et 2024/1500, les organismes de promotion de l'égalité doivent pouvoir accéder aux statistiques relatives aux droits et obligations découlant des directives européennes relatives

à la non-discrimination, lorsque qu'il est estimé que ces statistiques sont nécessaires pour procéder à une évaluation globale de la situation en matière de discrimination dans l'État membre et pour établir des rapports périodiques sur l'état de l'égalité dans leur État membre.

Documentation en vertu d'autres lois sur la sécurité des produits

Pour les produits, qui ne sont pas classés à haut risque, vendus par une entreprise auprès de consommateurs et consommatrices, des exigences peuvent être imposées en vertu du Règlement sur la sécurité générale des produits[36] (RSGP) et du Règlement sur l'ASM[37], qui comprennent des obligations et des exigences en matière de documentation comparables à celles du règlement sur l'IA. Ces exigences ne relèvent pas du champ d'application de la présente méthodologie car les pouvoirs appartiennent à l'ASM ; toutefois, elles sont mentionnées car le RSGP est considéré comme un filet de sécurité pour les systèmes d'IA qui ne sont pas à haut risque (considérant 166 du règlement sur l'IA) et ses exigences peuvent conduire à la mise à disposition d'une documentation pertinente, soit directement à l'individu, soit par l'organisme de promotion de l'égalité agissant en coopération avec l'ASM. Par exemple, les fabricants soumis à ce règlement sont tenus de veiller à ce que leur produit soit accompagné d'instructions claires et de consignes de sécurité à l'intention des consommateurs et consommatrices (article 9(7)), d'établir une documentation technique et de procéder à des évaluations des risques qui doivent être mises à la disposition de l'ASM (articles 9(2)-(3) et 15).

2.2. Informations supplémentaires disponibles pour les systèmes d'IA à haut risque

2.2.1. Renseignements supplémentaires à la disposition de la personne ayant déposé la plainte

Préparer et obtenir une réponse à une demande d'explication conformément à l'article 86 du règlement sur l'IA

Ce droit peut être invoqué en complément de la demande DAPC (voir Préparer et obtenir une réponse à une demande d'accès d'une personne concernée (DAPC) conformément à l'article 15 du RGPD). Toute personne faisant l'objet d'une décision qui l'affecte de manière significative et négative, prise sur le fondement de résultats d'un système d'IA à haut risque, a le droit d'obtenir des informations sur le rôle du système d'IA dans le processus décisionnel et les principaux éléments de la décision. Ces explications devraient être claires et utiles. Les réponses correctement traitées peuvent contenir des informations très pertinentes, y compris les données d'entrée et les critères d'influence de la décision (voir l'étape 3 pour plus de détails) aux fins de l'identification ultérieure d'une discrimination prima facie. Si une personne ayant déposé une plainte n'a pas encore présenté ce type de demande, elle devrait être

36 Règlement (UE) 2023/988.
37 Règlement (UE) 2019/1020.

encouragée à le faire avant de poursuivre toute plainte ou enquête. Pour obtenir un modèle, voir l'annexe B.

Ce droit ne s'applique que dans la mesure où il n'est pas déjà prévu par le droit de l'Union. Par exemple, si le système algorithmique est un système entièrement automatisé qui traite des données à caractère personnel de la personne concernée, il se peut que le droit à une explication soit prévu par le RGPD aux articles 13-15 (règlement (UE) 2016/679). Lorsque l'organisation intimée rejette la demande d'explication pour ce motif, une demande de suivi au titre du RGPD (autre droit de l'Union applicable) serait l'étape suivante (sauf si elle a déjà été exercée). Pour obtenir un modèle, voir l'annexe A.

2.2.2. Informations publiques

Base de données de l'UE

Voir également 1.2.1 Base de données de l'UE .

À l'étape 1, vous avez effectué une vérification préliminaire de la base de données de l'UE. Si l'organisation intimée figure sur la liste, vous devriez avoir déjà repéré si elle est reprise dans la base de données comme suit :

- un fournisseur, mentionné dans la section A ou B de la base de données :
  - pour un cas d'utilisation à haut risque pertinent dans la situation de la personne ayant déposé la plainte ;
  - pour un cas d'utilisation destiné à un domaine à haut risque, mais relevant de l'exception prévue à l'article 6(3), du règlement sur l'IA ;
- un déployeur, mentionné dans la partie C de la base de données :
  - un déployeur pour un cas d'utilisation pertinent dans la situation de la plainte ; quelle organisation est le fournisseur ?

Vous trouverez ci-après les principaux points d'intérêt de la base de données (règlement sur l'IA, annexe VIII).

Section A : systèmes d'IA à haut risque

- Coordonnées du fournisseur et de la personne représentante autorisée ;
- Dénomination commerciale ou autre référence permettant la traçabilité de l'identification du système d'IA (cette information pourrait être utile, par exemple, lorsqu'un système à haut risque est utilisé dans le cadre d'un autre système) ;
- Usage prévu ;
- Informations utilisées par le système (données, entrées) et logique de fonctionnement (ne s'applique pas aux systèmes de mise en application de la loi, de migration, d'asile et de contrôle aux frontières) ;
- Statut du système d'IA (par exemple, est-il toujours sur le marché ?) ;
- Type, numéro et date d'expiration du certificat délivré par l'organisme notifié, nom ou numéro d'identification de cet organisme et copie scannée du certificat (non applicable aux systèmes de mise en application de la loi, de migration, d'asile et de contrôle aux frontières) ;

- États membres dans lesquels le système est ou a été commercialisé (utile pour la coopération avec d'autres autorités) ;
- Une copie de la déclaration de conformité de l'UE (règlement sur l'IA, article 47) ;
- Instructions d'utilisation au format électronique (ne s'applique pas aux systèmes de mis en application de la loi ou de migration, d'asile et de contrôle aux frontières).

Section B : système d'IA pour lequel le fournisseur a conclu qu'il ne présentait pas de haut risque

- Coordonnées du fournisseur et de la personne représentante autorisée ;
- Dénomination commerciale ou autre référence permettant l'identification et le traçage du système d'IA ;
- Usage prévu ;
- Article 6(3), condition et justification de la raison pour laquelle le système n'est pas considéré comme étant à haut risque (non applicable aux services de mise en application de la loi, de migration, d'asile ou de contrôle aux frontières) ;
- Le statut du système d'IA (est-il toujours sur le marché ?) ;
- États membres où le système est ou a été commercialisé.

Section C : déployeurs de systèmes à haut risque (règlement sur l'IA, article 49(3))

- Nom et coordonnées du déployeur (uniquement les autorités publiques, les institutions, organes et agences de l'UE – ou les personnes agissant en leur nom) ;
- URL de l'entrée du système d'IA dans la base de données de l'UE (telle qu'enregistrée par le fournisseur) ;
- Synthèse de l'analyse d'impact sur les droits fondamentaux (AIDF) (ne s'applique pas aux services de mise en application de la loi, de migration, d'asile ou de contrôle aux frontières) ;
- Synthèse de l'analyse d'impact sur la protection des données (AIPD).

Tous les systèmes d'IA à haut risque ne sont pas répertoriés dans la section publique de la base de données de l'UE.

Le tableau ci-dessous indique dans quels cas d'utilisation les informations doivent être fournies dans la base de données.

Tableau 3. Informations disponibles dans la base de données de l'UE et nationale sur les systèmes à haut risque

Section publique (article 49(1)-(3)), (à l'exception des déployeurs privés, sauf s'ils agissent pour le compte d'un organisme public)	Section non publique (article 49(4))	Mise en œuvre/ enregistrement au niveau national (article 49(5))
Toutes les informations décrites ci-dessus sont requises et sont incluses dans la section publique de la base de données.	Une quantité limitée d'informations est requise et n'est incluse que dans une section sécurisée non publique de la base de données.	Les exigences d'enregistrement ne s'appliquent qu'au niveau national.
Éducation et formation professionnelle	Mise en application de la loi	Infrastructure critique
Emploi, gestion de la main d'œuvre et accès à l'emploi indépendant	Migration, asile et contrôles aux frontières	
Services et allocations essentielles	Systèmes d'IA à haut risque en cours de test	
Administration de la justice et processus démocratiques	Données biométriques utilisées à des fins de mise en application de la loi, de migration, d'asile et de contrôle aux frontières	

Restrictions d'accès à l'information

Conformément à l'article 49(4)(d) du règlement sur l'IA, pour les systèmes d'IA à haut risque dans les domaines de la mise en application de la loi, de la migration, de l'asile et du contrôle aux frontières (annexe III, points 1, 6 et 7), les informations relatives aux systèmez d'IA enregistrées dans la base de données de l'UE sont inscrites dans une section non publique et accessibles uniquement à la Commission européenne et aux ASM compétentes. À cette fin, l'autorité de protection des données ou le Contrôleur européen de la protection des données (CEPD, pour la Commission) sera l'autorité de surveillance compétente (article 74(8) et (9)) et aura accès aux informations contenues dans la base de données.

Cela ne signifie pas que les documents sur lesquels reposent ces enregistrements sont également inaccessibles. Toutefois, cela signifie qu'une justification plus solide est nécessaire pour obtenir l'accès aux documents et que cet accès pourrait être refusé. Les OPE qui sont des autorités relevant de l'article 77 ont le pouvoir de

demander et d'accéder à toute information ou documentation créée ou conservée en vertu du règlement (voir ci-dessous, étape 2.2.3). Toute demande de ce type doit être suffisamment justifiée pour démontrer qu'elle est nécessaire pour l'exercice effectif de leur mandat, dans la limite de leur compétence.

Le règlement précise également très clairement que les activités de surveillance du marché « ne portent en aucune manière atteinte à l'indépendance des autorités judiciaires ni n'interfèrent d'une autre manière avec leurs activités lorsque ces autorités agissent dans l'exercice de leurs fonctions judiciaires » (article 74(8)). Cela signifie qu'en ce qui concerne toute information pertinente qui devrait être mise à la disposition de l'OPE ou d'une personne par le biais de procédures judiciaires, les restrictions de l'article 74(8) ne s'appliquent pas. La disponibilité de contenu sensible, même par le biais de ces procédures, est susceptible d'être régie par d'autres exigences de confidentialité et par des intérêts de sécurité publique et nationale (voir Confidentialité).

Si des bases juridiques pour le partage d'informations entre les organismes de promotion de l'égalité, les autorités de surveillance du marché et d'autres autorités relevant de l'article 77 sont établies (voir Coopération intersectorielle), les organismes de promotion de l'égalité pourraient accéder aux informations obtenues par les autorités de surveillance du marché dans le cadre des articles 71(4) et 49(4) (Conseil de l'Europe 2025b).

2.2.3. Informations supplémentaires fournis par l'organisation intimée ou d'autres parties

Documentation du règlement sur IA et obligations de rapports

Voir également Règlement européen sur l'IA, article 77 : pouvoirs des autorités chargées de la protection des droits fondamentaux.

Le règlement sur l'IA impose aux fournisseurs et aux déployeurs de systèmes d'IA à haut risque de créer et de conserver divers types de documentation tout au long du cycle de vie du système. Les organismes de promotion de l'égalité qui sont des autorités relevant de l'article 77 ont le pouvoir de demander et d'accéder à tout document créé ou conservé en vertu du règlement sur l'IA dans une langue et un format accessibles, nécessaire pour s'acquitter efficacement de leurs mandats dans les limites de leur compétence[38] (Règlement européen sur l'IA, article 77).

Cette documentation est reprise dans le tableau des différents types de documentation technique ci-dessous.

38 Les organismes de promotion de l'égalité qui ne relèvent pas de l'article 77 ne disposent pas d'une base juridique directe supplémentaire en vertu du règlement sur l'IA pour accéder à la documentation pertinente pour les systèmes d'IA à haut risque et devront s'appuyer sur leurs pouvoirs existants. On peut toutefois se demander s'ils seront en mesure d'acquérir une compréhension suffisante du fonctionnement des systèmes d'IA à haut risque pour s'acquitter efficacement de leurs missions.

Il est recommandé de demander toutes ces informations car l'ensemble de ces documents fournit une image claire de l'utilisation, de la conception, des données et des risques du système.

L'étape 3 fournit des explications détaillées sur ce que ces informations montreront et sur la manière dont vous pouvez les utiliser.

La documentation sera probablement organisée sur la base de normes de documentation universelles et sectorielles élaborées par l'industrie ces dernières années, incluant des références à des « fiches techniques » et à des « fiches modèles » en annexe IV. Pour les organismes de promotion de l'égalité, il serait utile de connaître ces deux types de documentation (Mittelstadt 2024). La structure et le contenu requis pour ces types de documentation seront liés par les normes du Comité technique mixte 21 (JTC 21) (voir ci-dessous, Normes harmonisées).

Voir aussi Coopération intersectorielle.

Voici une liste des exigences qui doivent être respectées pour les systèmes d'IA à haut risque en vertu du règlement sur l'IA. Bien que toutes les exigences ne requièrent pas explicitement de la documentation, il est prudent de supposer qu'au moins une documentation interne existera pour chacune de ces exigences.

Tableau 4. Exigences applicables aux systèmes d'IA à haut risque en vertu du règlement sur l'IA

Exigence (article)	Description de l'obligation	Rôle
Système de gestion des risques (article 9)	Système de gestion des risques qui identifie et analyse les risques raisonnablement prévisibles pour la santé, la sécurité ou les droits fondamentaux, en fonction des mesures de gestion des risques.	Fournisseur
Données et gouvernance des données (article 10)	Les jeux de données d'entraînement, de validation et de test sont soumis à la gouvernance des données. Les pratiques comprennent les choix de conception, la collecte et l'origine des données, la préparation pertinente, les hypothèses formulées, l'examen des biais et les mesures appropriées pour prévenir ces derniers.	Fournisseur
Documentation technique pour les systèmes d'IA à haut risque (article 11, annexe IV)	Description générale du système d'IA et description détaillée de certains éléments et processus. Une description plus détaillée est proposée dans le tableau ci-dessous.	Fournisseur

Exigence (article)	Description de l'obligation	Rôle
Enregistrement (article 12)	Les systèmes devraient enregistrer les événements pertinents pour identifier un risque pour la santé, la sécurité ou les droits fondamentaux et faciliter le suivi après la mise sur le marché. Ces événements comprennent l'enregistrement de la période d'utilisation, la base de données de référence, les données d'entrée et l'identification des personnes physiques impliquées dans la vérification des résultats.	Fournisseur
Transparence et fourniture d'informations aux déployeurs (article 13) et contrôle humain (article 14)	Le fonctionnement des systèmes devrait être suffisamment transparent. Les informations comprennent les caractéristiques, les capacités et les limites du système, ainsi que son objectif prévu et les mesures de performance. En outre, le contrôle humain et les ressources nécessaires au fonctionnement du système devraient être inclus.	Du fournisseur au déployeur
Système de gestion de la qualité (article 17)	Documentation des politiques, procédures et instructions, y compris la stratégie en matière de conformité, les techniques de conception et de contrôle de la qualité, l'examen avant, pendant et après le développement. Comprend également des spécifications techniques et des systèmes pour la gestion des données. Enfin, la documentation concernant la réponse aux incidents graves, la communication avec les autorités, la tenue de dossiers, la gestion des ressources et la division des responsabilités.	Fournisseur
Journaux générés automatiquement (article 19)	Les fournisseurs doivent conserver les journaux générés automatiquement pendant au moins six mois. Les institutions financières tiennent les registres conformément à la loi sur les services financiers pertinente.	Fournisseur

Exigence (article)	Description de l'obligation	Rôle
Obligations incombant aux déployeurs de systèmes d'IA à haut risque (article 26)	Les déployeurs doivent conserver les journaux générés automatiquement pendant au moins six mois. Les institutions financières tiennent les registres conformément à la législation sur les services financiers pertinente.	Déployeur
Analyse d'impact sur les droits fondamentaux (article 27)	Bien que des exceptions s'appliquent, les déployeurs doivent procéder à une évaluation de l'impact que l'utilisation du système peut avoir sur les droits fondamentaux. Cette analyse comprend les processus conformes à l'objectif visé, la description du calendrier et de la fréquence, les catégories de personnes touchées, les risques spécifiques de préjudice, les mesures de contrôle humain et les mesures prises en cas de matérialisation du risque.	Déployeur
Déclaration UE de conformité (article 47, annexe V)	Le fournisseur doit établir une déclaration de conformité de l'UE (lisible par machine) écrite et signée pour chaque système d'IA à haut risque, à la disposition des autorités nationales pendant 10 ans après sa mise sur le marché, indiquant qu'il satisfait aux exigences énoncées à l'article 2 du règlement.	Fournisseur
Marquage CE (article 48)	Les systèmes sont soumis aux principes généraux énoncés à l'article 30 du règlement (CE) n° 765/2008 et portent un marquage CE.	Fournisseur
Surveillance après commercialisation (article 72)	Les fournisseurs doivent mettre en place un système de surveillance après l'introduction sur le marché, qui collecte, documente et analyse les données pertinentes sur les performances du système, afin de permettre une évaluation de la conformité.	Fournisseur
Signalement d'incidents graves (article 73)	Les fournisseurs sont tenus de signaler les incidents graves aux autorités de surveillance du marché de l'État membre concerné.	Du fournisseur à l'ASM. L'ASM doit impliquer les OPE.

La documentation technique exigée des fournisseurs en vertu de l'article 11 du règlement sur l'IA, en combinaison avec l'annexe IV du règlement sur l'IA, constituera une ressource particulièrement utile pour les OPE.

Le tableau suivant met en lumière cette documentation technique et indique aux organismes de promotion de l'égalité les éléments requis qui peuvent être pertinents aux fins de la mise en œuvre de cette méthodologie.

Tableau 5. Différents types de documentation technique visés à l'article 11 et à l'annexe IV (adapté avec l'autorisation de Mittelstadt 2024)

Type de documentation	Explication
1. Description générale du système d'IA	Ce cadre inclut : ▸ usage prévu ; ▸ interactions avec le système ; ▸ gestion des versions et réflexion ; ▸ apparence dans les logiciels et le matériel, ou dans les produits dont il fait partie ; ▸ matériel sur lequel il est destiné à être utilisé ; ▸ interface utilisateur et instructions pour le déployeur.
2. Spécifications de conception	Ces spécifications incluent : ▸ procédés et étapes pour le développement du système ; ▸ spécifications de conception ; logique, algorithme et choix clés, y compris hypothèses et choix d'optimisation ; ▸ pertinence des paramètres choisis ; ▸ rendement attendu et sa qualité ; ▸ compromis possibles en matière de conformité.
3. Architecture du système	Explication de la structure des composants logiciels, de leur interaction et de leur intégration aux ressources informatiques utilisées pour le développement, l'entraînement, les tests et la validation.
4. Exigences en matière de données	Les exigences en matière de données comprennent des fiches techniques décrivant les méthodes d'entraînement, les techniques et les jeux de données utilisés avec des informations sur leur origine, leur portée et leurs caractéristiques, les procédures d'obtention, d'étiquetage et les méthodes de nettoyage.
5. Contrôle humain	Évaluation des mesures de contrôle humain mises en œuvre conformément à l'article 14, avec les mesures techniques nécessaires pour faciliter l'interprétation du système d'IA par les déployeurs.

Type de documentation	Explication
6. Mesures des performances	Les procédures de validation et de test utilisées, ainsi que les informations sur les principales caractéristiques des données de test, les paramètres utilisés pour mesurer l'exactitude, la robustesse et la conformité, ainsi que les effets potentiellement discriminatoires. Contient également tous les journaux et les rapports de test, y compris les modifications prédéterminées.
7. Capacités	Des informations détaillées sur la surveillance, le fonctionnement et le contrôle du système d'IA sont nécessaires, en particulier concernant ses capacités et ses limites. Elles incluent les degrés de précision pour des groupes spécifiques.
8. Gestion des risques	Le système de gestion des risques décrit à l'article 9.
9. Surveillance après l'introduction sur le marché	Une description détaillée du système mis en place pour évaluer les performances du système après l'introduction sur le marché conformément à l'article 72.
10. Autres législations de l'UE	Normes harmonisées appliquées, copie de la déclaration européenne de conformité (article 47).

Normes harmonisées

Les normes harmonisées sont des normes techniques en cours d'élaboration par les organismes européens de normalisation (CEN/CENELEC) qui, si elles sont respectées, conféreront aux fournisseurs de systèmes d'IA à haut risque une « présomption de conformité » au règlement sur l'IA. Dans la pratique, ces normes couvriront probablement un large éventail d'éléments d'information, y compris ceux nécessaires pour aider à comprendre le mode de fonctionnement du système d'IA, ses caractéristiques, ses capacités, ses forces, ses limites et ses performances, conformément au texte juridique (Soler Garrido et al. 2024)

Les normes ci-après, en cours d'élaboration, ont été jugées particulièrement importantes pour définir les exigences en matière de documentation technique à la disposition des organismes de promotion de l'égalité[39]

- **JT021036** « Intelligence artificielle – Concepts, mesures et exigences pour la gestion des biais dans les systèmes d'IA » : l'objectif de cette norme est de définir « les concepts, les mesures et les exigences permettant l'évaluation et

39. Pour une liste et une description de toutes les normes en cours d'élaboration, voir CEN/CENELEC, « CEN/CLC/JTC 21 Work Programme », disponible en anglais en suivant ce lien : CEN - CEN/CLC/JTC 21, consulté le 25 juin 2026. D'autres normes, y compris des travaux de normalisation internationaux préexistants (comme les normes ISO/CEI), peuvent également être pertinentes lorsqu'il s'agit de traiter les biais et la transparence dans les systèmes d'IA. Les normes internationales ne sont généralement pas gratuites et sont soumises à des protections de la propriété intellectuelle, ce qui compromet la capacité des organismes de promotion de l'égalité à y avoir accès.

le traitement des biais dans les systèmes d'IA. Ils incluent les biais indésirables du fournisseur d'IA et du déployeur d'IA en fonction de leur spécification du système d'IA, dans le contexte du règlement sur l'IA ».

- **JT021008** « Cadre de fiabilité de l'IA », parties 1 (journalisation), 2 (précision et robustesse) et 3 (transparence et supervision humaine) : l'objectif de cette norme est de définir la terminologie, les concepts, les exigences et les directives sur la journalisation, la précision, la robustesse, la transparence et la supervision humaine.
- **JT021024** « Gestion des risques liés à l'IA » : l'objectif de cette norme est de définir les exigences et de fournir des orientations pour la gestion des risques des systèmes d'IA tout au long du cycle de vie de ce dernier. Ce document exige des fournisseurs qu'ils établissent des critères objectifs d'acceptabilité des risques, mais ne précise pas les niveaux de risque acceptables.

Il est peu probable que ces normes définissent des seuils ou expliquent en quoi un traitement différent entre des personnes ou des groupes de personnes constituerait une discrimination interdite. Selon Mittelstadt (2024), l'OPE devrait être conscient que les normes établiront les attentes quant à la façon dont les biais peuvent être mesurés et atténués dans les systèmes d'IA, mais ne détermineront pas si ces biais sont illégaux ou contraires à l'éthique.

Normes, évaluations de conformité et certificats

Les systèmes à haut risque conformes aux normes harmonisées sont présumés conformes aux exigences de la section 2 du règlement sur l'IA (exigences relatives aux systèmes d'IA à haut risque). Toutefois, il importe de noter que le respect des obligations de transparence ou d'autres obligations n'indique pas que l'utilisation d'un système d'IA ou de ses résultats est licite en vertu d'une réglementation ou d'une autre législation et devrait être sans préjudice des autres obligations de transparence imposées aux déployeurs de systèmes d'IA par le droit de l'Union ou le droit national (considérant 137). Les fournisseurs doivent lister les normes harmonisées appliquées au système d'IA. Lorsque des normes harmonisées n'ont pas été utilisées, les fournisseurs doivent expliquer de quelle manière les exigences du règlement sur l'IA (chapitre III, section 2 du règlement sur l'IA) ont été respectées, par exemple en utilisant d'autres normes et spécifications techniques pertinentes. Ces explications devraient permettre aux organismes de promotion de l'égalité de déterminer si une norme pertinente en matière de partialité, de fiabilité ou de gestion des risques harmonisée a été suivie.

Impact d'une présomption de conformité

Une présomption de conformité, quelle qu'elle soit, ne fait pas obstacle à l'ouverture d'une enquête ou à l'adoption de mesures en cas de risques de discrimination ou autre pour les droits humains. Tout d'abord, le règlement sur l'IA lui-même reconnaît qu'un système d'IA conforme au règlement peut toujours présenter un risque et l'article 82 prévoit une procédure pour faire face à cette situation. De même, l'article 7(3), du Règlement sur la sécurité générale des produits (UE) 2023/988 (qui peut s'appliquer aux systèmes d'IA qui ne sont pas à haut risque) prévoit que la présomption de conformité à l'exigence générale de sécurité du règlement n'empêche pas l'ASM de

prendre toutes les mesures appropriées en vertu du règlement lorsqu'il existe des éléments de preuve démontrant que, malgré la présomption, le produit est dangereux.

Remarque : en vertu du règlement (UE) 2019/1020 sur les ASM, lorsqu'un opérateur économique présente des rapports de test ou des certificats attestant de la conformité de son produit à la législation d'harmonisation de l'Union, l'ASM doit « tenir dûment compte » du rapport ou du certificat.

Analyse d'impact sur les droits fondamentaux

L'obligation de procéder à une analyse d'impact sur les droits fondamentaux (AIDF) ne s'applique qu'à certains déployeurs de systèmes d'IA à haut risque, comme le prévoient l'article 27 du règlement sur l'IA et l'annexe III 5(b) et (c) du règlement sur l'IA. Elle inclut :

a. les déployeurs de systèmes d'IA à haut risque qui sont des organismes publics ou des entités privées fournissant des services publics ;
b. les déployeurs de systèmes d'IA à haut risque utilisés pour évaluer la solvabilité des personnes physiques ou pour établir leur cote de crédit (à l'exception des systèmes d'IA utilisés pour la détection de fraudes financières), et ceux utilisés pour évaluer les risques ou déterminer les prix dans le domaine de l'assurance-vie et de l'assurance-maladie.

L'Omnibus numérique sur l'IA reconnaît un chevauchement entre l'AIDF et l'AIPD. Il peut donc s'avérer nécessaire de demander l'AIPD si elle est recoupée plutôt qu'incorporée à l'AIDF[40]

Après l'achèvement de l'AIDF, le déployeur est tenu, en vertu de l'article 27(3) et (5) du règlement sur l'IA, de notifier les résultats de l'AIDF à l'ASM compétente en soumettant un modèle (standardisé) complété. En outre, une synthèse des conclusions de l'AIDF doit être présenté par le déployeur dans le cadre de son processus d'enregistrement dans la base de données de l'UE, pour ceux qui ont des obligations d'enregistrement en vertu de l'article 49(3) du règlement sur l'IA. Les déployeurs de la seconde catégorie (ceux du secteur du crédit et des assurances) ne sont pas soumis à cette obligation. Les organismes de promotion de l'égalité devraient donc trouver une synthèse dans la section C de la base de données de l'UE, telle qu'elle figure à l'annexe VIII, uniquement pour les AIDF obligatoires menées par des organismes publics ou des déployeurs privés relevant de la première catégorie.

Remarque : la mise en œuvre de ces deux obligations est limitée dans le cas d'autorités publiques déployant des systèmes d'IA à haut risque dans les domaines de la mise en application de la loi, de la migration, de l'asile et du contrôle aux frontières. En vertu de l'article 49(4) du règlement sur l'IA, aucune synthèse des conclusions de l'AIDF ne doit être soumise à la base de données de l'UE (pas même dans sa section sécurisée et non publique). En outre, en vertu de l'article 27(3) du règlement sur l'IA, ces déployeurs, dans des cas exceptionnels (y compris pour des raisons de sécurité publique) peuvent être

40. Règlement (UE) 2026/1744 du Parlement européen et du Conseil du 8 juillet 2026 modifiant les règlements (UE) 2024/1689, (UE) 2018/1139 et (UE) 2023/1230 en ce qui concerne la simplification de la mise en œuvre des règles harmonisées concernant l'intelligence artificielle (train de mesures omnibus numérique sur l'IA), disponible : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=OJ:L_202601744 (consulté le 24 juillet 2026).

exemptés de l'obligation de notifier les résultats de l'AIDF aux autorités de surveillance du marché (ASM) suite à la dérogation à la procédure d'évaluation de la conformité décrite à l'article 46(1) du règlement sur l'IA (voir également l'encadré ci-dessous sur le domaine de la mise en application de la loi). Les possibilités pour les organismes de promotion de l'égalité d'accéder à l'information dans de tels cas sont donc limitées en vertu du règlement sur l'IA.

Les déployeurs d'autres catégories de systèmes d'IA à haut risque (par exemple une entreprise de recrutement qui utilise l'IA pour présélectionner ou sélectionner des candidat·es) ne sont pas légalement tenus d'effectuer une AIDF. Cependant, cela n'implique pas nécessairement qu'une AIDF ne sera pas effectuée dans ces cas. Certaines organisations peuvent le faire en vertu de la législation nationale, conformément à des cadres tels que les Principes directeurs relatifs aux entreprises et aux droits de l'homme des Nations Unies, ou pour s'aligner sur leurs propres engagements en matière de diligence raisonnable ou de droits humains. Dans ces cas, les organismes de promotion de l'égalité peuvent demander leur divulgation dans le cadre de leurs pouvoirs de collecte d'informations.

Les fournisseurs de systèmes d'IA à haut risque ne sont pas soumis à l'obligation de procéder à une analyse d'impact sur les droits fondamentaux du type spécifique visé à l'article 27 du règlement sur l'IA. Toutefois, la réalisation d'une AIDF en tant qu'outil général d'identification, d'évaluation et de traitement des incidences négatives potentielles ou réelles sur les droits fondamentaux, ancrée dans le cadre de l'évaluation de l'impact sur les droits humains (AIDH), fait partie de l'évaluation de la conformité et de l'obligation du fournisseur en matière de système de gestion des risques et de gestion de la qualité, comme le prévoient les articles 9(2), 17(1)(g) et 43 du règlement sur l'IA (Mantelero 2024). Ce type d'analyse d'impact suivrait généralement une approche normalisée et pourrait également être demandé, avec le reste de la documentation conservée en vertu du règlement sur l'IA, par les organismes de promotion de l'égalité exerçant les pouvoirs qui leur sont conférés en vertu de l'article 77 du règlement sur l'IA (voir ci-dessus au paragraphe 2.1.3).

Outre les exigences légales prévues par le règlement sur l'IA, les lois, politiques ou pratiques nationales ou régionales peuvent exiger la réalisation et, dans certains cas, la publication d'analyses d'impact sur les droits humains ou les droits fondamentaux. Dans ces cas, les organismes de promotion de l'égalité peuvent demander leur divulgation dans le cadre de leurs pouvoirs de collecte d'informations.

Analyse d'impact sur la protection des données (AIPD)

Tout traitement de données à caractère personnel susceptible d'entraîner un haut risque pour les droits des personnes concernées en raison de sa nature, de sa portée, de son contexte ou de ses finalités nécessite que la personne responsable du traitement effectue une analyse d'impact sur la protection des données (AIPD), conformément à l'article 35(1) du RGPD. En conséquence, les AIPD sont, par définition, très pertinentes dans les cas de systèmes d'IA à haut risque impliquant le traitement de données à caractère personnel, la prise de décision automatisée et le profilage (cf. article 35(3) du RGPD).

Toutefois, la présentation de synthèses d'AIPD dans la base de données de l'UE (section C) n'est requise que pour certains groupes restreints de déployeurs de systèmes d'IA à haut risque, qui sont tenus d'enregistrer leur utilisation de l'IA dans la base de données de l'UE en vertu de l'article 49(3). Ces groupes incluent les autorités publiques, les institutions, organes et organismes de l'Union – ou les personnes agissant en leur nom qui utilisent des systèmes d'IA à haut risque. Pour le reste des déployeurs et des systèmes d'IA à haut risque, il n'existe aucune obligation légale directe en vertu du RGPD ou du règlement sur l'IA de publier ou de fournir l'AIPD à un organisme de promotion de l'égalité. Tout contrôle concernant l'analyse d'impact est laissé à l'initiative des APD – les canaux de collaboration avec les APD pourraient s'avérer pertinents à cet égard. Cela ne devrait pas décourager l'organisme de promotion de l'égalité de formuler une telle demande, dans le cadre des pouvoirs qui lui sont conférés en matière de collecte d'informations en vertu du droit national. Certaines organisations peuvent accepter de fournir l'AIPD si celle-ci est disponible ou peuvent même choisir de publier une synthèse de celle-ci. L'AIPD peut être demandée lors de la correspondance préalable à l'action ou si un litige est engagé.

La proposition de règlement de l'Omnibus numérique[41] exige que l'EDPB établisse des listes uniques, à l'échelle de l'UE, des activités de traitement qui nécessitent ou non une AIPD, ainsi qu'une méthodologie et un modèle communs[42], qui devraient remplacer le système fragmenté actuel des listes nationales et harmoniser ces listes au niveau de l'UE.

2.2.4. Notification de l'ASM ou enquête *ex officio*

L'organisme de promotion de l'égalité peut s'attendre à recevoir des informations de l'ASM sur les systèmes d'IA qui représentent un risque pour les droits en matière de non-discrimination ou qui sont impliqués dans des incidents graves portant atteinte à ces droits. Pour plus de détails sur le type d'informations attendues, voir 1.1 C : Information par l'ASM.

Les OPE pourraient obtenir la documentation relative au règlement sur l'IA des ASM si une ASM a déjà accédé à cette documentation de la part du fournisseur. Cela pourrait accélérer le processus d'accès pour les OPE. Les organisations intimées ne peuvent empêcher les ASM de partager la documentation avec les OPE sous couvert de la protection du secret des affaires, en particulier lorsque les OPE demandent la documentation par l'intermédiaire de l'ASM (Conseil de l'Europe, 2025b). Pour un exemple analogue de cette application dans la pratique, voir CK c. Dun & Bradstreet Austria GmbH et Magistradt der Stadt Wien, paragraphe 76.

41 Article 3 Proposition de règlement du Parlement européen et du Conseil modifiant les règlements (UE) 2016/679, (UE) 2018/1724, (UE) 2018/1725, (UE) 2023/2854 ainsi que les directives 2002/58/CE, (UE) 2022/2555 et (UE) 2022/2557 en ce qui concerne la simplification du cadre législatif numérique, et abrogeant les règlements (UE) 2018/1807, (UE) 2019/1150 et (UE) 2022/868 ainsi que la directive (UE) 2019/1024 (Omnibus numérique), eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:52025PC0837, consulté le 25 juin 2026.

42 Le Comité européen de la protection des données a mené une consultation en mai 2026 sur un modèle d'AIPD, disponible ici : https://www.edpb.europa.eu/public-consultations/template-for-data-protection-impact-assessment_fr.

Lorsque l'affaire est traitée dans le cadre d'une enquête ex officio ou d'une notification par une ASM, il convient de prendre en considération les éléments suivants :

- sollicitez la divulgation des documents de l'ASM qui sont nécessaires à l'enquête de l'OPE ;
- organisez une table ronde ou un atelier avec les OSC concernées ou d'autres organisations intéressées et utiliser le débat pour recueillir des informations pertinentes ;
- sensibilisez et consultez les parties prenantes pertinentes en vue d'identifier les personnes ou groupes touchés, notamment par des enquêtes ou d'autres processus participatifs de collecte d'informations.

Domaine de mise en application de la loi

Les possibilités de collecte d'informations seront plus limitées dans le contexte de la mise en application de la loi, ce qui rendra plus difficile l'identification de l'utilisation de l'IA ou des systèmes algorithmiques et de toute discrimination qui en résulterait[43].

Règlement sur l'IA

- En vertu de l'article 49(4) du règlement sur l'IA, l'enregistrement des systèmes d'IA à haut risque dans les domaines de la mise en application de la loi, de la migration, de l'asile et du contrôle aux frontières doit non seulement avoir lieu dans une section non publique sécurisée de la base de données de l'UE visée à l'article 71 du règlement sur l'IA, mais aussi ne comporter qu'un ensemble spécifique d'informations limitées, à l'exclusion par exemple des informations sur les données sous-jacentes et la logique de fonctionnement. En outre, cette section de la base de données non publique est accessible par les ASM désignées pour les systèmes d'IA de mise en application de la loi et de contrôle aux frontières conformément à l'article 74(8), et en principe non accessible aux OPE. Pour de plus d'informations, voir Restrictions d'accès à l'information au point 2.2.2.
- Conformément à l'article 27(1) du règlement sur l'IA, les autorités de mise en application de la loi sont tenues, en tant qu'autorités publiques, de procéder à une analyse d'impact sur les droits fondamentaux (AIDF) avant de déployer un système d'IA à haut risque. Toutefois, deux exceptions s'appliquent à cet égard. Premièrement, en vertu de l'article 49, paragraphe 4, point c) du règlement sur l'IA, ces déploiements ne soumettent pas de résumé des conclusions de la AIDF ou de la AIPD à la base de données de l'UE – pas même dans sa section sécurisée et non publique. Deuxièmement, elles peuvent, dans des cas exceptionnels, y compris pour des raisons de sécurité publique, être autorisées par l'ASM à mettre sur le marché ou à mettre en service un système d'IA à haut risque sans se conformer aux procédures d'évaluation de la conformité ou, dans des cas « urgents », sans autorisation (article 46(1) et article 46(2) du règlement sur l'IA). En vertu de l'article 27(3) du règlement

43. Pour une comparaison entre les droits en matière de données algorithmiques dans le RGPD et la DPJ, voir González Fuster (2020).

sur l'IA, cela inclurait également une exemption de l'obligation de notifier les résultats de l'AIDF à l'ASM.

Note: les articles 27 et 46 du règlement sur l'IA ne prévoient aucune exemption à l'obligation de mener l'AIDF dans son ensemble, et la condition énoncée à l'article 46, paragraphe 1, est spécifiquement limitée à une « période restreinte pendant laquelle les procédures d'évaluation de la conformité nécessaires sont en cours ». L'article 46(2), exige en outre que, en cas d'urgence, l'autorisation soit demandée « pendant ou après l'utilisation, sans retard injustifié ».

Il est important de noter que si l'autorisation est refusée par la suite :

- l'utilisation du système à haut risque sera arrêtée avec effet immédiat ;
- tous les résultats de cette utilisation seront immédiatement écartés.

Les voies possibles pour que les organismes de promotion de l'égalité puissent accéder aux informations relatives à la fois à l'AIPD et à l'AIDF sont examinées ci-dessous.

Directive « Police-Justice »

En outre, la directive « Police-Justice » (DPJ)[44] prévoit plusieurs exceptions et restrictions possibles aux droits à l'information des personnes par le biais de transpositions nationales, si nécessaire pour éviter de porter préjudice à la prévention, à la détection, aux enquêtes ou aux poursuites pénales ou à l'exécution de sanctions pénales et pour protéger la sécurité publique et nationale (article 13(3) et 15 de la DPJ).

En conséquence, les pratiques diffèrent d'un État membre à l'autre en ce qui concerne le type d'information à fournir et la manière dont il est possible d'y accéder. Il est donc crucial de se référer à la législation nationale dans ce domaine.

Prise de décision automatisée

- L'article 11 de la DPJ interdit certaines formes de traitement automatisé, y compris le profilage, et exige que ce traitement soit explicitement autorisé par le droit de l'Union ou de l'État membre, sous réserve de garanties appropriées pour les droits et libertés de la personne concernée. Le niveau minimal de protection requis est le droit d'obtenir une intervention humaine de la part de la personne responsable du traitement. Il sera nécessaire d'examiner les cadres nationaux régissant le traitement automatisé par les services de mise en application de la loi afin de déterminer quelles informations sur le traitement automatisé devraient être disponibles. Pour différentes transpositions dans différents États membres, voir Vogiatzoglou et Marquiene 2022 ; voir également le considérant 38.

44 La référence à la DPJ en ce qui concerne les sources d'information n'est pas considérée comme exhaustive. D'autres instruments relevant du domaine plus vaste de la liberté, de la sécurité et de la justice, comme la décision du Conseil SIS II, VIS et EURODAC peuvent également contenir des règles de fond et de procédure relatives aux droits des personnes concernées, y compris les droits à l'information et à son accès.

- Les articles 13 et 14 de la DPJ ne stipulent pas explicitement les obligations en matière d'information concernant l'utilisation de la prise de décision automatisée, y compris le profilage.

Divulgation d'informations

- En vertu de l'article 13(1) de la DPJ, les personnes responsables du traitement des données sont tenues de fournir certaines informations minimales, notamment leur identité et leurs coordonnées, le droit des personnes concernées d'accéder à leurs données à caractère personnel et les finalités prévues du traitement des données.

Ces informations ne concernent pas une personne concernée donnée, mais une procédure de traitement donnée et toutes les personnes concernées potentiellement affectées par cette procédure. Elles sont généralement mises à disposition sur le site internet dédié de l'autorité compétente. Il a été signalé qu'il n'est pas toujours facile ou possible de trouver les informations pertinentes sur le site Web de l'entité centralisée régissant les autorités chargées de la mise en application de la loi, l'autorité de la police nationale ou le ministère compétent (Vogiatzoglou et Marquiene 2022).

- L'article 13(2), de la DPJ élargit le champ des informations à partager avec les personnes concernées, y compris la base juridique du traitement et, le cas échéant, les catégories de destinataires avec lesquels les données à caractère personnel sont partagées et, si nécessaire, d'autres informations, en particulier lorsque les données à caractère personnel sont collectées à l'insu de la personne concernée. En vertu de l'article 13(3), cette mesure peut être retardée, restreinte ou même omise en vertu du droit national si une telle mesure répond aux critères de nécessité et de proportionnalité dans une société démocratique.

Ces informations, si elles sont suffisamment précises, peuvent être proposées sur le site Web, avec les informations minimales précédemment citées.

Droit d'accès

- En vertu de l'article 14 de la DPJ, les personnes concernées ont le droit d'accéder à davantage d'informations sur les activités de traitement de données que les informations générales mises à la disposition du public, y compris les finalités du traitement, les catégories de données concernées et les destinataires ou catégories de destinataires auxquelles les données sont divulguées, ainsi que la communication des données en cours de traitement et leur origine.

Contrairement au RGPD, l'article ne prévoit pas explicitement le droit d'être informé de l'existence d'une prise de décision automatisée, ni de la logique sous-jacente, ni des conséquences potentielles pour la personne concernée. Toutefois, les autorités chargées de la mise en application de la loi peuvent être tenues, en vertu de la législation nationale, de divulguer des informations sur la manière dont une personne a fait l'objet de profilage.

- En vertu du considérant 43 de la DPJ, les informations fournies ne doivent pas nécessairement inclure une copie réelle des données traitées, comme dans le cadre du RGPD. Au contraire, une synthèse complète est suffisante. Le contenu

de la réponse – lorsque l'accès n'est pas refusé pour des motifs formalistes – est souvent réduit à des réponses normalisées, brèves et dépourvues d'informations substantielles, mais néanmoins jugées formellement conformes (Erdogan 2024).

- En vertu de l'article 15 de la DPJ, le droit d'accès peut être retardé, restreint ou même omis selon le droit national, si une telle mesure est nécessaire et proportionnée dans une société démocratique pour la protection, entre autres, de la sécurité publique ou nationale, ou pour questionner sans entrave, mener des enquêtes ou des procédures officielles ou judiciaires.

Demande d'accès

- L'article 14 et l'article 17(3) de la DPJ prévoient deux types de procédure pour l'exercice du droit d'accès. Tout d'abord, la procédure « d'accès direct » prévue à l'article 14 de la DPJ, selon laquelle les individus peuvent s'adresser directement à l'organe de mise en application de la loi, à savoir les personnes responsables du traitement, pour poser des questions sur leurs données à caractère personnel. Ensuite, une procédure exceptionnelle, l'accès « indirect », dans laquelle le droit d'accès est exercé par l'autorité nationale de protection des données (APD) agissant en tant qu'intermédiaire.

l'accès indirect est la solution par défaut, ont été critiquées pour leurs lacunes structurelles et procédurales (Dimitrova et de Hert 2024), qui limitent le contrôle effectif des activités de traitement des données et la communication des informations pertinentes aux personnes concernées.

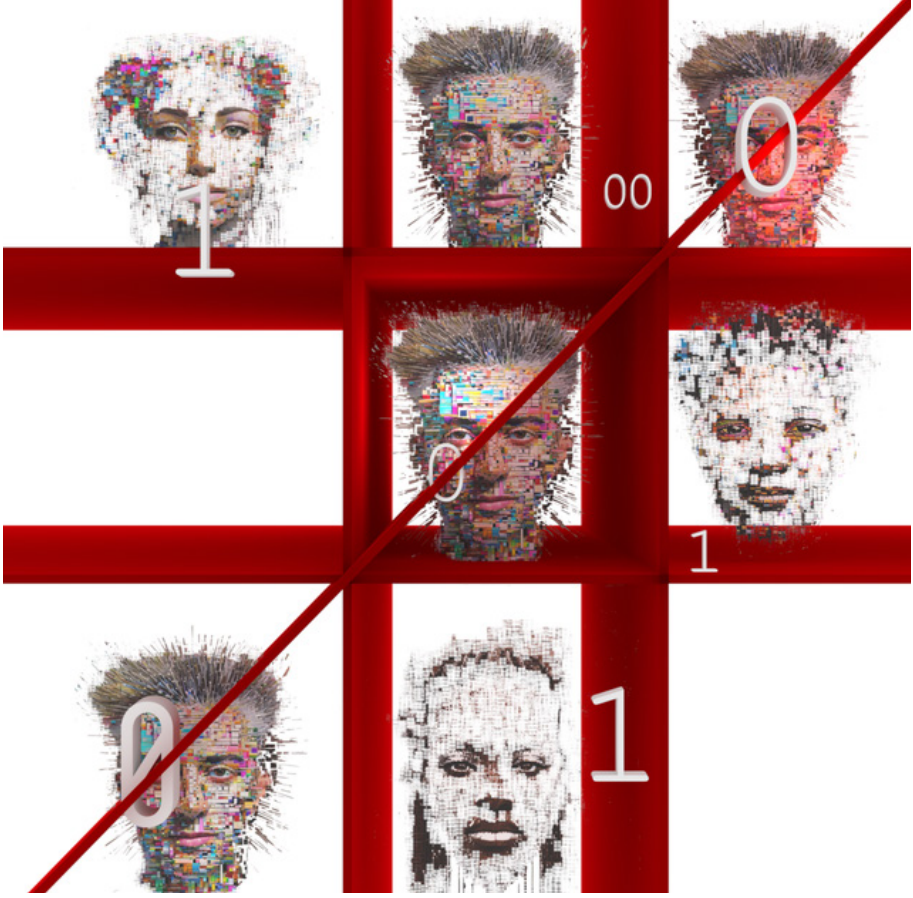
Remarque : lorsque des informations sont collectées par une autorité compétente à des fins de mise en application de la loi, elles relèvent du champ d'application de la DPJ – par exemple, les données relatives aux antécédents judiciaires d'une personne. Lorsque les informations sont collectées à partir d'une autre source publique ou à des fins autres que la mise en application de la loi (comme l'état civil, le statut professionnel ou l'éducation), elles relèvent du champ d'application du RGPD. Dans certains systèmes d'IA utilisés par les forces de l'ordre, les deux types de données d'entrée peuvent être utilisés, et il faut donc veiller à déterminer si les droits à l'information applicables sont régis par le RGPD ou la DPJ. Cette situation devient encore plus complexe lorsque des entreprises privées sont impliquées dans le développement de logiciels ou dans le traitement ou le contrôle de données pertinentes (voir Erdogan 2023). En cas de doute, il convient de solliciter l'avis d'un·e spécialiste du droit.

Registre des activités de traitement

- En vertu de l'article 24 de la DPJ, les personnes responsables du traitement doivent tenir un registre des activités de traitement, y compris des informations sur les finalités du traitement, les catégories de destinataires auxquelles les données à caractère personnel ont été ou seront divulguées et, le cas échéant, l'utilisation du profilage. Ces informations sont mises à la disposition des autorités chargées de la protection des données sur demande (pour les possibilités de collaboration avec les APD, voir Coopération intersectorielle).

Analyse d'impact sur la protection des données (AIPD)

Dans le contexte de la mise en application de la loi, l'article 27 de la DPJ exige une analyse d'impact sur la protection des données lorsqu'un type de traitement utilisant de nouvelles technologies est susceptible d'entraîner un haut risque pour les droits et libertés des personnes physiques. Les obligations d'AIPD varieront d'un État membre à l'autre en fonction des transpositions nationales. L'article 29 du Groupe de travail sur la protection des données recommande que les législations nationales imposent aux personnes responsables du traitement l'obligation d'effectuer une AIPD en relation avec les décisions automatisées. Toutefois, comme dans le cas de l'AIDF, les autorités de mise en application de la loi ne sont pas tenues de soumettre une synthèse de celle-ci à la base de données de l'UE (article 49(4)(c) du règlement sur l'IA).



Étape 3

Méthodologie d'évaluation des informations reçues

L'étape 3 présente les types de discrimination causés par les systèmes d'IA, qu'ils soient utilisés pour la prise de décision automatisée ou pour aider les décideurs et décideuses.

L'introduction aux concepts définit les types de discrimination et les applique à un contexte d'utilisation de l'IA. Les nouveaux concepts pour les organismes de promotion de l'égalité peuvent inclure la discrimination par proxy et les variables. Elle explique également comment identifier et établir une présomption de discrimination (prima facie).

La partie opérationnelle de l'étape 3 est conçue pour aider les personnes en charge du dossier à évaluer les informations reçues lors de l'étape 2 et la feuille d'enregistrement d'échantillon en Annexe K peut être utilisée en parallèle pour enregistrer ces informations. Dans les étapes 3.2 à 3.6, la personne en charge du dossier reçoit une brève explication, une série de questions à prendre en considération, des sources

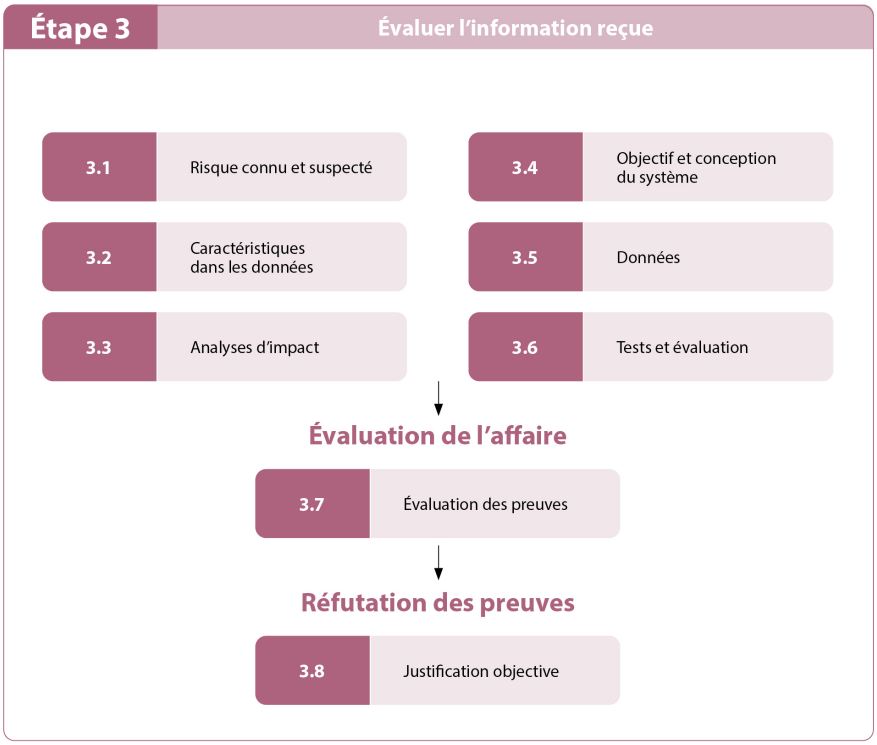
potentielles d'information qui peuvent aider à répondre à ces questions et des exemples d'indicateurs de discrimination ou de pratique discriminatoire.

L'étape 3.7 rassemble toutes ces conclusions et fournit un guide pour évaluer si une présomption de discrimination est établie et surtout, ce qu'il faut faire lorsque les éléments de preuve disponibles sont insuffisants ou qu'il y a un manque de transparence

L'étape 3.8 fournit des lignes directrices sur l'évaluation de la probabilité que l'organisation intimée soit en mesure de réfuter ou de justifier toute discrimination.

Les personnes en charge du dossier ne doivent pas nécessairement être en mesure de répondre à toutes les questions ou d'avoir accès à tous les documents ou éléments de preuves. L'objectif est plutôt de leur permettre de tirer le meilleur parti des informations dont elles disposent et d'en apprécier la pertinence.

Illustration 4. Étape 3 – Évaluation des informations reçues



À l'étape 2, vous devriez avoir engagé le processus de collecte de la documentation pertinente et d'autres preuves. À l'étape 3, vous utiliserez ces éléments pour évaluer si vous êtes en mesure d'établir une présomption ou un commencement de preuve prima facie de discrimination, ainsi que la probabilité que l'organisation intimée soit en mesure de renverser cette présomption ou de fournir une justification.

Pour bon nombre des sous-étapes de l'étape 3, une expertise technique peut être utile. Il devrait s'agir d'un·e expert·e ayant une bonne compréhension des statistiques, de

la science des données et de l'IA. Les concepts qui sont au cœur du développement des systèmes algorithmiques comme les données d'entraînement, les étiquettes de vérité terrain, les résultats et les entrées ne sont pas expliqués plus en détail dans cette méthodologie. Ils sont supposés être des connaissances de base dans les domaines de la statistique, de la science des données et de l'IA. Lorsque vous employez des expert·es techniques, assurez-vous qu'ils ou elles comprennent à la fois la science des données et les principes de non-discrimination.

Dans certains cas, les ASM ou d'autres agences de l'État membre disposeront de l'expertise technique requise qui pourra être utilisée par les organismes de promotion de l'égalité (voir Coopération intersectorielle).

En outre, l'expertise sectorielle peut parfois être utile pour comprendre les tendances des inégalités historiques et existantes dans le secteur où le système est appliqué. En cas de recours à une expertise sectorielle, il peut être utile d'employer des expert·es ayant une expérience en sciences sociales et en recherche empirique, car cette expérience aide à combler le fossé entre l'expertise sectorielle et l'expertise technique. Si ce n'est pas possible, il peut être nécessaire que l'OPE mène d'autres recherches sur les inégalités historiques et autres afin de pouvoir évaluer la possibilité qu'elles aient affecté le système. Par exemple, la surveillance excessive connue de certaines communautés ou les disparités de revenus entre les genres.

Une recommandation visant impliquer des expert·es est mentionnée, le cas échéant, à chaque sous-étape.

Remarque : comme indiqué dans la méthodologie, le nouvel article 77(1)(b) sur l'exigence de coopération et d'assistance mutuelle peut permettre à l'ASM ou à une autre autorité ou un autre organisme public, en vertu de la loi, de fournir l'expertise requise.

Remarque : en cas d'insuffisance d'informations disponibles, vous devrez :

- soit tenir compte de l'impact du manque de transparence sur la charge de la preuve (voir transparence à l'étape 3.7) ;
- soit envisager de demander un évaluation du système et un test effectué par l"ASM (voir étape 4.2.1)).

Un niveau raisonnable de connaissance du droit de la non-discrimination est requis pour utiliser cette méthodologie mais les concepts de base sont reformulés au début de cette étape pour mieux montrer de quelle manière ils se connectent aux systèmes algorithmiques.

Introduction aux concepts

Types de discrimination

Discrimination directe

Une personne est traitée de manière moins favorable qu'une autre se trouvant dans une situation comparable, en raison d'une caractéristique qu'elle possède et qui

relève d'un motif protégé (et, dans les affaires au sein de l'UE concernant l'âge/le handicap et les exigences professionnelles essentielles et déterminantes, aucune justification n'est admise).

Exemple de discrimination algorithmique directe

En Finlande, un homme s'est vu refuser un prêt par une institution financière, qui a utilisé un système de notation fondé sur des données telles que le lieu de résidence, le sexe, l'âge et la langue maternelle pour évaluer le risque de défaut de remboursement. Il s'est avéré que l'homme aurait obtenu le prêt s'il avait été une femme ou si le suédois avait été sa langue maternelle. De plus, aucune évaluation individualisée du risque n'a été réalisée. Le prêt lui a simplement été refusé sur la base d'hypothèses statistiques liées à son sexe et à sa langue maternelle. Saisi par le Médiateur contre les discriminations d'une demande de requête, le Tribunal national finlandais pour la non-discrimination et l'égalité a conclu, en 2018, qu'il s'agissait d'un cas de discrimination directe.

Lectures supplémentaires

- Évaluation de la solvabilité, autorité, discrimination multiple directe, sexe, langue, âge, lieu de résidence, motifs financiers, amende conditionnelle, Tribunal de Finlande (2018)

Discrimination indirecte

Une disposition, un critère ou une pratique apparemment neutre qui désavantage particulièrement les membres d'un groupe protégé par rapport à d'autres personnes se trouvant dans une situation comparable, à moins que :

- la disposition, le critère ou la pratique est objectivement justifié(e) par un objectif légitime ; et
- les moyens d'atteindre cet objectif sont appropriés et nécessaires.

Aucune décision ne fait encore autorité sur la question de savoir si certains ou tous les processus décisionnels automatisés constituent une disposition, un critère ou une pratique. Toutefois certains praticien·nes soutiennent fermement qu'ils le constituent presque certainement : « À notre avis, les processus décisionnels automatisés qui utilisent des algorithmes sont presque certainement une disposition, un critère ou une pratique aux fins d'une plainte pour discrimination indirecte » (Allen et Masters 2019).

Exemple de discrimination algorithmique indirecte (affaire Deliveroo)

Le système algorithmique de « réservation en libre-service » utilisé par Deliveroo pour l'attribution des équipes de travail s'est révélé être indirectement discriminatoire, en particulier en raison des convictions (participation à des actions collectives telles que les grèves) et du handicap. En s'appuyant sur des statistiques de participation et de fiabilité, des critères apparemment neutres, le système désavantageait dans la pratique les employé·es qui s'abstiennent de travailler pour des raisons légitimes,

comme une grève, une maladie ou des responsabilités familiales, en réduisant leurs scores et, par conséquent, leur accès aux futures opportunités de travail.

« L'aveuglement » des statistiques utilisées, c'est-à-dire la non-prise en compte par le système des raisons d'indisponibilité liées aux motifs protégés, a conduit à traiter de manière identique des situations différentes, entraînant un incidence disproportionnée sur les groupes protégés.

Lectures supplémentaires

- *Filcams Cgil Bologna e.a. c. Deliveroo Italia S.R. (2020)*
- Purificato (2021), disponible en anglais.

Discrimination multiple

La discrimination peut être fondée sur plusieurs motifs. Par exemple, une personne peut être victime d'une discrimination fondée sur la « race » et sur le sexe ou sur le handicap. Plusieurs termes sont utilisés pour désigner la discrimination fondée sur plusieurs motifs. Parmi les termes utilisés, on peut citer « discrimination multiple », « discrimination cumulée », « discrimination mixte », « discrimination combinée » et « discrimination intersectionnelle » (Fredman 2016). Leurs significations présentent toutes des différences subtiles. Pour cette méthodologie, il est important de reconnaître que les fondements de la discrimination peuvent être multiples (Xenidis 2021).

Discrimination intersectionnelle

« En substance … l'intersectionnalité est une façon de penser l'identité et sa relation au pouvoir » (Crenshaw 2015).

La discrimination intersectionnelle se distingue des autres formes de discrimination multiple en ce qu'elle ne résulte pas d'une simple addition de discriminations fondées sur un seul motif ; différentes caractéristiques se recoupent et ne peuvent être dissociées. Les structures sociales, économiques, culturelles et institutionnelles s'imbriquent, s'entrecroisent et forment un système complexe qui crée (et perpétue) l'inégalité (Crenshaw, 1991). Par exemple, des catégories sociales telles que l'âge et le sexe peuvent se chevaucher pour créer des avantages ou des désavantages selon les circonstances. Les inégalités peuvent varier en fonction des circonstances et des faits et dépendent parfois du moment et du lieu. Ainsi, une femme non migrante peut être « moins » égale qu'un homme mais plus égale qu'une femme migrante. Les inégalités sociales peuvent également façonner les rapports de pouvoir. Ces structures d'inégalités imbriquées se reflètent dans les données du monde réel utilisées pour concevoir, construire et tester des systèmes algorithmiques. En outre, les approches d'apprentissage automatique utilisées pour construire des systèmes d'IA sont capables de trouver des schémas d'interaction qui ne sont pas connus ou pas évidents à partir d'analyses de données de base. Ainsi, les systèmes algorithmiques peuvent conduire à une discrimination intersectionnelle (Xenidis 2021). Bien que la discrimination intersectionnelle ait été rejetée par la CJUE dans l'affaire Parris (C-443/15), il a été suggéré que cette question doit être réexaminée, notamment à la lumière de la reconnaissance croissante du phénomène dans le droit de l'Union,

comme la directive 2023/970/CE sur la transparence des rémunérations et la directive (UE) 2024/1385 relative à la lutte contre la violence à l'égard des femmes et la violence domestique (Howard 2024).

Discrimination par proxy

Une discrimination fondée sur l'utilisation de caractéristiques, appelées proxy, qui ne sont elles-mêmes attribuées à aucun des motifs de discrimination explicitement prohibés (voir caractéristiques protégées) mais qui sont corrélées à un motif protégé, est qualifiée de discrimination par proxy. Bien que la législation de l'UE ne reconnaisse pas explicitement la discrimination par proxy ou ne fasse pas référence à celle-ci, il s'agit d'une construction théorique qui aide à comprendre comment la discrimination algorithmique peut survenir.

- La discrimination par proxy peut constituer une discrimination directe s'il existe un lien indissociable entre les caractéristiques choisies (neutres) et les caractéristiques explicitement protégées (grossesse et genre, par exemple : Dekker c. Stichting Vormingscentrum voor Jong Volwassenen (VJV-Centrum) Plus, 1990 ; Dita Danosa c. LKB Līzings SIA, 2010 ; Andersen c. Region Syddanmark, 2010 ; VL c. Szpital Kliniczny im. dra J. Babińskiego SPZOZ w Cracovie, 2021.
- La discrimination par proxy peut également être une discrimination indirecte si la disposition, le critère ou la pratique, y compris les proxys, ont des répercussions disproportionnées sur les personnes ayant des caractéristiques protégées (pour une explication de cette situation dans la pratique, voir Jyske Finans c. Ligebehanlingsnævnet, 2017 (C-668/15)).

Affaire clé : Dekker c. Stichting Vormingscentrum (1990)

Un employeur contrevient directement au principe de l'égalité de traitement (directive du Conseil 76/207/CE) relative à la mise en œuvre de l'égalité de traitement entre hommes et femmes en ce qui concerne l'accès à l'emploi, à la formation et à la promotion professionnelles et les conditions de travail lorsqu'il refuse de conclure un contrat de travail avec une candidate apte à l'emploi, au motif que l'emploi d'une femme enceinte pourrait avoir des conséquences défavorables, la règlementation nationale en matière d'assurance maladie assimilant la grossesse à la maladie. La Cour a reconnu que seules les femmes pouvaient se voir refuser un emploi pour cause de grossesse et qu'il s'agissait, dès lors, d'une discrimination directe fondée sur le sexe.

Variables proxy

Dans le contexte du droit de la non-discrimination, les « proxys » ou « variables proxy » désignent des caractéristiques ou des combinaisons de caractéristiques qui ne sont pas elles-mêmes listées parmi les motifs de discrimination explicitement interdits dans le cadre juridique pertinent, mais qui sont corrélées à ces motifs de telle sorte que leur utilisation peut conduire à une discrimination illicite. Le code postal en est un exemple : un code postal n'est pas discriminatoire en soi ; toutefois, si de nombreuses personnes migrantes ou personnes d'une origine ethnique particulière

vivent dans une certaine région, il peut servir de substitut à un motif prohibé, à savoir l'origine ethnique.

Les proxys peuvent être identifiés par la surreprésentation : si des personnes avec un attribut protégé spécifique sont surreprésentées dans un groupe, les identifiants de ce groupe sont probablement des proxys. Ainsi, dans l'exemple ci-dessus, dès lors que des personnes d'origines différentes sont surreprésentées dans certains quartiers, le code postal peut être considéré comme un proxy.

Lorsqu'un proxy est utilisé comme donnée d'entrée d'un système algorithmique, cela peut constituer un indice de discrimination potentielle. Il convient de noter que, selon le type de proxy, son utilisation peut révéler tant une discrimination directe qu'indirecte (Weerts et al. 2024). Certaines caractéristiques sont connues, grâce à la recherche sur le sujet et à la jurisprudence pertinente, pour être des proxys de caractéristiques protégées. Les proxys peuvent également être identifiés en comprenant et en analysant le contexte spécifique dans lequel un système algorithmique est appliqué, en tenant compte des inégalités historiques et démographiques existantes.

Pour plus de détails et des exemples de différents types de proxy, reportez-vous à l'Annexe E.

Dans l'affaire Chez (C-83/14), une mesure collective – consistant à installer les compteurs d'électricité dans un quartier « rom » à une hauteur comprise entre six et sept mètres plutôt qu'à moins de deux mètres comme dans les autres quartiers – a été imposée à un quartier habité principalement, mais pas exclusivement, par des personnes d'origine rom. Si cette mesure s'avérait avoir été introduite et/ou maintenue pour des raisons liées à l'origine ethnique commune à la plupart des habitants et habitantes du quartier concerné, elle serait considérée comme une discrimination directe ou indirecte, y compris à l'encontre des habitants et habitantes de ce quartier qui ne sont pas d'origine rom. Il incombait à la juridiction nationale de déterminer cette question en tenant compte de toutes les circonstances pertinentes de l'affaire ainsi que des règles relatives au renversement de la charge de la preuve prévues à l'article 8(1) de la directive 2000/43.

Exemple d'algorithme : vérification de bourses d'études universitaires

Cet exemple devrait aider à expliquer comment fonctionnent les proxys. En l'espèce, le fait de « vivre près de parents » et « suivre une formation professionnelle » ont agi comme des indicateurs de « provenance des personnes migrantes non européennes ».

L'Agence exécutive pour l'éducation des Pays-Bas (DUO) effectue des contrôles pour s'assurer que les bourses d'études universitaires versées sont correctement accordées. Ce processus est connu sous le nom de vérification des bourses d'études (abréviation en néerlandais CUB). Des soupçons sont apparus, selon lesquels le processus CUB, désormais interrompu, était discriminatoire à l'égard des étudiant·es issu·es de l'immigration. Cette suspicion était fondée sur les procédures de recours engagées par des étudiant·es qui avaient été informé·es avoir reçu des bourses d'études à tort, dans lesquelles les avocat·es se sont rendu compte que leurs client·es étaient

majoritairement issu·es de l'immigration. L'enquête a mis en lumière le fait qu'un système de notation des risques de fraude a été utilisé dans ce processus pour hiérarchiser les enquêtes sur l'utilisation (il)légitime des bourses.

Algorithm Audit a effectué une analyse du système algorithmique, des différences de traitement dans les différentes étapes du processus CUB et a examiné si les entrées du système pouvaient être considérées comme des proxy[45]. Le système utilise trois entrées : l'âge de l'élève, le niveau d'instruction et la distance entre leur adresse et de celle de leurs parents. La nationalité, les antécédents migratoires ou toute autre information sur la « race » ou l'appartenance ethnique n'ont pas été utilisés dans le système de notation des fraudes.

Les principales constatations d'Algorithm Audit étaient les suivantes :

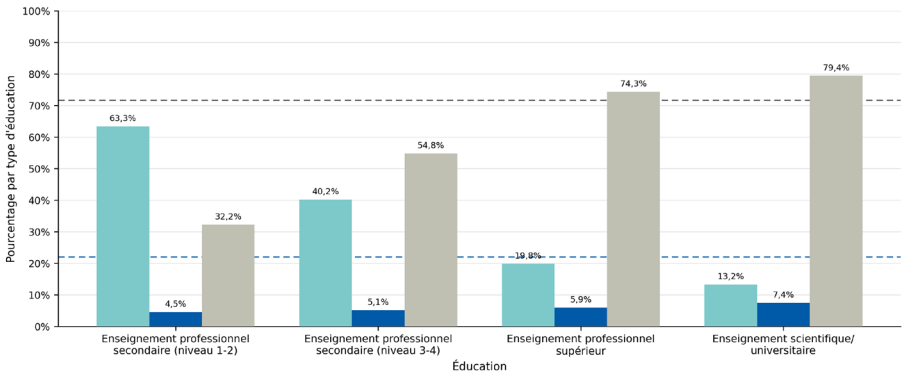
- le processus CUB a globalement conduit à un traitement inégal des étudiant·es issu·es de l'immigration non européenne ;
- Le profil de risque attribuait des scores de risque plus élevés aux étudiant·es issu·es de l'immigration ; les étudiant·es issu·es de l'immigration non européenne étaient deux fois plus susceptibles d'être classé·es dans une catégorie à haut risque que les étudiant·es d'origine néerlandaise ;
- la sélection manuelle supplémentaire des étudiant·es pour le contrôle a renforcé la surreprésentation plutôt que de la réduire. Les étudiant·es issu·es de l'immigration non européenne étaient 6,2 fois plus susceptibles d'être sélectionné·es pour un contrôle, qui impliquait une visite invasive à domicile, que les étudiant·es d'origine néerlandaise.

Un tribunal néerlandais a conclu que cette pratique n'était pas objectivement justifiée et constituait une discrimination indirecte (Rechtbank Overijssel 2024).

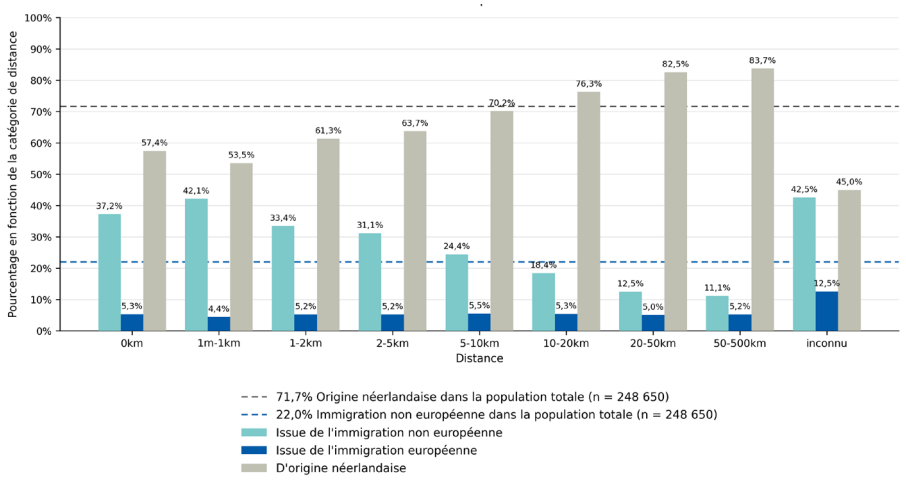
Des analyses détaillées ont montré que le système de notation des risques attribuait des scores de risque plus élevés aux élèves des niveaux d'éducation inférieurs (enseignement professionnel) et aux élèves vivant plus près de leurs parents. Il a été démontré que les deux variables sont (fortement) corrélées avec le contexte de migration et sont donc des proxys pour le contexte migratoire (voir les graphiques ci-dessous).

45 Algorithm Audit a obtenu un accès complet à l'algorithme et aux données internes de DUO. En raison du RGPD, DUO ne collecte ni ne traite d'informations sur l'origine des étudiant·es aux fins du processus CUB. Pour réaliser ces analyses, DUO et Algorithm Audit ont collaboré avec l'office national néerlandais de la statistique (Netherlands Statistics ou CBS). CBS a fourni des statistiques agrégées sur le pays de naissance et le pays d'origine en fonction des populations définies par Algorithm Audit pour chacune des étapes du processus (catégorie de risque, sélection pour le contrôle, contrôle des résultats, procédure d'appel, résultat de l'appel) et des variables d'entrée (éducation, âge, distance par rapport aux parents).

Graphique 1 : répartition des étudiant·es issu·es de l'immigration (non-) européenne et des étudiant·es d'origine néerlandaise par type d'éducation dans la population subventionnée par les bourses



Graphique 2 : répartition des étudiant·es issu·es de la migration (non-) européenne et des étudiant·es d'origine néerlandaise par catégorie de distance dans la population subventionnée par les bourses



Source graphique : Algorithm Audit

Pour une explication complète, voir Algorithm Audit 2024a et 2024b.

Éléments de comparaison

Un élément de comparaison est une personne qui se trouve dans une situation comparable à celle de la personne ayant déposé la plainte, mais qui ne partage pas la caractéristique protégée présumée être le motif de la discrimination.

Remarque : il est possible que deux groupes de personnes soient considérés comme étant dans une situation analogue aux fins d'une plainte mais pas d'une autre. Il est

donc toujours important d'évaluer l'élément de comparaison à la lumière de l'objectif de la mesure contestée (Conseil de l'Europe et Agence des droits fondamentaux de l'Union européenne 2018 : 44-49).

Le choix du bon élément de comparaison sera important lors de l'évaluation des preuves dans les systèmes algorithmiques, en particulier lors de l'évaluation de l'objectif du système et des mesures d'entraînement, de test et de validation, ainsi que des évaluations et tests proprement dits. Il peut notamment être nécessaire d'évaluer si des données de caractéristiques protégées ont été utilisées pour entraîner le système algorithmique et quels choix ont été faits lors des tests du système. Cette question est examinée plus en détail aux sections 3.3, 3.5 et 3.6.

Établir une présomption de discrimination

Conformément aux directives relatives à la non-discrimination, dès lors qu'une personne s'estime lésée par le non-respect à son égard du principe de l'égalité de traitement et établit, devant une juridiction ou une autre instance compétente, des faits qui permettent de présumer l'existence d'une discrimination directe ou indirecte, il incombe à la partie défenderesse de prouver qu'il n'y a pas eu violation du principe de l'égalité de traitement[46]. Voir également étape 3.8.

Cela signifie que la personne ayant déposé la plainte doit présenter suffisamment d'éléments de preuve pour faire naître une présomption de discrimination.

La charge de la preuve est alors renversée et il appartient à l'organisation intimée d'apporter la preuve contraire. En pratique, les juridictions nationales disposent d'une marge d'appréciation pour déterminer si les éléments de preuve présentés par la personne ayant déposé la plainte sont suffisants pour entraîner ce renversement de la charge de la preuve[47].

Les éléments de preuve peuvent comprendre des données statistique[48] , des documents écrits (s'ils comportent un contenu lié à la discrimination ou au cas particulier de la personne requérante) et des rapports et analyses provenant de sources pertinentes (voir données recueillies à l'étape 2), des tests de situation (voir étape 3.6), des évaluations de bon sens, des témoignages et des conclusions tirées de preuves circonstancielles. La CJUE souligne l'importance de tenir compte de « toutes les circonstances entourant la pratique litigieuse » (C-83/14 Chez, paragraphes 80-84). Par conséquent, des données statistiques et qualitatives peuvent être utilisées pour établir une présomption de discrimination et il apparaît que les deux approches devraient en fait être utilisées.

Qu'est-ce que des « données sur l'égalité » ?

Dans le contexte de la méthodologie, les données quantitatives (y compris statistiques) et qualitatives sont désignées comme des « données sur l'égalité », suivant la définition

46 Directive 2000/78/CE, article 10, directive 2006/54/CE, article 19, directive 2000/43/CE, article 8.
47 Dans la pratique, ce sera souvent différent pour la discrimination directe et indirecte. Pour une lecture plus approfondie, voir Farkas et O'Farrell (2015).
48 Pour en savoir plus sur les statistiques en tant qu'éléments de preuve, voir ci-dessous, « Approche des données statistiques ».

donnée dans les « Orientations pour améliorer la collecte et l'utilisation des données relatives à l'égalité » (Commission européenne, 2021) comme suit :

> toute information permettant de décrire et d'analyser la situation en matière d'égalité. Cette information peut être de nature quantitative ou qualitative. Ces données incluent les données agrégées qui reflètent les inégalités ainsi que leurs causes ou leurs effets dans la société. Parfois, les données qui sont collectées principalement pour des raisons qui ne concernent pas l'égalité peuvent être utilisées afin de produire des données sur l'égalité. (Commission européenne 2021: 7)

Les données sur l'égalité sont utilisées, entre autres, pour « fournir des preuves fiables dans des affaires administratives ou judiciaires concernant la discrimination grâce à des données qui mettent en évidence discrimination directe ou indirecte ; et soutenir le plaidoyer et la sensibilisation dans le domaine de l'égalité et de la non-discrimination » (Commission européenne 2021: 8).

Lectures supplémentaires

- Handbook on identifying and using equality data in legal casework, Ilieva, 2024 ;
- Strategic Litigation Handbook, Equinet, 2017 ;
- Handbook: artificial intelligence, judicial decision-making and fundamental rights, JuLIA Project, 2024.

Comment utiliser l'étape 3

Les sections ci-dessous exposent les questions qui peuvent aider les organismes de promotion de l'égalité à déterminer s'il existe une preuve prima facie de discrimination. L'objectif est de montrer où, dans les informations recueillies à l'étape 2, les éléments de preuve pertinents peuvent être trouvés, ce qu'il faut rechercher dans les éléments de preuve et comment les utiliser pour prouver la discrimination.

Chaque section de cette partie suit la structure de base suivante.

- Une explication de ce qui devrait être évalué par l'OPE.
- Des questions à prendre en considération lors de l'évaluation des informations.
- Les sources d'information potentielles.
- Des indicateurs de discrimination.

Les sections 3.7 et 3.8 fournissent ensuite un cadre pour évaluer si la preuve recueillie est suffisante pour étayer une affaire prima facie de discrimination impliquant un système d'IA ou algorithmique (3.7 Évaluer si la preuve est suffisante) et pour évaluer la probabilité que l'organisation intimée puisse la réfuter ou justifier une discrimination (3.8 L'organisation intimée peut-elle démontrer que la différence de traitement n'est pas discriminatoire ?). Ces informations aideront à déterminer si des mesures doivent être prises à l'encontre de l'organisation intimée. Des recoupements seront nécessairement faits au cours du processus et l'examen des preuves fait partie ces tâches.

Il est extrêmement improbable qu'un OPE puisse accéder à toutes les informations ou répondre à toutes les questions de cette étape. Les OPE sont plutôt destinés à servir de guide sur les types d'informations qui pourraient être disponibles et sur la manière dont elles pourraient être évaluées à la fois individuellement et collectivement.

Une suggestion de fiche de consignation a été fournie à l'annexe K. Celle-ci doit servir d'aide et peut être adaptée aux besoins individuels de l'OPE.

Lectures supplémentaires

Certains des concepts de l'étape 3 seront nouveaux ou peu familiers. Les ressources ci-dessous les expliquent plus en détail.

- Wachter, Mittelstadt et Russell, 2020, Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. Computer Law & Security Review, Tome 41, juillet 2021.
- Mittelstadt, 2024, How to use the Artificial Intelligence Act to investigate bias and discrimination: A guide for equality bodies, Equinet.

3.1. Quels sont les risques connus ou suspectés de discrimination ?

L'utilisation de systèmes algorithmiques peut s'accompagner d'un risque inhérent de pratiques discriminatoires dans certains contextes. Lorsqu'une discrimination algorithmique a déjà été démontrée dans un cas d'utilisation similaire dans un secteur similaire, cela pourrait indiquer un risque plus important de discrimination qui s'applique également à l'affaire examinée. Par exemple, lorsqu'il est démontré que la reconnaissance faciale utilisée dans les services de police est discriminatoire, il est probable que la reconnaissance faciale utilisée à d'autres fins le sera également.

3.1.1. Le type de système algorithmique ou le domaine dans lequel il est appliqué est-il explicitement mentionné dans le règlement sur l'IA pour ses risques de discrimination ?

Le règlement sur l'IA fait référence à la fois aux « secteurs » et aux « cas d'utilisation » au sein des secteurs.

Lorsque des systèmes algorithmiques sont utilisés dans des secteurs où il existe des inégalités historiques et/ou où il a été démontré que des discriminations se produisent fréquemment, les suspicions de discrimination devraient s'accroître. Les systèmes algorithmiques sont fondés sur des données réelles. Les inégalités existantes et les biais humains se reflètent dans les données et apparaissent donc souvent et peuvent être exacerbés dans le système algorithmique.

Biais historiques : exemple d'algorithme, l'outil de recrutement Amazon

Amazon a développé un outil de recrutement par IA entraîné sur des CV historiques qui reflétaient un ensemble de candidats majoritairement masculins. En conséquence, le système « a appris par lui-même que les candidats masculins étaient préférables ». Bien que le système n'utilise pas explicitement le genre ou le sexe en tant que donnée d'entrée, il s'appuie sur des données proxys corrélées avec le fait d'être de sexe féminin – comme le mot « femmes » (« capitaine de club d'échecs féminin », par exemple) ou les diplômes d'établissements d'enseignement pour femmes – et décline systématiquement les candidatures des candidates (Dastin 2018).

Nombre des cas d'usage classés comme présentant un risque élevé par le règlement sur l'IA ont été jugés relever de la catégorie des cas à haut risque par l'UE précisément en raison d'un risque de discrimination. Les considérants 54 et 56 à 60 du règlement sur l'IA mentionnent explicitement le risque de discrimination, la perpétuation d'inégalités historiques et des vulnérabilités spécifiques. Même si un système n'est pas technologiquement complexe et/ou n'emploie pas l'IA, l'utilisation de systèmes algorithmiques (reposant sur des règles) pour ces cas d'utilisation comporte un risque similaire de discrimination. L'annexe I décrit les domaines d'IA à haut risque et les cas d'utilisation.

Un cas d'utilisation explicitement mentionné dans le règlement sur l'IA justifie un niveau d'examen supplémentaire. En outre, des données sur l'égalité, tant qualitatives que quantitatives, peuvent être disponibles pour ces cas et figurent souvent dans la jurisprudence, les études universitaires, celles menées par des organisations de la société civile ou les rapports gouvernementaux. Voir aussi la section 2.1.2 Jurisprudence existante et autres sources.

Remarque : si le système est un système d'IA à haut risque et s'il est enregistré dans la base de données de l'UE, y compris dans la section B[49], il est par définition explicitement mentionné dans le règlement sur l'IA et peut donc présenter un risque élevé de discrimination.

Questions à envisager

- Le système algorithmique est-il un « type » de technologie connu pour créer un risque de discrimination, comme le profilage (article 6(3) et considérant 53), l'analyse prédictive des risques (considérant 42) ou la biométrie et la reconnaissance des émotions (annexe III(1), considérants 44 et 54) ?

49 Lorsqu'un système figure dans la section B de la base de données de l'UE, il a été auto-évalué par un fournisseur comme étant utilisé dans un secteur à haut risque mais ne présentant pas de risque sur le fondement d'un ensemble restreint d'exceptions (voir règlement sur l'IA, article 6(3)). Bien que la documentation exigée en vertu du règlement ne soit probablement pas disponible, ce système est appliqué dans un domaine explicitement mentionné dans le règlement pour ses risques de discrimination.

- La décision ou le traitement vécu était-il lié à l'un des secteurs suivants ?
  - Éducation et formation professionnelle (annexe III(3), considérant 56).
  - Emploi, gestion des effectifs et accès au travail indépendant (annexe III(4), considérant 57).
  - Accès aux prestations et services publics (bien-être) et jouissance de ces services (annexe III(5)(a)).
  - Notation de crédit (annexe III(5)(b), considérant 58).
  - Assurance vie et maladie (annexe III(5)(c), considérant 58).
  - Accès aux services de soins de santé et éligibilité à ces services (annexe III(5)(a), considérant 58).
  - Mise en application de la loi (annexe III(6), considérant 59).
  - Migration, asile et contrôle aux frontières (annexe III(7), considérant 60).
  - Administration de la justice (annexe III(8)(a), considérant 61).
  - Administration des processus démocratiques (annexe III(8)(b), considérant 61).
- Si l'affaire était liée à l'un des secteurs susmentionnés, l'affaire couvre-t-elle un cas d'utilisation listé dans la section pertinente de l'annexe III (y compris le considérant correspondant) ?

Sources potentielles d'information pour évaluer un risque de discrimination généralement établi

Règlement européen sur l'IA

- Article 6, annexe III et considérants – décrivent les types d'IA à haut risque.
- Article 5 – décrit les pratiques interdites en matière d'IA et les considérants 28 à 45.

Autres

- Lignes directrices de la Commission sur les pratiques interdites en matière d'intelligence artificielle (IA), telles que définies par le règlement sur l'IA.
- À venir : lignes directrices de la Commission sur la classification des systèmes d'IA à haut risque et exigences connexes[50].

Indicateurs d'un risque de discrimination

Si l'un de ces indicateurs s'applique, il ne constitue pas en lui-même une preuve de discrimination. Toutefois, ces indicateurs pourraient indiquer un système présentant un risque de discrimination :

- si le système est un système d'IA à haut risque dans les sections A ou C de la base de données de l'UE ;
- hors haut risque :

50. Au moment de la révision de cette méthodologie, la Commission a publié un projet de lignes directrices à des fins de consultation : La Commission sollicite un retour d'information sur le projet de lignes directrices pour la classification des systèmes d'intelligence artificielle à haut risque, 19 mai 2026.

- si le système est destiné à être utilisé dans un secteur dans lequel il existe un risque identifié de discrimination (annexe III ou considérants) ; et/ou
- s'il s'agit d'un type de technologie dans lequel il existe un risque connu de discrimination.

Indiquez si le système est un système à haut risque ou non ou s'il est destiné à être utilisé dans l'un des secteurs ou domaines identifiés comme présentant un risque de discrimination dans le règlement sur l'IA.

Suspicion de classification incorrecte en vertu du règlement sur l'IA
Si un organisme de promotion de l'égalité soupçonne qu'un système a été auto-évalué à tort comme n'étant pas un système d'IA à haut risque ou comme n'étant pas reconnu comme une pratique d'IA interdite (voir les suspicions de systèmes d'IA interdits à l'étape 1.4), il doit en informer l'ASM concernée, qu'un cas de discrimination puisse être établi ou non (voir également étape 4).

3.1.2. Un risque connu de discrimination dans ce secteur ou avec cette technologie a-t-il été identifié par les autorités nationales ou régionales, les OSC nationales ou internationales ou les organisations de défense des droits humains ?

Les rapports émanant des institutions de médiation, de la société civile et des milieux universitaires, ainsi que la jurisprudence existante, peuvent faire apparaître des résultats inégaux pour un motif de protection spécifique.

Utilisez les renseignements recueillis à l'étape 2 pour évaluer si ce système ou ce type de système est déjà connu ou soupçonné de produire un résultat discriminatoire. Toute information pertinente peut également vous aider à réaliser les étapes suivantes.

À cette étape, une expertise sectorielle serait bénéfique.

Sources potentielles d'information pour évaluer un risque de discrimination généralement établi

- Suivi, recherche ou enquêtes de l'OPE
- Rapports des OSC nationales et internationales
- Jurisprudence existante
- Recherches sur internet

► Expert·es

Indicateurs de discrimination

- Une discrimination a été identifiée ou soupçonnée par d'autres dans l'utilisation spécifique de systèmes algorithmiques par l'organisation intimée qui fait l'objet de la présente affaire.
- La discrimination a été identifiée ou soupçonnée par d'autres dans des utilisations très similaires des systèmes algorithmiques, comme le montre l'affaire, ce qui indique un schéma potentiel de discrimination.

Consignez la zone dans laquelle il existe un risque connu, la technologie et toute autre information et preuve pertinente. Par exemple :

Secteur : mise en application de la loi

Technologie : reconnaissance faciale, biométrie

Sources d'information : recherches menées par une OSC/un milieu universitaire, article 5 du règlement sur l'IA (systèmes interdits), article 6 (haut risque) et annexe III.

Cas d'utilisation qui devraient être considérés comme étant à haut risque ou interdits en vertu du règlement sur l'IA

est solide et cohérent mais qu'il ne figure pas dans les listes de l'annexe III, les OPE devraient envisager de signaler le cas d'utilisation à la Commission européenne pour qu'il soit ajouté à la liste prévue à l'article 7 (voir étape 4.3.2).

Si l'OPE estime qu'une utilisation de l'IA contredit fondamentalement les valeurs de l'Union de respect de la dignité humaine, de la liberté, de l'égalité, de la démocratie et de l'État de droit et des droits fondamentaux consacrés dans la Charte, mais qu'elle n'est ni interdite ni considérée comme un risque élevé en vertu du règlement sur l'IA (voir étape 4.3.2).

3.2. Existe-t-il des caractéristiques protégées ou des proxys connus dans les entrées ou les jeux de données ?

Les données d'entrée d'un système algorithmique définissent les données de sortie et donc le résultat. Par conséquent, si une caractéristique protégée (voir caractéristiques protégées) ou un proxy (voir discrimination par proxy) est utilisé en tant qu'entrée dans le système algorithmique, il devrait y avoir une suspicion de discrimination.

Dans cette étape, l'expertise sectorielle est utile.

Questions à envisager

- Les données à caractère personnel traitées révèlent-elles des caractéristiques protégées ou des proxys pour des caractéristiques protégées ?

- Les caractéristiques protégées ou les proxys connus (voir annexe E) sont-ils mentionnés dans l'un des documents pertinents ?
- Les caractéristiques protégées ou les proxys connus sont-ils mentionnés directement en tant que données d'entrée dans le système (que ce soit dans le cadre d'un entraînement, de tests ou en cours d'utilisation) ? La formulation pourrait être la suivante : « X, Y et Z sont utilisés pour prédire/calculer ».
- L'une des entrées du système ou une combinaison de ces entrées est-elle un proxy probable à un attribut protégé, fondé sur les connaissances sectorielles ? La réponse à cette question nécessitera peut-être de consulter des statistiques nationales ou des expert·es sectoriel·les ayant des connaissances sur la manière dont les inégalités existent et sont liées dans un secteur donné. Par exemple :
  - dans les pays où il existe une déduction fiscale pour les frais liés aux soins de santé, les personnes handicapées sont susceptibles d'être surreprésentées parmi celles qui utilisent cette déduction. Par conséquent, l'utilisation de cette déduction est une approximation probable de l'invalidité ;
  - dans de nombreux pays, les personnes issues de l'immigration sont sous-représentées parmi les titulaires de diplômes universitaires, de sorte que le niveau d'éducation est probablement un indicateur de l'origine raciale ou ethnique.
- Certains des proxys mentionnés sont-ils inextricablement liés à une caractéristique protégée ? Voir Annexe E – Proxys.
- La suspicion qu'une (sélection d') entrée(s) est un proxy peut-elle être étayée par des preuves ? Par exemple, des statistiques sont-elles disponibles auprès d'un bureau national de statistiques ou dans des recherches universitaires, montrant une surreprésentation ou une sous-représentation d'un groupe ayant une caractéristique protégée dans un sous-groupe spécifique des entrées ? Par exemple :
  - si le diplôme d'enseignement est l'un des facteurs soupçonnés d'être un indicateur de l'origine ethnique, existe-t-il des statistiques montrant que les personnes issues de l'immigration sont surreprésentées dans le groupe des titulaires de diplômes professionnels ?

Remarque : la collecte de ces preuves peut nécessiter la contribution spécialisée de collaborateurs et collaboratrices ayant de l'expérience en analyse de données.

- Le système est-il entraîné, testé ou validé sur la base d'informations contenant des caractéristiques protégées ? (Voir étape 3.5.)

Sources potentielles d'information pour comprendre les entrées dans le système

Personne ayant déposé la plainte

- Liste des informations fournies par la personne ayant déposé la plainte à l'organisation intimée.
- Explications, motivations et justifications fournies à la personne ayant déposé la plainte au sujet du traitement.

RGPD, articles 13 et 14 : informations fournies à la personne concernée

- Informations fournies à la personne concernée (personne ayant déposé la plainte) au moment où elle a fourni les données (article 13 du RGPD).
- Informations fournies par une personne responsable du traitement lorsque les données à caractère personnel n'ont pas été obtenues directement auprès de la personne concernée (article 14 du RGPD).
- La logique sous-jacente dans les cas de prise de décision automatisée (articles 13(2)(f)) et article 14(2)(g)).

RGPD, réponse DAPC

- Les catégories de données à caractère personnel concernées [article 15(1)(b)] : la personne responsable du traitement doit fournir une copie des données à caractère personnel en cours de traitement. Elles peuvent inclure des données fournies par des tiers (article 14 du RGPD).
- Potentiellement : variables ou paramètres d'entrée utilisés pour la prise de décision et manière dont ils sont apparus (par exemple à l'aide d'informations statistiques) et leur pondération respective au niveau agrégé (opinion d'AG Pikamäe dans SCHUFA Holding, 2023, paragraphe 58).

Règlement européen sur l'IA, article 86 : demande d'explication

- Données d'entrée et principaux éléments de la décision.

Disponibilité publique

- Base de données de l'UE pour les systèmes d'IA à haut risque : section A systèmes d'IA à haut risque – une description des informations utilisées par le système (données, entrées) et de sa logique de fonctionnement.
- FAQ sur la plateforme.
- Communications publiques de l'entreprise.

Organisation intimée

De la part du fournisseur :

Règlement européen sur l'IA, article 11 et annexe IV : documentation technique

- Annexe IV(2)(b), la logique générale du système d'IA et des algorithmes, la justification et les hypothèses formulées, y compris en ce qui concerne les personnes ou groupes de personnes auxquels le système est destiné ; les principaux choix de classification ; ce que le système est conçu pour optimiser et la pertinence des différents paramètres ; la description de la production attendue et de la qualité de la production du système.
- Annexe IV(2)(d), description générale des jeux de données, informations sur leur provenance, leur portée et leurs principales caractéristiques.

De la part du déployeur :

Règlement européen sur l'IA, article 27 : analyse d'impact sur les droits fondamentaux

- Les catégories de personnes physiques et de groupes susceptibles d'être affectés par le système algorithmique (article 27(1)(c), du règlement sur l'IA).

Autres

- Clauses contractuelles.
- Recherche documentaire, jurisprudence existante et statistiques sur la surreprésentation et la sous-représentation dans le domaine pertinent.

Indicateurs de discrimination

- Utilisation directe des caractéristiques protégées dans le système.
- Utilisation de proxys qui sont inextricablement liés à un attribut protégé dans le système.
- Utilisation de proxys connus dans le système.
- Utilisation de caractéristiques protégées ou de proxy pour entraîner, tester ou valider le système.

Tout ce qui précède exige la preuve que ces entrées conduisent à une inégalité de traitement. Les indicateurs a et b révèlent une discrimination directe. Les indicateurs c et d révèlent une discrimination indirecte.

- Enregistrez l'utilisation directe des caractéristiques protégées. Où sont-elles utilisées et comment ?
- Consignez toute utilisation de caractéristiques protégées ou de proxys qui ont ou qui sont soupçonnés d'avoir un impact sur le résultat du système. Étudiez la conception technique du système et l'influence des proxys (suspectés) sur les résultats du système. Le fait qu'un proxy soit utilisé en tant qu'entrée ne suffit pas en soi comme preuve de discrimination. La prochaine étape consiste à évaluer les preuves pour comprendre si l'utilisation de ces entrées mène à un traitement inégal[51].
- Pour tout proxy, inclure des preuves ou des arguments expliquant pourquoi l'OPE les considère comme des proxys..

3.3. Analyse des risques pour les droits fondamentaux

51. Une entrée dans le système qui indique un motif protégé ou un proxy connu, et qui ne chevauche pas un motif protégé, peut ne pas suffire en soi à étayer une allégation de discrimination. Voir par exemple le cas STAR au Danemark. Dans ce cas, l'outil de prédiction a utilisé « l'ascendance » ou « l'origine » en tant que caractère prédictif. Le conseil a déterminé que le mot « ascendance » utilisé dans l'outil de profilage « n'indique rien au sujet de l'origine ethnique ». www.retsinformation.dk/eli/accn/W20220951825 (disponible en danois). Pour plus de détails, voir Equinet 2022: 23-24.

(analyses des risques, AIDH, AIDF et AIPD)

Les fournisseurs et les déployeurs de systèmes d'IA à haut risque ont des obligations diverses et variées pour identifier et documenter les risques de discrimination. Pour les systèmes algorithmiques qui ne présentent pas de risque élevé, les fournisseurs et/ou les déployeurs d'IA peuvent bien avoir identifié et documenté les risques de discrimination dans le cadre de leur conformité au RGPD (en particulier dans le cadre d'une AIPD) ou de la diligence raisonnable de leur entreprise sur les droits humains (Nations Unies, 2011). De nouvelles obligations peuvent également être introduites en raison de la Convention-cadre du Conseil de l'Europe sur l'IA.

Les risques de discrimination sont généralement identifiés dans une analyse d'impact sur les droits fondamentaux (AIDF), une analyse d'impact sur la protection des données (AIPD) et/ou d'autres procédures de gestion des risques. S'ils sont correctement réalisés, ces processus devraient identifier et consigner tout risque de discrimination présenté par un système d'IA ou le traitement de données à l'aide d'un système d'IA, ainsi que toute mesure d'atténuation prise et leur impact.

Dans les documents techniques et les évaluations des risques, des partis pris ou l'équité peuvent être mentionnés au lieu de l'inégalité ou de la discrimination.

À cette étape, des compétences techniques et sectorielles seraient bénéfiques.

Analyse d'impact sur les droits fondamentaux : règlement sur l'IAt

L'obligation d'effectuer une AIDF ne s'applique qu'à certains déployeurs de systèmes d'IA à haut risque (voir étape 2.2 pour plus de détails).

L'AIDF (et d'autres analyses) devrait inclure un éventail d'informations qui devrait mettre en évidence les lacunes potentielles et les atteintes aux droits, y compris le droit à la non-discrimination. Elle comprend :

- des informations sur la manière dont le déployeur utilisera le système d'IA à haut risque comme prévu par le fournisseur ;
- la durée et la fréquence d'utilisation du système ;
- comment les mesures de contrôle humain ont été mises en œuvre ;
- les catégories de personnes physiques et de groupes susceptibles d'être concernés ;
- les risques spécifiques de préjudices portant atteinte aux droits fondamentaux, y compris le droit à la non-discrimination, compte tenu des informations fournies par le fournisseur ;
- les mesures pour la gouvernance interne et les mécanismes de plainte que le déployeur utilisera si le risque se matérialise.

Si le déployeur n'a pas effectué une AIDF mais aurait dû le faire ou n'a pas inclus d'informations pertinentes, reportez-vous à l'étape 4.

AIPD

Une AIPD[52] devrait prendre en considération tous les risques pour les droits et libertés des personnes physiques résultant du traitement de données à caractère personnel, et pas seulement les risques pour la vie privée. Cela inclut également les risques de discrimination (article 29 du Groupe de travail sur la protection des données, 2017a). L'OPE peut avoir besoin de coopérer avec les autorités chargées de la protection des données.

Questions à envisager

Les informations sur les différentes analyses des risques effectuées par le fournisseur et le déployeur doivent être recoupées les unes avec les autres et toute divergence doit être notée.

- Le fournisseur ou le déployeur a-t-il identifié des risques pertinents pour les droits fondamentaux que l'autre n'a pas identifiés ?
- Lorsqu'une partie a effectué plus d'une analyse des risques ou plus d'un type d'analyse des risques, y a-t-il une cohérence en ce qui concerne les risques identifiés ? Des risques identifiés dans une analyse sont-ils absents dans une autre (à moins qu'il n'y ait une bonne explication, par exemple le risque ne survient que dans le cas d'un enfant et il est spécifiquement pris en compte dans l'analyse des risques pour les enfants) ?

Pour tous les systèmes algorithmiques, y compris l'IA à haut risque :

- La discrimination (ou un autre risque pour les droits fondamentaux) a-t-elle été identifiée par le fournisseur ou le déployeur ?
- Le fournisseur a-t-il mis en place un processus de gestion des risques ?
- Comment les procédures de gestion des risques et les évaluations d'impact ont-elles été utilisées pour déterminer les biais ou les inégalités dans les données ou le système ? Comment le biais est-il défini ou délimité ?
- Y a-t-il des risques non atténués ? Dans la mesure du possible, vérifiez les risques mentionnés dans la documentation technique par rapport au système de gestion des risques.

Remarque : lorsqu'une discrimination a été identifiée et que des mesures de gestion ont été mises en place, une enquête plus approfondie sera nécessaire (voir étape 3.6, tests et évaluation). Si le système présente un risque élevé et que l'OPE est une autorité relevant de l'article 77, il peut demander l'accès aux documents relatifs à l'évaluation des risques visés à l'article 9.

Dans les cas impliquant un système d'IA à haut risque :

- un processus de gestion des risques est-il conforme aux exigences de l'article 9 ?

52. Le Comité européen de la protection des données a mené une consultation en mai 2026 sur un modèle d'AIPD, disponible ici : www.edpb.europa.eu/public-consultations/template-for-data-protection-impact-assessment_fr.

Dans les cas impliquant un système d'IA à haut risque, lorsqu'une AIDF est requise de la part du déployeur :

- l'AIDF a-t-elle été effectuée correctement ? Il ne devrait pas s'agir d'un simple exercice de cases à cocher ;
- le déployeur a-t-il impliqué les parties prenantes concernées lors de la réalisation de l'analyse des risques ? (Le considérant 96 prévoit cette possibilité) ;
- les parties prenantes auraient-elles dû être consultées, par exemple les personnes susceptibles d'être affectées par le système ? (Voir affaire Bridges ci-dessous) ;
- si un processus de consultation a eu lieu, envisagez de communiquer avec les personnes consultées, mais sachez qu'elles peuvent être assujetties à une obligation de confidentialité ou à un accord de non-divulgation ;
- vérifiez que tous les risques de discrimination identifiés dans la documentation technique (article 11) sont mentionnés dans la documentation du système de gestion des risques (article 9). Si ce n'est pas le cas, une présomption de discrimination peut naître.

Sources potentielles d'information sur les risques et analyses des risques

Personne ayant déposé la plainte

RGPD, articles 13-15, informations fournies à la personne concernée

- L'existence d'une prise de décision automatisée et/ou d'un profilage (articles 13-15, 21 et 22 du RGPD).
- La logique sous-jacente dans l'ADM utilisée (articles 13(2)(f), 14(2)(g) et 15(1)(h) du RGPD).

Disponibilité publique

- Base de données de l'UE pour les systèmes d'IA à haut risque : section C déployeur : synthèse de l'AIPD et de l'AIDF (sauf mise en application de la loi ou contrôle aux frontières).

Organisation intimée

Du fournisseur :

Règlement européen sur l'IA, article 9, système de gestion des risques

- Les fournisseurs sont tenus de mettre en place un système de gestion des risques en ce qui concerne les systèmes d'IA à haut risque. Il s'agit d'un processus itératif continu qui exige :
  - l'identification des risques : l'identification et l'analyse des risques connus et raisonnablement prévisibles pour les droits fondamentaux (article 9(2)(a)) ;
  - la gestion des risques : l'adoption de mesures de gestion des risques appropriées et ciblées (article 9(2)(d)), y compris la transmission d'informations et la prestation de formations à destination des déployeurs en vue d'éliminer ou de réduire les risques (article 9(5)(c) et article 13) ;

- des tests : assurez-vous que le système a été testé afin d'identifier les mesures de gestion des risques les plus appropriées et les plus ciblées (article 9(6) ;
- des groupes particuliers : lors de la mise en œuvre du système, il convient de se demander si celui-ci est susceptible d'avoir un impact négatif sur les personnes de moins de 18 ans et, le cas échéant, sur d'autres groupes en situation de vulnérabilité (article 9(9)).

Cependant, la gestion du risque n'est requise que pour les risques qui « peuvent être raisonnablement atténués ou éliminés au fil du développement ou de la conception du système d'IA à haut risque, ou à la fourniture d'informations techniques adéquates ».

Règlement européen sur l'IA, article 10, données et gouvernance des données (voir étape 3.6 entraînement, tests et données de validation)

Règlement européen sur l'IA, article 11 et annexe IV, documentation technique

- L'annexe IV(2)(g) exige que la documentation technique contienne des informations sur les effets potentiellement discriminatoires.
- Annexe IV (3) : des informations détaillées sur la surveillance, le fonctionnement et le contrôle du système d'IA, en particulier en ce qui concerne ses capacités et ses limites de performance, y compris les degrés de précision pour des personnes ou groupes de personnes spécifiques à l'égard desquels le système est destiné à être utilisé et le niveau global de précision attendu par rapport à sa destination ; les résultats imprévus prévisibles et les sources de risques pour la santé et la sécurité, les droits fondamentaux et la discrimination eu égard à la destination du système d'IA ; les mesures de contrôle humain nécessaires conformément à l'article 14, y compris les mesures techniques mises en place pour faciliter l'interprétation des résultats des systèmes d'IA par les déployeurs ; les spécifications relatives aux données d'entrée, le cas échéant.

Du déployeur :

Règlement européen sur l'IA, article 13, transparence et fourniture d'informations aux déployeurs

- Instructions d'utilisation fournies au déployeur par le fournisseur (article 13(3)) :
  - les circonstances connues ou prévisibles de l'utilisation prévue ou d'une mauvaise utilisation susceptible de mettre en péril les droits fondamentaux ;
  - le cas échéant, ses performances vis-à-vis de personnes ou de groupes de personnes spécifiques à l'égard desquels le système est destiné à être utilisé ;
  - le cas échéant, les spécifications des données d'entrée ou toute autre information pertinente en ce qui concerne les jeux de données d'entraînement, de validation et de test utilisés, compte tenu de l'objectif prévu du système d'IA à haut risque.

Règlement européen sur l'IA, article 26(9), analyse d'impact sur la protection des données : réalisée à l'aide des informations fournies par le fournisseur/RGPD, article 35, analyse d'impact sur la protection des données (AIPD)

Règlement européen sur l'IA, article 27(1)(d), analyse d'impact sur les droits fondamentaux

- Les risques spécifiques de préjudices portant atteinte aux droits fondamentaux, y compris au droit à la non-discrimination, compte tenu des informations partagées par le fournisseur.

Autres

- Analyse réalisée en utilisant l'analyse des risques et des impacts des systèmes d'IA du Conseil de l'Europe du point de vue des droits humains, de la démocratie et de l'État de droit (méthodologie HUDERIA).
- Certaines organisations réaliseront et publieront volontairement une AIDH ou une AIDF, une AIPD ou une analyse de l'impact sur les droits de l'enfant. Une analyse « éthique » peut être réalisée.
- Toute autre analyse (du risque) partagée par l'organisation intimée.
- Demandes d'accès à l'information, le cas échéant.

Affaire clé : *R (Bridges) c. Chief Constable of South Wales Police*, 2020

(Les chiffres entre crochets […] désignent les numéros de paragraphe du jugement.)

L'affaire Bridges, jugée par la Cour d'appel du Royaume-Uni, portait sur la question de savoir si une autorité publique avait violé les droits humains ou avait manqué à son obligation d'évaluer les incidences sur l'égalité en vertu de la loi de 2010 sur l'égalité nationale avant de mettre en œuvre la reconnaissance faciale en temps réel. L'obligation d'égalité dans le secteur public impose à une autorité publique de tenir dûment compte de la nécessité d'éliminer la discrimination, le harcèlement, la victimisation et d'autres comportements répréhensibles, de promouvoir l'égalité des chances et de favoriser de bonnes relations entre les personnes qui partagent une caractéristique pertinente et celles qui n'en partagent pas.

La Cour a jugé que cette obligation :

- doit être respectée avant et au moment où une politique (d'utilisation de la reconnaissance faciale) est envisagée ;
- doit être mise en œuvre de manière effective, avec rigueur et ouverture d'esprit – il ne s'agit pas d'un simple exercice de cases à cocher ;
- ne peut être déléguée, ce qui peut créer de réelles difficultés pour les autorités publiques qui déploient des systèmes fournis par des prestataires tiers, en particulier si les informations fournies sur les performances d'un système sont insuffisantes ou si l'autorité ne dispose pas d'une expertise ou d'une expérience suffisantes pour comprendre l'information ou le système [199] ;

En outre, si des documents pertinents pour l'évaluation ne sont pas disponibles, leur acquisition est obligatoire – ce qui peut nécessiter des consultations supplémentaires avec les groupes concernés. Enfin, à condition qu'un examen rigoureux de l'obligation ait été effectué afin de pouvoir apprécier convenablement l'incidence potentielle de

la décision sur les objectifs en matière d'égalité et l'opportunité de les promouvoir, il appartient à la personne détenant l'autorité de décision de juger du poids à accorder aux facteurs qui éclairent la décision.

Dans l'ensemble, la Cour a statué que le respect de l'obligation « exige des mesures raisonnables pour se renseigner sur ce qui n'est peut-être pas encore connu d'une autorité publique au sujet de l'impact potentiel d'une décision ou d'une politique proposée sur les personnes ayant les caractéristiques pertinentes, en particulier aux fins actuelles de la "race" et du sexe » [181].

Le seul fait d'intégrer un être humain dans le processus décisionnel a été jugé insuffisant pour satisfaire à cette obligation légale, dès lors que celle-ci régit le processus et non le contenu de la décision.

La Cour a reconnu que les êtres humains commettent des erreurs [185].

Des observations semblables sont susceptibles de s'appliquer aux évaluations des risques prévues par le règlement sur l'IA.

Indicateurs de discrimination

Un risque de discrimination qui a été identifié mais qui n'a pas été atténué, suivi ou justifié devrait être considéré comme un indicateur fort.

- Référence directe à la discrimination, aux biais ou à l'injustice dans l'AIDF.
- Référence directe à la discrimination, aux biais ou à l'injustice dans l'AIPD.
- Dans le cas où le biais est mentionné comme étant un risque, la signification et les types de biais ne sont pas définis ou ils sont définis de façon inappropriée dans la documentation.
- Aucune procédure de gestion des risques n'est en place ou la procédure est inadéquate pour les risques identifiés de discrimination, de biais ou d'injustice.
- Le risque non atténué de discrimination est clairement indiqué dans la documentation.
- Tests insuffisants.
- Référence directe à un risque de violation des droits fondamentaux dans l'AIDF ou l'AIPD :
  - lorsqu'il existe un risque de violation des droits fondamentaux, cela devrait signaler la nécessité d'une enquête plus approfondie afin de garantir que le système est non discriminatoire.
- Une AIDF ou une AIPD incomplète, superficielle ou inexacte justifierait une enquête plus approfondie.
- Le fait de ne pas consulter les parties prenantes concernées pour identifier les risques lorsqu'elles auraient clairement dû être inclus (par exemple, un système conçu pour prendre des décisions concernant les personnes handicapées qui n'a consulté aucune personne handicapée ou organisation représentative avant le déploiement peut ne pas avoir réalisé l'AIDF avec le soin requis).

- Consignez tout risque de discrimination identifié par le fournisseur ou le déployeur et toute mesure identifiée pour atténuer ce risque.
- Examinez et notez s'il y a des défaillances claires et évidentes dans l'évaluation des risques, l'AIPD ou l'AIDF.

3.4. Quel est l'objectif du système algorithmique et comment l'atteint-il ?

Pour comprendre si une discrimination algorithmique a eu lieu, l'OPE devrait comprendre le système lui-même, au moins en termes généraux. Comprendre l'objectif et la conception du système et son fonctionnement aide à évaluer les éléments suivants.

- Quel est l'objectif du système ?
  - L'objectif est-il lié à un domaine où une discrimination historique a eu lieu à l'égard de la ou des caractéristiques protégées ?
  - Les données utilisées sont-elles appropriées à l'objectif visé (voir étape 3.5) ?
- Comment les entrées contribuent aux résultats ?
  - Une entrée identifiée à l'étape 3.2 influence-t-elle le résultat et conduit-elle à une différence potentielle de traitement ?
- Si le système algorithmique influe soit sur le processus décisionnel dans son ensemble, soit sur l'impact final pour le groupe concerné :
  - quelle est l'ampleur du rôle du système ? Par exemple, est-il le seul implique dans la prise de décision ou aide-t-il à la prise de décision ?
  - le système algorithmique est-il l'ensemble du système ou juste une partie d'un système beaucoup plus grand qui peut impliquer d'autres systèmes algorithmiques ?

Dans les informations sur l'utilisation prévue et la conception, vous devriez généralement chercher à voir si les choix de conception semblent appropriés pour l'utilisation prévue et l'utilisation réelle. À cette étape, l'expertise technique et l'expertise sectorielle sont utiles.

Les informations dont disposent les OPE devraient en fin de compte permettre à ceux-ci de comprendre ce que fait le système d'IA, comment il le fait grâce à des informations utiles sur la logique sous-jacente, l'importance du traitement et les conséquences envisagées pour l'individu et les raisons pour lesquelles divers choix en matière de conception et de données ont été faits. Lorsque le système algorithmique est considéré comme une ADM admissible ou un système d'IA à haut risque utilisé pour une prise de décision efficace, la disponibilité de cette information est requise. Toutefois, si la documentation ne permet pas cette compréhension, il peut y avoir des problèmes liés au manque de transparence (voir ci-dessous, transparence, à l'étape 3.7).

La conception du système peut également être particulièrement importante plus tard lorsque vous examinerez les moyens par lesquels l'organisation intimée peut

chercher à réfuter toute présomption de discrimination ou à justifier les décisions prises.

Questions à envisager

Objectif et utilisation

- Comment le système algorithmique est-il censé être utilisé ?
  - Quel est l'objectif poursuivi ?
  - Comment est-il utilisé dans la pratique ?
  - Dans les cas où le système est fourni par un fournisseur : est-il utilisé par le déployeur comme prévu par le fournisseur ?
- Pour qui le système est-il conçu ?
  - Quels ont été les choix de conception (voir l'encadré ci-dessous) faits à l'égard des personnes ou groupes de personnes pour lesquels le système est destiné à être utilisé ?
- Des hypothèses ont-elles été formulées au sujet des personnes ou des groupes de personnes pour lesquels le système est destiné à être utilisé ? Par exemple, on s'attend à ce que ce système ne soit utilisé que par des adultes, ou à ce qu'il soit utilisé dans un secteur particulier ou pour une population particulière. Ces hypothèses sont-elles rationnelles et justifiées ?
- Comment le système influence-t-il le traitement des personnes ?
- Dans les cas où il y a un être humain responsable de faire des choix qui influencent le traitement des personnes (Human-in-the-Loop) :
  - existe-t-il des lignes directrices ou des protocoles destinés aux personnes responsables de la décision pour les aider à comprendre comment utiliser le système d'IA ou interpréter les résultats ?
  - la personne concernée a-t-elle la compétence, la formation et l'autorité nécessaires pour outrepasser une décision prise par un système d'IA ?
  - la personne concernée a-t-elle été formée pour comprendre et interpréter les résultats du système et reconnaître les cas où des biais peuvent survenir et où une intervention est nécessaire ?

Remarque : la simple existence d'un être humain dans la boucle ne signifie pas que le système n'influence pas le traitement des personnes, puisque l'être humain dans la boucle est plus ou moins influencé par les résultats du système.

Fonctionnement du système

- Quel est l'objectif d'optimisation du système ?
  - Le système a-t-il été conçu pour optimiser quelque chose de (légèrement) différent de son cas d'utilisation prévu ? Par exemple, a-t-il été optimisé pour prédire les erreurs, mais est-il maintenant utilisé pour prédire la fraude ?
- Le système utilise-t-il des caractéristiques de groupe pour produire des résultats sur un individu (profilage) ?

- Comment le système gère-t-il les écarts ?
  - Comment le système gère-t-il les écarts par rapport au comportement attendu (erreur ou activité inhabituelle, par exemple) ?
  - Tous les écarts sont-ils traités de la même manière, quelle que soit la raison sous-jacente ?
- Le système a-t-il été conçu dans un souci d'équité et d'égalité ?
  - L'équité ou la non-discrimination faisaient-elles partie de la conception
    - La notion d'équité propre au contexte est-elle explicitement définie ?
    - Quelles mesures ont été prises pour favoriser l'équité et l'égalité ?

Influence des entrées

- Comment le choix d'un modèle ou d'un algorithme (pré-formé) a-t-il été fait ?
  - Cette procédure est-elle documentée et disponible pour les déploiements du système ?
- Quels sont les critères qui influencent le résultat du système ?
  - Les motifs protégés pertinents et/ou les variables proxys dans les données sont-ils traités ou utilisés comme base pour une ADM ou un profilage (comme les entrées pour créer un profil, les informations sur le profil) ?
  - Existe-t-il des indicateurs sur le poids/l'influence d'entrées spécifiques sur les sorties ?
- Quels sont les critères qui ont le plus de poids dans le système ou la décision
  - L'organisation intimée pourrait-elle soutenir que les motifs protégés et/ou les proxys ont un impact négligeable sur le résultat ?
- Quel est l'impact cumulatif de plusieurs variables indirectes prises ensemble, telles que des facteurs tels que code postal + éducation ?
- Les données utilisées ou auxquelles on accorde le plus de poids sont-elles arbitraires ou non pertinentes au regard de la finalité à laquelle elles sont utilisées ?

Sources potentielles d'information pour comprendre l'objectif, la conception et le fonctionnement du système

Personne ayant déposé la plainte

RGPD, articles 13-15, informations fournies à la personne concernée

- L'existence d'une prise de décision automatisée et/ou d'un profilage (articles 13-15, 21 et 22 du RGPD).
- La logique sous-jacente dans l'ADM utilisée (articles 13(2)(f), 14(2)(g) et 15(1)(h) du RGPD).

RGPD, article 15(3) : DAPC

- La portée, la nature et le niveau de détail des informations qui seront fournies en réponse à une demande de DAPC peuvent varier. Dans le cas d'une ADM

admissible, il est nécessaire d'inclure les éléments suivants, mais pour d'autres (parties de) systèmes algorithmiques, ces informations peuvent également être incluses :

- les variables ou paramètres d'entrée utilisés pour la prise de décision et leur origine (par exemple, l'utilisation d'informations statistiques) et leur poids respectif (avis de AG Pikamäe SCHUFA Holding, 2023, paragraphe 58) ;
- l'effet d'une variable d'entrée sur la note ; l'explication de la raison pour laquelle la personne concernée s'est vu attribuer une note particulière ; la liste des catégories de profils possibles et des scénarios contrefactuels ;
- le principe de calcul sous-jacent : explications détaillées sur la méthode de calcul de la valeur de la note ; affectation de la personne concernée à un résultat d'évaluation, liste des catégories de profils (Barros Vale et Zanfir-Fortuna 2022) ;
- l'existence de la prise de décision automatisée et/ou du profilage et des informations utiles sur la logique sous-jacente, ainsi que sur la signification et les conséquences envisagées de ce traitement pour la personne concernée (article 15(1)(h)).

Règlement européen sur l'IA, article 86, règlement sur l'IA : droit à une explication

- Le rôle du système d'IA dans le processus décisionnel. Par exemple : « Le système d'IA [nom/désignation du système d'IA] a été utilisé pour [description de la fonction spécifique du système d'IA, comme l'analyse des données, l'évaluation des critères, l'analyse vocale pour la reconnaissance des émotions, etc.]. Le système d'IA a contribué à la procédure de prise de décision en [analysant les données et en fournissant une évaluation des paramètres pertinents, par exemple] ».
- La logique de traitement : « Le système a analysé les données d'entrée à l'aide de [description des algorithmes ou des modèles, comme les arbres de décision] pour effectuer [description de l'objectif, comme l'évaluation des risques] ».
- Résultat : « Sur le fondement de cette analyse, le système d'IA a généré les résultats suivants, qui ont été intégrés dans la décision finale : [Description du résultat, un score, une classification, par exemple] ».
- Impact de la décision prise sur la personne affectée : « La décision a l'impact potentiel suivant sur vous : [impact sur la solvabilité, accès à certains services, impact financier, etc.]. Ces impacts sont pertinents car ils pourraient affecter considérablement votre capacité à [expliquer les conséquences, telles que la participation à des programmes ou la réception d'offres de financement] ».
- Pondération des critères : « Les principaux critères d'influence ont été pondérés comme suit : [brève explication de l'importance de chaque critère] ».

Organisation intimée

De la part du fournisseur

Règlement européen sur l'IA, section A(5) et (6) de l'annexe VIII relative aux systèmes d'IA à haut risque

- Description de la destination du système d'IA, description de base et concise des informations utilisées par le système (données, entrées) et de sa logique de fonctionnement.
- Pour ce qui est du contrôle de la mise en application de la loi et de la gestion des migrations, des informations supplémentaires peuvent être mises à la disposition de l'ASM concernée dans la base de données

Règlement européen sur l'IA, article 11 et annexe IV, documentation technique

- (1)(a) La destination prévue.
- (1)(d) La description de toutes les formes sous lesquelles le système d'IA est mis sur le marché ou mis en service (comme les progiciels intégrés au matériel, les téléchargements ou les API).
- (1)(h) Description de base de l'interface utilisateur fournie au déployeur.
- (2)b) Spécifications de conception et logique générale. Voir ci-dessus, 3.2 : Existe-t-il des caractéristiques protégées ou des proxy connus dans les entrées ou les jeux de données ?

De la part du déployeur

Article 13, Transparence et fourniture d'informations aux déployeurs

- Le fournisseur doit fournir (article 13(2)) des instructions d'utilisation pour le déployeur.
- Article 13(3)(b)(iv), le cas échéant, les capacités et caractéristiques techniques du système d'IA à haut risque pour fournir des informations utiles pour expliquer ses résultats, et (3)(b)(v), ses performances concernant des personnes ou groupes de personnes spécifiques sur lesquels le système est destiné à être utilisé.
- L'article 13(3)(c), les modifications apportées au système d'IA à haut risque et à ses performances qui ont été prédéterminées par le fournisseur au moment de l'évaluation initiale de la conformité, le cas échéant.

Autres

- Recherche documentaires
- Clauses contractuelles
- FAQ sur la plateforme
- Communications publiques de l'entreprise

Indicateurs de discrimination

Indicateurs forts

Ensemble, ces indicateurs pourraient étayer une preuve prima facie.

- Les critères suspectés (identifiés à l'étape 3.2) ont une influence non négligeable sur le résultat du système.

  - Lorsqu'une caractéristique protégée ou un proxy inextricablement lié à une caractéristique protégée a une influence non négligeable sur le résultat du système, cela indique clairement une discrimination directe.
  - Lorsqu'un proxy (suspecté) a une influence non négligeable sur le résultat du système, cela indique une discrimination indirecte.
- La sortie du système détermine (au moins en partie) le traitement des individus.

Autres indicateurs

- L'utilisation prévue, les hypothèses et/ou les choix de conception sont soupçonnés de refléter ou d'amplifier les inégalités historiques, conduisant à un traitement inégal.
- Les choix de conception favorisent un groupe de personnes plutôt qu'un autre.
- Une optimisation du système pour les groupes historiquement avantagés.
- Le système n'est pas utilisé aux fins prévues pour lesquelles il a été développé/fourni.

- Consignez l'objectif du système et ce qu'il optimise.
- Enregistrez les proxys et les caractéristiques protégées qui influencent les résultats du système et, le cas échéant, les mesures de leur poids/influence sur le résultat.

3.5. Données d'entraînement, de test et de validation

Des données sont utilisées pour développer et tester les systèmes algorithmiques. Même les systèmes algorithmiques fondés sur des règles qui sont conçus manuellement sont validés sur la base de données. Les données sont toujours une représentation imparfaite de la réalité. Les biais et les inégalités existantes sont reflétés dans les données. L'évaluation des biais et des inégalités dans les données peut être utilisée pour soutenir que le système qui en résulte conduit à des résultats discriminatoires. Notez que la compréhension des biais dans les données ne montrera pas en soi un traitement inégal. Cependant, avec d'autres sources d'analyses complémentaires, il contribue à établir une preuve prima facie et peut être utilisé pour motiver des demandes de tests supplémentaires.

Lorsque des données statistiques sur l'égalité sont disponibles, cette section est également pertinente pour examiner la validité et la fiabilité des données utilisées pour calculer ces statistiques. Dans ce cas, cette section doit être lue conjointement avec la section 3.6.

Dans les données, vous devriez généralement chercher à voir quelles informations l'organisation intimée a partagées sur la façon dont elle a entraîné et testé son système. Sur cette base, vous pourrez peut-être identifier les principaux problèmes de discrimination ; par exemple, aucune femme n'a été incluse dans le jeu de données, les données d'entraînement provenaient d'une ville, mais le système est utilisé dans une zone rurale, etc.

Dans d'autres cas, vous devrez peut-être faire appel à une expertise technique et sectorielle pour déterminer si le système est formé et testé sur des jeux de données représentatifs, pertinents, mesurés de manière précise, appropriée et généralisable.

Parmi les questions clés à prendre en considération figurent une variété de biais. Ces biais ne sont pas indépendants les uns des autres : un biais en influencera un autre.

- S'il est probable qu'il y ait un biais historique dans les données. Dans tous les domaines où il existe une inégalité historique ou actuelle, cette inégalité se reflète dans les données, à moins que cela ne soit corrigé avec soin. On peut citer par exemple la surveillance policière excessive des communautés noires et les périodes d'interruption d'activité liée à la garde d'enfants ou à d'autres responsabilités de prise en charge, qui ont un impact sur l'égalité entre les femmes et les hommes.
- S'il y a un biais de représentation dans les données. Par exemple, la population cible reflète-t-elle la population d'utilisation ? Par exemple, les données qui sont représentatives de Lisbonne peuvent ne pas être représentatives de Porto. Les données représentatives d'Helsinki il y a 30 ans ne reflètent peut-être pas la population d'aujourd'hui. D'autres problèmes peuvent découler de la composition de l'échantillon et de la pondération des données. Par exemple, dans les données sur la santé, il pourrait y avoir une sous-représentation de femmes et/ou une surreprésentation des femmes enceintes. Chacun de ces éléments pourrait fausser les données et avoir un impact sur les résultats. En outre, faites attention à la façon dont les minorités sont représentées dans les données. Par exemple, lorsque certains accents n'ont pas été inclus dans un système de reconnaissance vocale, ce système fonctionnera moins bien pour ces accents. Si les personnes porteuses de hijab n'ont pas été incluses dans les données pour la reconnaissance faciale, il est possible que les personnes porteuses de hijab ne soient pas bien reconnues.
- S'il y a un biais de sélection dans les données. C'est le cas lorsqu'il n'y a que des données pour un sous-groupe spécifique de données. Par exemple, dans l'acceptation de prêt, il n'y a que des informations sur les personnes en défaut de paiement de leur prêt pour le groupe qui a reçu le prêt dans le passé. Les données relatives aux patrouilles de police ne comprennent que des informations sur les zones où la police a patrouillé. Les données sur les fraudes peuvent inclure uniquement le groupe qui a fait l'objet d'une enquête pour fraude dans le passé. Les données sur le recrutement contiennent moins d'informations sur les candidats et candidates dont les CV n'ont jamais été sélectionnés pour d'autres étapes. Cela pourrait fausser les données, en particulier si les personnes ayant une caractéristique protégée spécifique étaient historiquement surreprésentées ou sous-représentées dans ce contexte.
- Présence éventuelle d'un biais de mesure dans les données. C'est le cas lorsque les étiquettes (la mesure) ne reflètent pas le concept qu'elles sont censées mesurer. Par exemple, si vous voulez mesurer la fraude dans les prestations sociales, mais que vous décidez que toutes les allocations incorrectes de prestations sont pertinentes à considérer, le résultat sera que vous incluez l'erreur innocente ainsi que la fraude. Cela peut avoir des répercussions disparates sur certains groupes, par exemple ceux qui ont un faible taux d'alphabétisation, un faible niveau d'instruction et ceux qui ne parlent pas une langue nationale comme langue maternelle. Autre forme de biais de mesure : lorsque les biais humains (ou la discrimination humaine)

influencent les étiquettes enregistrées qui sont maintenant considérées comme la vérité fondamentale. Dans toute situation où les étiquettes de vérité terrain sont le résultat d'une évaluation humaine avec un certain degré de discrétion individuelle, il y a un risque que les biais humains entraînent un biais de mesure. Les évaluations des enseignants et enseignantes, les données de recrutement ou les enquêtes sur les fraudes en sont des exemples.

Questions à envisager

Examinez les questions ci-dessous en gardant à l'esprit les biais et les inégalités. Toutes les questions ci-dessous sont subjectives, nuancées et spécifiques au contexte. Pour y répondre, il faudra bien comprendre le secteur et les inégalités relatives existantes, ainsi que les systèmes et la technologie utilisés. Le résultat de cette évaluation fournira principalement des orientations pour une évaluation plus approfondie ou peut être utilisé pour établir une preuve prima facie.

Notez que dans le cas de systèmes algorithmiques conçus manuellement, il n'y aura pas de données d'entraînement. Dans certains systèmes d'IA, les données d'entraînement n'incluent pas d'étiquettes. Dans ce cas, il y a toujours une attente de données de test, qui incluent une forme ou une autre d'étiquettes pour vérifier si le modèle fournit des résultats corrects ou acceptables.

Données d'entraînement, de test et de validation

- Comment les jeux de données d'entraînement, de test et de validation ont-ils été choisis ?
- Quelles données d'entraînement/test/validation ont été recueillies ?
- Quelles sont les sources de ces données ?
- Dans quel but les données ont-elles été collectées à l'origine ? Y a-t-il des problèmes potentiels découlant de leur utilisation dans un contexte différent ? Par exemple, si des données ont été recueillies aux fins d'évaluation de la santé, leur réutilisation à des fins d'assurance soulève-t-elle de potentiels problèmes de biais de mesure ?
- Quelles méthodes ont été utilisées pour recueillir les données ?
  - Ces méthodes semblent-elles raisonnables, en ce sens qu'elles donnent raisonnablement lieu à des données fiables et pertinentes pour la tâche prévue du système algorithmique ?
  - A-t-on veillé à ce que les données soient représentatives et mesurées de manière appropriée ? Peut-on raisonnablement s'attendre à ce que les données se généralisent au domaine et au contexte dans lesquels les systèmes algorithmiques sont maintenant appliqués ?
  - Des données supplémentaires ont-elles été recueillies (ou étiquetées) pour être utilisées en tant que données d'entraînement/test/validation ? Un échantillon aléatoire a-t-il été utilisé pour recueillir des données ?
- Lorsque le jeu de données provient d'un tiers, comment l'a-t-il obtenu ? Les données fournies par des tiers ont-elles été vérifiées ?

- Y a-t-il des inégalités historiques susceptibles de se refléter dans les données (voir la section sur les biais ci-dessus) ? Considérez également la discrimination multiple/intersectionnelle ici.
- Les minorités sont-elles suffisamment reflétées dans les données ? C'est particulièrement pertinent pour les systèmes algorithmiques qui sont appliqués dans des secteurs où la culture, le contexte et l'apparence physique peuvent influencer les résultats.

Étiquettes

- Dans quelle mesure les étiquettes sont-elles fiables ?
- Les étiquettes correspondent-elles à l'objectif des systèmes ? En d'autres termes, les étiquettes mesurent-elles ce que le système est censé prédire ?
- La justesse peut-elle être déterminée objectivement ? Cela a-t-il été déterminé ?
- Dans les cas où il y a un jugement humain impliqué dans l'étiquetage, les biais humains auraient-ils faussé les étiquettes ? Y a-t-il des raisons de soupçonner une discrimination ?

Mesures d'atténuation existantes

- Comment la qualité des données a-t-elle été évaluée ?
  - Quelles mesures ont été prises pour résoudre les problèmes de qualité découverts, tels que compléter ou supprimer des données ?
- Quelles mesures ont été prises pour que les données soient représentatives, pertinentes, mesurées avec précision et généralisables ?
- Comment les biais potentiels et les inégalités dans le jeu de données ont-ils été atténués ?
  - Des mesures visant à atténuer les biais[53] ont-elles été mises en œuvre ?
  - Pour quels biais ont-elles été mises en œuvre ? Existe-t-il des biais reconnus pour lesquels aucune mesure d'atténuation n'a été mise en œuvre ?
  - Quelles techniques ont été utilisées ?
  - Existe-t-il des informations sur l'efficacité de ces mesures d'atténuation des biais ?

Sources potentielles d'information pour évaluer les données utilisées

Personne ayant déposé la plainte

Règlement européen sur l'IA, article 86, droit à une explication

RGPD, articles 13-15, informations fournies à la personne concernée

53 Les mesures d'atténuation de biais pourraient être toutes les mesures prises pour prévenir les biais dans les données ou dans un système. Il pourrait s'agir d'un échantillonnage aléatoire, d'une collecte de données supplémentaires pour accroître la représentativité des données, de protocoles stricts pour l'évaluation des cas, d'un contrôle par deux personnes différentes lors de l'évaluation des cas, d'une nouvelle pondération des données et de nombreuses autres mesures.

Disponibilité publique

Base de données de l'UE pour les systèmes d'IA à haut risque

- Section A, systèmes d'IA à haut risque – informations utilisées par le système, y compris les données et les entrées.
- Section C, déployeur – synthèse de l'AIPD et de l'AIDF (sauf si elles sont utilisées pour la mise en application de la loi ou la gestion des migrations).

Organisation intimée

De la part du fournisseur :

Règlement européen sur l'IA, annexe IV(2)(d), documentation technique

- « Le cas échéant, les exigences relatives aux données en ce qui concerne les fiches décrivant les méthodes et techniques d'entraînement et les jeux de données d'entraînement utilisés, y compris une description générale de ces jeux de données et des informations sur leur provenance, leur portée et leurs principales caractéristiques; la manière dont les données ont été obtenues et sélectionnées; les procédures d'étiquetage (par exemple pour l'apprentissage supervisé), les méthodes de nettoyage des données (par exemple la détection des valeurs aberrantes). »

Règlement européen sur l'IA, article 10, données et gouvernance des données

- Les fournisseurs sont tenus d'avoir mis en œuvre des pratiques appropriées afin de garantir que les jeux de données soient « pertinents, suffisamment représentatifs et, dans toute la mesure possible, exempts d'erreurs et complets au regard de la destination. Ils possèdent les propriétés statistiques appropriées, y compris, le cas échéant, en ce qui concerne les personnes ou groupes de personnes à l'égard desquels le système d'IA à haut risque est destiné à être utilisé. Ces caractéristiques des jeux de données peuvent être remplies au niveau des jeux de données pris individuellement ou d'une combinaison de ceux-ci ».

Ceci couvre au moins :

- des choix de conception pertinents ;
- la collecte et utilisation des données (article 10(2)(b), (e) et (g)) ;
- les opérations de traitement des données, y compris l'annotation et l'étiquetage (article 10(2)(c)) ;
- la formulation d'hypothèses, notamment en ce qui concerne l'évaluation et la représentation (article 10(2)(d)) ;
- en tenant explicitement compte des biais, ils sont tenus d'examiner et de prendre « des mesures appropriées pour détecter, prévenir et atténuer les éventuels biais » repérés dans les jeux de données qui sont susceptibles « d'avoir une incidence négative sur les droits fondamentaux ou de se traduire par une discrimination interdite par le droit de l'Union ».

Règlement européen sur l'IA, article 17(1)(f), gestion de la qualité

- Les politiques, procédures et instructions écrites doivent comprendre : « les systèmes et procédures de gestion des données, notamment l'acquisition, la collecte, l'analyse, l'étiquetage, le stockage, la filtration, l'exploration, l'agrégation, la conservation des données et toute autre opération concernant les données qui est effectuée avant la mise sur le marché ou la mise en service de systèmes d'IA à haut risque et aux fins de celles-ci ».

De la part du déployeur :

Règlement européen sur l'IA, article 13(3.b)(vi), instructions d'utilisation

- Le fournisseur devrait fournir « le cas échéant, les spécifications relatives aux données d'entrée, ou toute autre information pertinente concernant les jeux de données d'entraînement, de validation et de test utilisés, compte tenu de la destination du système d'IA à haut risque » (dans les informations fournies au déployeur par le fournisseur).

Règlement européen sur l'IA, Article 26(4)

- Dans la mesure où le déployeur exerce un contrôle sur les données d'entrée, il veille à ce que les données d'entrée soient pertinentes et suffisamment représentatives au regard de la destination du système d'IA à haut risque.

Autres

- Initiatives de documentation
- Fiches techniques pour les jeux de données
- Modèles de cartes pour modèles de rapports
- Gouvernance des données appliquée et normes de qualité – lorsque l'organisation intimée mentionne l'utilisation de normes, ces normes peuvent être utilisées pour comprendre ce à quoi l'organisation intimée s'est engagée.
- Pour ce qui est du contrôle de la mise en application de la loi et de la gestion des migrations, des informations supplémentaires peuvent être mises à la disposition de l'ASM concernée dans la base de données.

Indicateurs de discrimination

Les indicateurs ci-dessous montrent des problèmes avec les données, qui peuvent contribuer à une preuve prima facie en indiquant généralement que des données non fiables ou biaisées sont utilisées et peuvent être utilisées pour motiver des demandes de tests supplémentaires. Il est probable que cette liste soit incomplète. D'autres considérations peuvent découler de ce qui précède et ne pas être énumérées ici.

- Problèmes généraux liés aux données :
  - non représentatives (zone géographique, secteur) ;
  - source inappropriée (réutilisée) ;
  - pas à jour ;
  - biais de sélection ;

- biais de mesure dans le sens où l'étiquette reflète un concept différent de ce qui est prévu ou les étiquettes sont généralement peu fiables.

- Indicateurs de problèmes de qualité des données liés à la discrimination :
  - les données ne sont pas représentatives de toutes les personnes affectées ;
  - forte probabilité de refléter des inégalités historiques ou persistantes ;
  - le biais de sélection est susceptible d'entraîner une surreprésentation ou une sous-représentation de groupes ayant des motifs protégés ;
  - le biais de mesure est susceptible d'affecter des groupes ayant des motifs protégés ;
  - la source, la création ou l'étiquetage des données comporte un indice de discrimination.
- Les décisions relatives aux données ne sont pas suffisamment documentées.
- Aucune mesure d'atténuation ou l'efficacité des mesures d'atténuation ne peut pas être démontrée.

Indicateurs de non-conformité avec les systèmes à haut risque du règlement sur l'IA

Bien que la non-conformité au règlement sur l'IA ne soit pas au centre de cette méthodologie d'évaluation, toute indication que les exigences du règlement ne sont pas respectées peut apporter un éclairage quant aux mesures correctives (voir étape 4).

- Le fournisseur ne peut pas démontrer qu'il a mis en œuvre des pratiques garantissant que les jeux de données sont « pertinents, suffisamment représentatifs et, dans toute la mesure possible, exempts d'erreurs et complets au regard de la destination. Ils possèdent les propriétés statistiques appropriées, y compris, le cas échéant, en ce qui concerne les personnes ou groupes de personnes à l'égard desquels le système d'IA à haut risque est destiné à être utilisé. Ces caractéristiques des jeux de données peuvent être remplies au niveau des jeux de données pris individuellement ou d'une combinaison de ceux-ci ». Ou bien ils ne peuvent fournir de documentation à ce sujet.
- Le fournisseur n'a pas effectué de test de biais ou n'a pas testé tous les biais pertinents.
- Le fournisseur n'a pas mis en œuvre de mesures efficaces pour prévenir et atténuer les biais qu'il a détectés ou soupçonnés.

- Consignez tout problème général de données, tout problème lié à la qualité des données, tout problème de documentation insuffisante et toute préoccupation concernant les mesures d'atténuation.
- Consignez tout problème de données lié à l'inégalité et à la discrimination, toute préoccupation concernant les mesures d'atténuation et, le cas échéant, un argument quant à la manière dont ces problèmes pourraient conduire à une inégalité de traitement dans l'affaire en cours.
- Consignez tout indicateur de non-conformité aux dispositions du règlement sur l'IA relatives aux données.

3.6. Tests et évaluation

Les résultats des tests et de l'évaluation peuvent fournir des informations précieuses pour déterminer s'il y a eu discrimination. La question centrale ici est de savoir si le système algorithmique fournit des résultats inégaux pour différents groupes d'individus. Il est possible de le démontrer, par exemple, à travers les statistiques sur les résultats du système.

Lorsque des données statistiques sur l'égalité sont disponibles, la validité et la fiabilité de ces données doivent être évaluées (voir ci-dessous). Il est donc pertinent d'examiner la validité et la fiabilité des données utilisées pour calculer ces statistiques. Par conséquent, cette section doit être lue conjointement avec la section 3.5.

À cette étape, une expertise technique et sectorielle peut être utile.

Afin de tester les résultats pour différents groupes ayant des caractéristiques protégées différentes, il faut disposer de données sur ces caractéristiques protégées. Par exemple, pour tester si un système produit des résultats différents pour les hommes et les femmes, il faut savoir pour chacun des sujets de données de test qui est un homme et qui est une femme. Ces données peuvent parfois manquer. Dans ce cas, les tests doivent se concentrer sur les caractéristiques disponibles, y compris celles qui sont ou pourraient être liées aux motifs protégés. Par exemple, il se peut que des données sur l'origine ethnique ne soient pas disponibles mais que des données sur la nationalité le soient, ou pour les images faciales, il se peut qu'il n'y ait pas de données disponibles sur les personnes racialisées, mais la valeur chromatique des images faciales pourrait être utilisée.

Acquisition de données statistiques (supplémentaires)

Lorsque les données statistiques ne sont pas disponibles ou sont insuffisantes pour évaluer la discrimination algorithmique, certaines options peuvent aider à acquérir des données statistiques (supplémentaires).

Évaluation et tests par l'ASM

Lorsqu'un OPE n'est pas en mesure de déterminer s'il y a eu violation des droits en vertu du droit de l'Union, l'article 77(3) du règlement sur l'IA lui permet de présenter une « demande motivée à l'ASM d'organiser des tests du système par des moyens techniques ». L'OPE doit collaborer étroitement avec les tests.

Le considérant 157 précise que la procédure prévue à l'article 77 devrait être appliquée aux systèmes d'IA à haut risque présentant un risque pour les droits fondamentaux, aux systèmes interdits mis sur le marché (article 6), mis en service ou utilisés en violation des pratiques interdites dans la loi (article 5) et aux systèmes d'IA à haut risque ou non, mis à disposition en violation des exigences de transparence du règlement et présentant un risque (articles 6(4) et 50).

Tests par l'OPE ou en collaboration avec des tiers

Les organismes de promotion de l'égalité peuvent créer leurs propres données sur l'égalité dans l'exercice de leurs mandats de supervision et autres en vertu

des directives et règlements des organismes chargés de l'égalité ou autrement. Toutefois, la prudence peut être de mise lors de l'utilisation de ces données, car il existe un risque qu'elles ne soient pas considérées comme indépendantes dans l'exercice d'une fonction de litige. On ne sait pas encore clairement comment les obligations de confidentialité énoncées à l'article 78 du règlement sur l'IA influeraient sur l'utilisation d'informations obtenues dans cet objectif dans le cadre d'un litige civil (cf. procédures pénales et administratives explicitement mentionnées).

Déterminez si la preuve sur laquelle vous voulez vous appuyer a été obtenue :

1. en raison de l'exercice par l'organisme de promotion de l'égalité de son rôle de juge indépendant et impartial ou de son rôle de supervision et de traitement des données ; ou
2. aux fins du litige..

Si les données ont été créées avec pour objectif la mention 1., l'indépendance de ces données pourrait être remise en question. Comme l'a noté Equinet :

> Lorsqu'un organisme de promotion de l'égalité plaide sur la base de ses propres données, le tribunal pourrait, en théorie, avoir du mal à considérer ces données comme externes et indépendantes. En outre, lorsqu'un organisme de promotion de l'égalité statue sur des affaires en se fondant (en partie) sur ses propres données en matière d'égalité, les parties à une affaire et d'autres observateurs pourraient remettre en question une telle pratique pour deux motifs au moins : l'organisme est-il un arbitre impartial s'il identifie et utilise de manière proactive des éléments de preuve qui favorisent la victime présumée [?] ; et ces éléments de preuve sont-ils crédibles si l'organisme les a produits, étant par définition favorables à l'égalité/aux victimes [?]. Ce type de questions potentielles de procès/d'égalité des chances (perçues) comme équitables pourraient affecter la crédibilité globale d'un organisme de promotion de l'égalité dans la société, y compris son rôle en matière de rapports. (Ilieva 2024)

Analyses statistiques du système algorithmique

Dans les cas où le système algorithmique et les données pertinentes sont à la disposition de l'OPE, celui-ci pourrait effectuer sa propre évaluation statistique. Les mesures d'équité et les méthodes d'application ont été mises en œuvre dans un large éventail de boîtes à outils d'équité en open source. Ces boîtes à outils peuvent être utilisées pour mesurer les écarts de performance dans les systèmes algorithmiques ou les jeux de données mis à leur disposition. Parmi les exemples les plus utilisés, citons IBM AI Fairlearn 360 et Microsoft Fairlearn, mais beaucoup d'autres sont disponibles. Il n'existe pas qu'une seule boîte à outils spécifique recommandée car chacune a ses propres avantages et inconvénients (Mittelstadt 2024).

Test de situation

Des tests de « situation » peuvent également être utilisés pour générer des preuves de discrimination. Le test de situation est une méthode expérimentale où des paires de sujets de test sont établies, qui ne diffèrent que par une caractéristique protégée, pour tester les différences de traitement.

Dans certains contextes, comme le recrutement algorithmique, l'une des options viables consiste à recueillir des données statistiques sur un système algorithmique spécifique par le biais de tests de situation (de discrimination). Dans divers pays, dont le Danemark, la Finlande, la France, les Pays-Bas, le Royaume-Uni et la Suède, les tests de situation ont été admis en tant que preuve valable (Commission européenne, 2019). Les tests de situation à grande échelle peuvent également produire des données statistiques et ont été appliqués dans de nombreux contextes différents, tels que l'accès à l'emploi, au logement et à d'autres types de biens et services. Les résultats des tests sont considérés comme des preuves recevables dans plusieurs pays de l'UE. Pour obtenir des résultats crédibles, l'OPE pourrait devoir collaborer avec des expert·es pour réaliser ce type de tests (Makkonen, 2016).

Collaboration avec des instituts de statistique/utilisation de statistiques générales

Les statistiques générales peuvent être utilisées pour argumenter que l'utilisation de proxys conduit à une différence de traitement. Par exemple, dans le cas néerlandais des contrôles des bourses d'études :

- l'algorithme différenciait au niveau de l'éducation ;
- selon les statistiques générales du Bureau néerlandais des statistiques, les étudiants et étudiantes issu·es de l'immigration non-européenne sont surreprésenté·es dans les niveaux d'enseignement les plus bas : il y a environ deux fois plus d'étudiant·es issu·es de l'immigration non-européenne que ce à quoi l'on pourrait s'attendre par rapport à l'ensemble des étudiant·es (StatLine – Mbo; studenten, niveau, leerweg, migratie 2005/'06-2021/'22, disponible en néerlandais) ;
- il s'ensuit, sans accéder à des données de base spécifiques pour les étudiant·es individuel·les, que l'utilisation de l'entraînement dans les systèmes algorithmiques conduit à un traitement inégal.

Approche des tribunaux en matière de données statistiques

Sur la nécessité de disposer de données statistiques

Les statistiques en elles-mêmes ne révèlent pas nécessairement une discrimination ou une pratique discriminatoire (Hugh Jordan c. Royaume-Uni, requête n° 24746/94, Cour européenne des droits de l'homme)[54]. Elles font partie des éléments de preuve pertinents qui peuvent être utilisés pour établir ou réfuter une preuve prima facie de discrimination directe ou indirecte. Les tribunaux peuvent généralement être plus enclins à accepter les preuves statistiques présentées par les personnes ayant déposé la plainte pour établir une preuve prima facie de discrimination directe, que de se fonder sur de telles preuves fournies par la défense pour les réfuter,

54 Les statistiques font apparaître une différence entre le nombre de personnes tuées de différentes religions, mais non pas si ces meurtres étaient contraires à la loi ou s'ils ont résulté d'un usage excessif de la force par les membres des forces de sécurité (paragraphe 154).

ce qui reflète le déséquilibre informationnel entre les parties et les difficultés reconnues auxquelles les personnes ayant déposé une plainte sont confrontées pour étayer leurs allégations de discrimination (Makkonen 2007). Il appartiendra à la juridiction nationale d'apprécier si les statistiques disponibles sont « valables » et peuvent être prises en compte (voir Enderby, Affaire C-127/92, paragraphe 16)..

Données statistiques et discrimination directe

Dans une affaire de discrimination directe, les statistiques peuvent révéler l'existence d'un schéma discriminatoire et constituer ainsi une preuve de discrimination directe dans le cadre d'une plainte individuelle. Toutefois, l'utilisation de preuves statistiques n'est généralement pas nécessaire pour établir une preuve prima facie de discrimination directe.

Données statistiques et discrimination indirecte

Les données statistiques sont particulièrement utiles dans le cas d'une discrimination indirecte, lorsqu'une personne ayant déposé une plainte peut utiliser des statistiques pour démontrer que la mise en œuvre d'une pratique ou d'une règle apparemment neutre a un effet ou un impact disproportionné sur un groupe de personnes, en particulier par rapport à d'autres personnes se trouvant dans une situation similaire (FRA 2018: 242).

En particulier « lorsqu'un demandeur est en mesure de démontrer, sur la base de statistiques officielles incontestées, l'existence d'une indication prima facie qu'une règle spécifique (bien que formulée de manière neutre) affecte en fait un pourcentage nettement plus élevé de femmes que d'hommes, il appartient au gouvernement défendeur de démontrer que cela résulte de facteurs objectifs non liés à une quelconque discrimination fondée sur le sexe » (Hoogendijk c. Pays-Bas, 2005).

Aucun seuil strict n'a été fixé mais un chiffre substantiel doit généralement être atteint pour montrer l'impact différentiel, c'est-à-dire un « nombre beaucoup plus grand » (Ingrid Rinner-Kühn, 1989) ou un « pourcentage considérablement plus faible » (Helga Nimz, 1991 et Maria Kowalska, 1990) ou « beaucoup plus » (M. A. De Weerd, 1994), bien que la présentation de données statistiques pour prouver la discrimination indirecte ne soit pas une exigence formelle (voir D.H. et al c. la République tchèque, 2007, § 188). Toutefois, des données statistiques précises ne sont pas requises, notamment si la partie qui prétend être victime de discrimination n'a pas accès à ces statistiques ou a des difficultés à y accéder (Minoo Schuch-Ghannadan, 2019).

Utilisation d'autres données et preuves

La preuve de pratiques ou d'attitudes adoptées à l'égard d'autres personnes appartenant à la même catégorie protégée peut suffire. Par exemple, la pratique consistant à placer des enfants dans des classes séparées en raison de leur maîtrise insuffisante du croate n'est appliquée qu'aux élèves roms, ce qui donne lieu à une présomption de traitement différent bien que les statistiques n'établissent pas la

discrimination. La Cour européenne des droits de l'homme a clairement indiqué que « la discrimination indirecte peut être prouvée sans preuves statistiques » (Oršuš et al. c. Croatie, 2010).

Dans les affaires impliquant des systèmes algorithmiques, d'autres informations et éléments de preuve factuels peuvent également suffire à établir une discrimination prima facie. Par exemple, dans l'affaire Deliveroo, le tribunal n'a pas exigé de tests statistiques approfondis ou d'analyses de données complexes. Au contraire, la conception même du système (telle qu'elle a été reconstruite à un niveau élémentaire à partir des informations disponibles comme les clauses contractuelles, les FAQ de la plateforme, les communications externes et internes de l'organisation intimée et les témoignages des individus affectés), combinée à une évaluation de ses effets prévisibles, suffisait à démontrer un désavantage disproportionné et à établir une présomption de discrimination (prima facie).

Évaluation des données statistiques sur l'égalité par l'organisation intimée

Aux fins du droit à la non-discrimination, les données statistiques doivent être à la fois valides et fiables.

Les tests de validité et de fiabilité devraient être appliqués tout au long de la procédure. Le processus en trois étapes suivant peut vous aider.

- L'approche adoptée par le fournisseur ou le déployeur pour les tests et l'évaluation a-t-elle été fondée sur ou produit-elle des données statistiques valides et fiables ?
- Votre preuve statistique cherche-t-elle à remettre en question l'organisation intimée de manière à la fois valide et fiable ?
- Des preuves statistiques en réfutation sont-elles valides et fiables ?

Pour que les données statistiques soient considérées comme « valides », les critères suivants doivent être remplis :

- les données semblent généralement importantes.
- elles couvrent suffisamment d'individus.
- elles capturent des phénomènes qui ne sont pas fortuits ou à court terme.
- dans l'ensemble, elles sont « pertinentes et suffisantes ».

La fiabilité des données statistiques implique la stabilité ou la cohérence de la mesure/du test appliqué. Une mesure de discrimination, par exemple, est fiable si la procédure de mesure donne les mêmes résultats lors d'essais répétés. Par exemple, si l'on mesure le nombre d'incidents d'égalité signalés chaque année, mais que le nombre d'organismes compétents (pour l'égalité) change au fil du temps, cela aura un effet négatif sur la fiabilité de cette mesure et aura également une incidence sur

sa comparabilité dans le temps. Aucune mesure n'est fiable. La fiabilité est donc toujours une question de degré[55].

Choix des mesures d'égalité

En science des données, plusieurs métriques ont été élaborées pour mesurer différentes notions de l'équité. Ces indicateurs n'ont pas été conçus à partir du droit à la non-discrimination et les chevauchements entre les mesures d'équité dans le domaine de la science des données et de l'IA, et les notions d'égalité, ne sont pas toujours clairs (Weerts et al. 2023). Malheureusement, il est mathématiquement impossible de créer un système qui soit juste selon toutes les métriques techniques.

Par conséquent, il est essentiel de déterminer ce que signifient une différence de traitement et un désavantage particulier (en tant que partie intégrante de la discrimination) dans le contexte de la présente affaire et de choisir les paramètres qui chevauchent au plus près cette différence. Notez que le simple fait de présenter une mesure d'équité qui montre que le système est « équitable » sur cette mesure ne constitue pas un système qui entraîne l'égalité de traitement. En cas de non-équité, il sera important de se demander pourquoi cette décision d'utiliser cet ensemble de mesures de l'égalité a été prise et quelles mesures ont été prises pour remédier à l'inégalité qui en résulte.

La parité démographique permet de déterminer si la probabilité d'une prédiction donnée est la même pour deux groupes. Cette mesure est particulièrement pertinente lorsqu'un résultat est désavantageux par rapport aux autres pour les personnes impliquées. Cela reflète mieux le concept d'inégalité de traitement dans des exemples tels que la prédiction de fraude (lorsqu'un groupe est soupçonné de fraude plus souvent que l'autre groupe) ou le recrutement (lorsqu'un groupe est sélectionné pour un entretien plus souvent que l'autre). Dans de nombreux cas, la parité démographique est la mesure pertinente, sauf s'il existe des raisons impérieuses de s'en remettre à un autre type de mesure. Notez que la parité démographique est aussi parfois appelée indépendance, équité entre les groupes, parité statistique, impact disparate, taux de sélection égal ou taux d'acceptation égal.

D'autres mesures pertinentes reposent sur des erreurs dans le système et testent si les systèmes entraînent plus d'erreurs pour un groupe que pour l'autre. Ce type de mesure est particulièrement pertinent lorsque le préjudice provient majoritairement d'erreurs dans les systèmes, et non du résultat lui-même. La reconnaissance faciale et les tests diagnostiques en médecine sont des exemples où les mesures fondées sur les erreurs sont les plus pertinentes. Voir le tableau 6 ci-dessous pour un aperçu des mesures potentiellement pertinentes.

Il est peu probable qu'une approche unique en matière de tests soit suffisante. Il convient plutôt de prendre en considération à la fois les modèles statistiques utilisés et leurs limites et les circonstances environnantes telles que l'organisation

55. Voir Commission européenne 2021: 20; Enderby, 1993, paragraphe 17 ; Seymour-Smith, 1999, paragraphe 62. Concernant les exigences générales en matière de statistiques dans les affaires de non-discrimination ; voir Enderby c. Frenchay Health Authority et Secretary of State for Health, paragraphe 17.

décisionnelle et les processus qu'ils adoptent (Barocas et al. 2023). En tant que tels, des « chiffres » sans contexte auront une signification insuffisante pour être valides ou fiables.

Même lorsque la parité démographique est choisie en tant que métrique principale, il est conseillé d'envisager également une métrique qui tient compte des erreurs. Les fournisseurs sont tenus de fixer un « niveau d'exactitude prévu » pour des groupes spécifiques de personnes (règlement sur l'IA, annexe IV(3)).

Les statistiques sur l'égalité peuvent être présentées sous forme de ratio ou de mesures distinctes entre les groupes. Quelques exemples sont présentés ci-dessous.

- Les statistiques de parité démographique dans un contexte de recrutement peuvent être exprimées comme suit : 25 % des candidats masculins sont invités à un entretien et 15 % des candidates féminines sont invitées après la sélection automatisée de CV. Ce constat peut également être exprimé sous forme de ratio : 0.15/0.25 = 0.6.
- L'égalité des chances pour le dépistage du cancer peut être exprimée comme suit : 75 % des hommes atteints de cancer sont détectés par l'outil de dépistage et 65 % des femmes atteintes de cancer sont détectées par l'outil de dépistage. Ce constat peut également être exprimé sous forme de ratio : 0.65/0.75 = 0.87.

Tableau 6. Description de mesures d'équité

(adapté avec la permission de Mittelstadt 2023)

Métrique	Description	Applicable lorsque	Exemple
Parité démographique	Les différents groupes doivent recevoir des résultats positifs selon le même ratio. La parité démographique pourrait être appelée indépendance, équité entre les groupes, parité statistique, impact disparate, taux de sélection égal ou taux d'acceptation égal.	▸ Situations où un résultat particulier lui-même (et non des erreurs) est désavantageux. ▸ Situations où les données historiques sont censées être préjudiciables. ▸ Situation où il n'y a pas d'étiquettes fiables (pas de vérité terrain convenue).	Recrutement, acceptation de prêt, détection de fraude, accès à l'éducation

Métrique	Description	Applicable lorsque	Exemple
Égalité des chances	Différents groupes devraient recevoir des faux négatifs selon le même ratio. On pourrait également parler d'équilibre du taux d'erreur de faux négatif.	▸ Situations où le préjudice écrasant provient de faux négatifs, et ▸ Il existe des étiquettes fiables (vérité terrain convenue).	Dépistage de maladies
Parité prédictive	Le pourcentage de faux positifs, parmi tous les positifs, devrait être le même pour les différents groupes. On pourrait également parler de test de résultat ou parité de valeur prédictive positive.	▸ Situations où le préjudice écrasant provient de faux positifs, et ▸ Il existe des étiquettes fiables (vérité terrain convenue).	Reconnaissance faciale par la police pour identifier une personne d'intérêt
Équilibre du taux de faux positifs	Différents groupes devraient recevoir des faux positifs selon le même ratio. On pourrait également parler d'égalité prédictive.	▸ Situations où le préjudice écrasant provient de faux positifs, et ▸ Il existe des étiquettes fiables (vérité terrain convenue).	Reconnaissance faciale par la police pour identifier une personne d'intérêt
Chances égalisées	Les taux de vrais positifs et de faux positifs devraient être les mêmes pour les différents groupes. On pourrait également parler d'égalité d'exactitude de la procédure conditionnelle et de mauvais traitements disparates. procedure accuracy equality and disparate mistreatment.	▸ Les situations où le préjudice causé par les faux positifs et les faux négatifs doit être pris en compte, et ▸ Il existe des étiquettes fiables (vérité terrain convenue).	Traitement de la maladie par une chirurgie à risque

Évaluation des statistiques fournies par l'organisation intimée

Pour évaluer la validité des indicateurs statistiques de l'égalité, l'OPE devrait au moins envisager trois problèmes.

1. D'une manière générale, la qualité et les biais des données utilisées pour calculer la métrique (voir étape 3.5).
2. La conception de la métrique et le choix des groupes comparatifs.

Lors de l'examen de la manière dont un fournisseur d'un système a identifié et mesuré les écarts de performance, les OPE devraient accorder une attention particulière à la manière dont la décision de mesure a été prise et aux groupes de personnes qui sont comparés. Il s'agit d'une décision importante dans le contexte de la législation sur la non-discrimination, où la définition de groupes appropriés « défavorisés » et « de comparaison » est souvent essentielle. (Mittelstadt 2024). Le choix des groupes peut (délibérément) cacher des résultats inégaux dans le monde réel.

Il est conseillé aux OPE :

- d'examiner quels groupes « défavorisés » et « de comparaison » appropriés seraient pris en considération, avant d'examiner les statistiques partagées par les organisations intimées ;
- de déterminer si l'organisation intimée a choisi et défini les mêmes groupes et, dans le cas contraire, déterminer quels renseignements sont ajoutés ou perdus dans les choix de l'organisation intimée ;
- de prendre en considération les groupes intersectionnels, car se concentrer uniquement sur les groupes à une seule caractéristique cache des inégalités intersectionnelles ;
- afin de comprendre les groupes de comparaison choisis par l'organisation intimée, d'examiner les explications fournies par celle-ci, en particulier toute information sur la mesure à travers laquelle une variation de données à caractère personnel prises en compte aurait conduit à un résultat différent ;
- dans les cas où l'organisation intimée n'a pas communiqué de données sur les métriques d'équité pertinentes ou a défini les groupes concernés de manière restrictive, d'envisager d'acquérir des données statistiques supplémentaires (voir l'encadré sur l'acquisition de données statistiques (supplémentaires) ci-dessus) (Mittelstadt 2024).

Les mesures statistiques d'égalité font souvent partie de rapports de performance plus larges qui comprennent un large éventail de résultats de tests mesurant divers aspects de la performance d'un système, comme la précision (fréquence à laquelle le système fait la bonne prédiction) et la robustesse (résilience face aux erreurs, aux comportements imprévus et aux changements du monde réel (déviation)). Il est utile pour les OPE de comprendre les performances en général, car les performances du système sont très pertinentes dans l'évaluation de la discrimination algorithmique, en particulier lorsqu'il s'agit d'envisager une justification objective potentielle (voir étape 3.8), où les questions essentielles impliquent de vérifier si le système est adéquat ou approprié et nécessaire.

3. Pertinence des mesures par rapport au système actuel tel qu'appliqué dans la pratique.

Les métriques doivent refléter le système algorithmique actuel. Les tests appliqués peuvent être devenus obsolètes ou le processus autour du système peut avoir changé (par exemple, le système met en œuvre une politique gouvernementale et cette politique a changé).

Notez que les paramètres rapportés devraient refléter le système après la mise en œuvre de toute mesure d'atténuation visant à « défavoriser » ou à prévenir la discrimination. Si ces mesures sont manuelles, l'OPE devrait s'assurer que les statistiques reflètent le résultat après cette intervention manuelle.

Lorsqu'un système d'IA à haut risque a reçu un certificat de conformité à la suite d'une évaluation, il est présumé conforme aux exigences en matière de données énoncées à l'article 10(4) du règlement sur l'IA, à savoir que les « jeux de données tiennent compte, dans la mesure requise par la destination, des caractéristiques ou éléments propres au cadre géographique, contextuel, comportemental ou fonctionnel spécifique dans lequel le système d'IA à haut risque est destiné à être utilisé ».

Il est probable que ces considérations guideront l'application du test de « validité » (voir plus sur les données statistiques à l'étape 3.6), mais il est important de noter que cette présomption peut être réfutée. Les organismes de promotion de l'égalité voudront peut-être examiner, en particulier, si les données utilisées répondent effectivement à ces exigences pour l'utilisation prévue du système d'IA. Si le système est utilisé autrement que pour son usage prévu, cette présomption ne s'appliquera pas.

Questions à envisager

Sur les tests de performance généraux

- Quels aspects des performances ont été testés ?
- Comment et où le rendement du système parmi les individus et les groupes de personnes a-t-il été enregistré ?
  - Ces chiffres seront-ils mis à jour au cours du cycle de vie du système ? Si oui, comment ?
- Quels impacts potentiels sur les individus et les groupes d'individus ont été pris en compte lors de la réalisation des tests de performance ?
- Une forme quelconque de préjudice a-t-elle été choisie comme point focal pour les tests (comme les faux positifs ou les faux négatifs) ?
- Quelles mesures ont été utilisées pour les tests de performance ?
- Des matrices de confusion ont-elles été produites au travers de tests de performance ?
- Comment les résultats des tests ont-ils été validés ?

- Les données utilisées pour effectuer les tests de performance sont-elles suffisamment représentatives, pertinentes et mesurées de manière précise et appropriée (voir étape 3.5) ?
- Les tests de performance effectués par l'organisation intimée répondent-ils aux critères de validité et de fiabilité ?

Données statistiques sur l'égalité

- Quel(le)s (types de) données statistiques sur l'égalité, mesures d'équité ou tests de biais ont été utilisés ?
  - Les paramètres choisis par l'organisation intimée reflètent-ils l'égalité de traitement pour le contexte spécifique ? (Habituellement, la parité démographique, à moins qu'il n'y ait des justifications claires pour d'autres paramètres, voir le tableau 6 ci-dessus).
  - Dans les cas où des mesures sont fondées sur des erreurs, les mêmes formes de préjudice ont-elles été considérées que dans les tests de performance réguliers ? (Par exemple, lorsque des tests de performances réguliers ont utilisé une métrique axée sur les faux positifs, les tests d'égalité ont-ils fait de même ?) Lors du test d'égalité fondé sur les performances, logiquement, la mesure de performance la plus pertinente doit également être utilisée pour ce test d'égalité. Un choix non motivé d'une mesure différente pourrait indiquer une omission dans le test d'égalité.
  - Quelle est la justification ou la base sur laquelle ces tests ont été choisis ?
  - D'autres ont-ils été pris en considération et/ou rejetés ? Si c'est le cas, pourquoi/pourquoi pas ?
  - La gamme de tests utilisée couvrant différentes caractéristiques protégées était-elle suffisamment large pour saisir différents types de discrimination potentielle ou d'impacts affectant différents groupes (intersectionnels) ?
  - Des groupes « de comparaison » pertinents ont-ils été choisis pour chaque test ? Ces considérations sont-elles incluses dans la documentation technique du système ?
- Y a-t-il eu un recours à des (types de) méthodes de mise en œuvre de l'équité ou de débiaisement ?
  - Quelle est la justification ou la base sur laquelle ces tests ont été choisis ?
  - D'autres ont-ils été pris en considération et/ou rejetés ? Si c'est le cas, pourquoi/pourquoi pas ?
  - Ces considérations sont-elles incluses dans la documentation technique du système ?
- Quels sont les avantages et quels préjudices ont résulté de la mise en œuvre de l'équité ou du débiaisement ?
  - Quels groupes de personnes ont ressenti/ressentiront un impact ?
  - Les effets secondaires sur l'égalité de traitement pour d'autres groupes ont-ils été étudiés ?
- Comment les résultats des tests ont-ils été validés ?

- Les données utilisées pour calculer les statistiques sont-elles suffisamment représentatives, pertinentes et mesurées de manière précise et appropriée (voir étape 3.5) ?
- Les données statistiques fournies par l'organisation intimée répondent-elles aux critères de validité et de fiabilité ?

Sources potentielles d'information pour évaluer les données statistiques sur l'égalité

Personne ayant déposé la plainte

RGPD, article 15(1)(h), des informations utiles sur la logique sous-jacente

Organisation intimée

De la part du fournisseur

Règlement européen sur l'IA, article 10, données et gouvernance des données

Règlement européen sur l'IA, article 11, documentation technique (annexe IV(2)(g) ;(3)) :

- « les procédures de validation et d'essai utilisées, y compris les informations sur les données de validation et d'essai utilisées et leurs principales caractéristiques; les indicateurs utilisés pour mesurer l'exactitude, la robustesse et le respect des autres exigences pertinentes énoncées au chapitre III, section 2, ainsi que les éventuelles incidences discriminatoires; les journaux de test et tous les rapports de test datés et signés par les personnes responsables, y compris en ce qui concerne les modifications prédéterminées visées au point f) ».

Règlement européen sur l'IA, article 11, documentation technique

Autres informations qui pourraient être demandées à l'organisation intimée :

- accès au système ;
- exemples de jeux de données (synthétiques) ;
- tout rapport sur les performances ou la fiabilité du système, y compris les matrices d'erreurs ;
- des informations sur le type de tests et la méthodologie de test effectués, y compris les normes (harmonisées) potentielles suivies ;
- tous les rapports sur les tests de biais et d'équité effectués ;
- des informations sur le type de tests et la méthodologie de test effectués, y compris les normes (harmonisées) potentielles appliquées ;
- AIDH ou AIDF volontaire, AIPD ou analyse d'impact sur l'éthique et les droits de l'enfant ;
- toute autre évaluation (du risque) partagée par l'organisation intimée ;
- demandes d'accès à l'information.

Indicateurs de discrimination

- Il existe des écarts dans les résultats ou des erreurs entre les groupes de personnes présentant des caractéristiques protégées, sur le fondement des indicateurs les plus pertinents qui reflètent l'égalité de traitement pour le cas spécifique.
- Les indicateurs les plus pertinents qui reflètent l'égalité de traitement pour un cas spécifique ne sont pas disponibles ou des indicateurs non pertinents sont utilisés.
- Il existe des écarts dans les résultats ou des erreurs entre les groupes de personnes présentant des caractéristiques protégées, en fonction d'autres paramètres (autres que celui qui reflète le plus fidèlement l'égalité de traitement pour le cas particulier).
- Tous les groupes affectés n'ont pas été inclus, ou les groupes affectés ont été définis de manière trop restrictive, sans tenir compte de la discrimination intersectionnelle.
- Des groupes corrects ont été identifiés et des mesures pertinentes ont été appliquées, mais les conclusions ne résistent pas à un examen minutieux, par exemple en raison de problèmes de données.
- Les rapports ne font mention que d'une seule mesure testée plutôt que de résultats de plusieurs mesures d'équité (comme la précision, le rappel, les taux d'erreur[56]). Des tests sur plusieurs indicateurs peuvent être nécessaires pour révéler différents types d'écarts de performances.

Indicateurs cumulatifs de discrimination

Tous les indicateurs susmentionnés peuvent fonctionner de manière cumulative. Ce ne sont là que quelques exemples.

- Risque de discrimination identifié par le fournisseur/déployeur (voir étape 3.3) mais les tests ne couvraient pas l'ensemble complet des individus et/ou groupes à risque en termes de préjudice – tenez compte notamment des proxys, des discriminations multiples et intersectionnelles. Par exemple, la discrimination a été testée sur le motif du genre ou du sexe, mais pas de la « race » ou de l'origine ethnique, tandis que le risque de discrimination raciale a été identifié.
- (Risque de) discrimination identifié(e) par des tiers (voir étape 3.1) mais les tests ne couvraient pas l'ensemble complet des individus et/ou groupes à risque en termes de préjudice – tenez compte notamment des proxys, des discriminations multiples et intersectionnelles. Par exemple, la discrimination a été testée sur le motif du genre ou du sexe, mais pas de la « race » ou de l'origine ethnique, alors que dans des cas similaires existants, des personnes ont été défavorisées en raison de la « race » et de la grossesse.
- Les critères d'entrée sont fortement liés aux motifs protégés ou des proxys sont utilisés dans des algorithmes (voir étape 3.3) et aucun test n'est effectué

56 Dans les analyses de données, différents indicateurs peuvent être utilisés pour refléter différents aspects de la performance d'un système. Par exemple, la précision se concentre sur l'exactitude des prédictions positives (en évitant les fausses alertes), et le rappel mesure la capacité à détecter toutes les cas positifs réels (en évitant de passer à côté d'événements).

pour comprendre si les exigences d'égalité de traitement sont satisfaites pour le groupe protégé.

- Consignez tout écart de production ou de traitement tel que révélé dans les statistiques.
- Consignez tout problème lié aux statistiques signalées (par exemple, un mauvais indicateur ou des problèmes de données).
- Notez tout écart dans le test.
- Combinez les résultats de cette étape avec ceux de l'étape précédente pour comprendre si, ensemble, un cas de discrimination prima facie peut être invoqué et/ou si l'organisation intimée a réussi à réfuter la présomption ou à justifier la discrimination (si vous examinez la justification, voir aussi 3.4 pour examiner l'objectif du système).
- Notez la conclusion.

Exemple : Smart Check Amsterdam

Ce modèle a été conçu pour filtrer les demandes d'aide sociale et identifier celles qui sont les plus susceptibles de comporter des erreurs majeures, les signaler et envoyer les affaires à un service d'enquête pour un examen plus approfondi. L'objectif était de réduire le nombre de personnes signalées pour enquête.

Quinze caractéristiques ont été prises en considération, notamment tout un éventail de facteurs tels que la question de savoir si la personne avait déjà demandé ou reçu des prestations, le nombre d'adresses qu'elle avait et la somme de ses actifs pour déterminer une cote de risque. Le modèle a délibérément évité d'utiliser des motifs protégés et a essayé d'éviter les « proxy ».

Selon les auteurs, « il a été entraîné sur un jeu de données englobant 3 400 enquêtes précédentes. Il en a découlé un risque d'intégration de biais historiques. Par exemple, si les personnes en charge du dossier avaient historiquement provoqué des taux d'erreurs plus élevés avec un groupe ethnique spécifique, le modèle pourrait apprendre à tort à prédire que ce groupe commet des fraudes à des taux plus élevés ».

Les fonctionnaires d'Amsterdam ont adopté une définition de l'équité fondée sur la répartition égale de la charge des enquêtes abusives entre les différents groupes démographiques. Il a également mené de vastes consultations. Tous les groupes de parties prenantes, de l'APD d'Amsterdam aux cabinets de conseil externes, ont donné leur approbation au projet, à l'exception du groupe représentant les personnes directement concernées (bénéficiaires des prestations).

Néanmoins, le système a été mis à l'essai. Le modèle initial était deux fois plus susceptible de signaler à tort des candidat·es non néerlandais·es et près de deux fois plus susceptible de signaler un·e candidat·e de nationalité non européenne qu'un·e candidat·e de nationalité européenne . Il était également 14 % plus susceptible de signaler à tort des hommes pour enquête.

La ville d'Amsterdam a également testé les données recueillies auprès des personnes en charge du dossier. Dans ce processus, il a été découvert que les êtres humains étaient plus susceptibles de signaler à tort les personnes néerlandaises et les femmes.

En raison de l'existence d'un biais, l'équipe a tenté d'utiliser la repondération des données d'entraînement pour remédier au problème. Cela signifiait que les personnes de nationalité non européennes réputées avoir commis des erreurs significatives dans leurs demandes se voyaient accorder un poids moindre dans les données et celles de nationalité européenne un poids plus important. Une fois le modèle repondéré, les personnes néerlandaises et non-néerlandaises risquent tout autant d'être signalées à tort.

Le problème de la nouvelle pondération des données est que, bien qu'elle puisse réduire l'équité à l'égard des personnes issues de l'immigration, elle ne garantirait pas l'équité entre les identités qui se croisent.

Néanmoins, le système a été déployé sur une base limitée pour être utilisé sur de vrai·es candidat·es. Sa performance devait être mesurée en fonction de deux critères clés : i) pourrait-il examiner les demandeur·ses sans biais ? ii) pourrait-il détecter la fraude à l'aide sociale mieux et plus équitablement que les êtres humains en charge de dossier ?

Au lieu de signaler moins de demandeur·es pour enquête, il en a signalé plus. Et ne valait pas mieux qu'un être humain pour identifier ceux et celles qui avaient besoin d'un examen approfondi. En outre, malgré la nouvelle pondération, le système signalait désormais à tort les demandeur·ses de nationalité néerlandaise et les femmes. En outre, les demandeur·ses d'aide sociale ayant des enfants étaient également plus susceptibles d'être signalé·es pour enquête.

En fin de compte, le projet n'a pas été poursuivi.

Ce qui précède est un résumé de Inside Amsterdam's high-stakes experiment to create fair welfare AI, 11 juin 2025, MIT Technology Review, Lighthouse Reports et Trouw E. G., Geiger G. et Braun J-C. Pour plus d'informations, voir Lighthouse Reports 2025.

3.7. Évaluer si les éléments de preuve sont suffisants pour etayer une preuve prima facie de discrimination impliquant un systeme d'IA ou algorithmique

- Passez en revue toutes les informations et les réponses aux questions de l'étape 3 pour déterminer si une preuve prima facie de discrimination a été établie. Utilisez les indicateurs de discrimination en tant que guide.
- À ce stade, supposons qu'il n'y ait pas d'explication adéquate aux faits démontrant une preuve prima facie (il peut y en avoir ; toutefois, cela serait pertinent pour évaluer la défense de l'organisation intimée (voir 3.8 ci-dessous)).
- De même, un manque de preuves ou un manque de transparence concernant un système ne devrait pas empêcher qu'une affaire soit engagée ou poursuivie à ce stade. Souvenez-vous qu'à l'étape 1, il a été déterminé que le rejet précoce des plaintes en cas de déséquilibre informationnel était probablement inapproprié : voir 1.4 Manjang c. Uber Eats UK Ltd & Autres (2021).

Vous devrez peut-être tenir compte de tout ou partie des éléments suivants.

- Qui est la victime et quels sont les motifs protégés en cause ?
- Qui est l'élément de comparaison approprié ? Pourquoi cet élément de comparaison a-t-il été choisi ?
- Y a-t-il eu discrimination directe ou indirecte ?
- Existe-t-il des formes multiples ou intersectionnelles de discrimination ?
- Avez-vous établi une discrimination par proxy ?
- Pourquoi la discrimination a-t-elle eu lieu (causalité) ? Examinez, par exemple, si la discrimination a eu lieu en raison de l'objectif du système, du développement du système, de l'utilisation de données au sein du système ou parce que le système n'est pas utilisé aux fins prévues ou est mal utilisé par le déployeur.

En ce qui concerne la discrimination indirecte, tenez compte également ce qui suit.

- Qu'est-ce que la disposition, le critère ou la pratique neutre ?
- En quoi les effets de la disposition, du critère ou de la pratique sont-ils nettement plus négatifs sur un groupe protégé ?
- La personne ayant déposé une plainte a-t-elle prouvé des faits permettant de conclure qu'elle a été traitée de manière moins favorable en raison de motifs protégés ?
- Si la différence de traitement n'est pas « directement » liée à une caractéristique protégée, est-elle suffisamment liée ? Ce sera particulièrement important lors de l'examen de la discrimination par proxy.
- Quel est le contexte de la discrimination ? L'organisation intimée peut-elle faire valoir avec succès que la décision a été prise pour un motif autre qu'une discrimination interdite ?

Compte tenu de toutes les circonstances, une preuve prima facie de discrimination a-t-elle été établie ?

Si oui, passez à la section 3.8.

Sinon, tenez compte des éléments suivants.

- Y a-t-il eu des manquements aux obligations en matière de documentation, de tests, de surveillance ou de déclaration qui pourraient expliquer la discrimination ou entraîner des lacunes dans la preuve ?
- Y a-t-il des mesures
  - Envisagez de présenter des demandes de divulgation de preuves fondées sur des considérations de transparence et d'accès à la justice au titre de l'article 13 de la Convention européenne des droits de l'homme (droit à un recours effectif devant une autorité nationale) et de l'article 47 de la Charte des droits fondamentaux de l'Union européenne (droit à un recours effectif et à un procès équitable).
  - Veuillez noter que dans certaines circonstances, lorsque l'organisation intimée cherche à faire valoir le secret des affaires ou les droits de tiers, des procédures confidentielles peuvent être disponibles par le biais de l'ASM, de l'APD ou des tribunaux. La jurisprudence actuelle suggère qu'il existe une tendance croissante à exiger que ces informations soient mises à disposition ; voir Dun & Bradstreet [67] (dans le contexte du RGPD) :

    > Dans l'hypothèse où le responsable du traitement considère que les informations à fournir à la personne concernée conformément à cette disposition comportent des données de tiers protégées par ce règlement ou des secrets d'affaires, au sens de l'article 2, point 1, de la directive 2016/943, ce responsable est tenu de communiquer ces informations prétendument protégées à l'autorité de contrôle ou à la juridiction compétentes, auxquelles il incombe de pondérer les droits et les intérêts en cause aux fins de déterminer l'étendue du droit d'accès de la personne concernée prévu à l'article 15 du RGPD. (voir aussi confidentialité)

Pour connaître la pertinence de la transmission des informations statistiques pour établir une preuve prima facie de discrimination, voir l'étape 3.6 sur les données statistiques .

Transparence

Une juridiction peut tirer des conclusions de l'absence de production de données statistiques pertinentes ou d'autres éléments de preuve. Une présomption de discrimination peut également être déduite des faits et des circonstances entourant l'affaire, comme les réponses d'une organisation intimée aux questions ou le défaut de se conformer à un code de bonnes pratiques pertinent. En ne fournissant pas d'informations, l'organisation intimée peut, en pratique, ne pas parvenir à renverser la présomption de discrimination qui pèse sur elle, conduisant la juridiction à présumer l'existence d'une discrimination (voir 3.8 pour les circonstances dans lesquelles une charge juridique de fournir des informations se présente).

- Des conclusions peuvent souvent être tirées d'une réponse équivoque, inadéquate ou évasive à un questionnaire ou d'une divulgation au cours du processus préalable à l'action ou de la procédure.
- Des conclusions peuvent également être tirées du refus de confirmer ou de nier l'existence d'informations.

Questions à envisager

- L'organisation intimée a-t-elle omis de divulguer certaines informations pertinentes, ce qui a entraîné un manque de transparence ? Par exemple, a-t-elle répondu de manière adéquate à toute demande dans le cadre du RGPD, à toute demande liée à l'accès à l'information, à toute correspondance préalable à l'action, à toute demande faite en vertu du règlement sur l'IA ou autre ?
- Peut-on tirer des conclusions des éléments de preuve disponibles ? Par exemple, les statistiques, les faits environnants, la discrimination connue ?
- Peut-on tirer des conclusions de l'absence de preuve ? Par exemple, l'absence de réponse aux questions, le non-respect des directives ou des codes de bonnes pratiques ?
- Dans toutes les circonstances, les faits et circonstances environnants, y compris toute absence de réponse à des questions ou à des demandes d'information comme une demande de DAPC ou de transmission d'une explication au titre de l'article 86 du règlement sur l'IA, permettent-ils de tirer une conclusion de discrimination ?

Exemples de la CJUE

- C-109/88, Handels- og Kontorfunktionærernes Forbund I Danmark c. Danfoss : un syndicat a intenté une action au nom de travailleuses d'une entreprise au motif qu'elles percevaient, en moyenne, une rémunération inférieure de 7 % à celles de leurs collègues masculins occupant le même poste ou un poste similaire. Dans cette affaire, la CJUE a jugé que, dès lors que le système de rémunération utilisé était totalement dépourvu de transparence, une présomption de discrimination devait être retenue et la charge de la preuve devait être renversée, la défense devant démontrer que l'écart de rémunération reposait sur des motifs objectifs et non discriminatoires.
- C-415/10 Galina Meister c. Speech Design Carrier Systems GmbH : la personne ayant déposé la plainte a invoqué une discrimination fondée sur le genre, l'âge et l'origine ethnique dans le processus de recrutement et a demandé au tribunal d'ordonner à l'employeur de rendre publique l'information indiquant si un autre candidat était employé à la fin du processus. La CJUE a considéré que, bien que les directives en matière de non-discrimination ne prévoient pas le droit d'accès à ce type d'information, le refus de la partie défenderesse de donner tout accès à l'information peut constituer un des éléments à prendre en compte pour établir les faits permettant de présumer l'existence d'une discrimination directe ou indirecte.

Exemples des conséquences du manque de transparence de la jurisprudence nationale

- **Pays-Bas affaire NL24.38560 (2025) (arrêt interlocutoire)**

La juridiction a décidé de sa propre initiative d'examiner l'existence d'un algorithme utilisé dans le traitement des demandes de visa et a chargé le ministère compétent de :

- fournir des informations pertinentes pour la prise de décision ;
- réaliser une AIDH ;

pour des raisons de transparence du processus décisionnel et du droit à une protection juridique effective au titre de l'article 13 de la Convention européenne des droits de l'homme et de l'article 47 de la Charte des droits fondamentaux de l'Union européenne.

- **Affaire du Tribunal fiscal britannique (premier niveau) : Elsbury c. The Information Commissioner (2025) UKFTT 915 (GRC) (2 août 2025)**

  > « L'incapacité de HMRC (Services des impôts britannique) à confirmer ou à nier qu'elle détient les informations demandées renforce la croyance fondée sur des indicateurs dans la correspondance de HMRC traitant des réclamations de recherche et développement (R&D) que l'IA est utilisée par les agent·es de HMRC – peut-être de manière non autorisée – sapant ainsi la confiance des contribuables dans le traitement des réclamations par HMRC, décourageant ainsi les personnes en demandes légitimes de faire des réclamations, entravant ainsi les objectifs politiques du régime d'allègement fiscal de R&D lui-même. »

- ***Filcams Cgil Bologna et al. c. Deliveroo Italia S.R. (2020) :***

  Le Tribunal de Bologne a constaté que l'entreprise ne s'était pas acquittée de la charge de la preuve, car elle n'avait ni démontré ni expliqué le fonctionnement concret de l'algorithme et n'avait fourni aucune preuve de son fonctionnement effectif. Elle ne précisait pas quels critères spécifiques étaient utilisés pour déterminer les statistiques des coursiers et coursières, et celles-ci n'étaient pas divulguées sur la plateforme, qui ne faisait référence qu'à la « fiabilité » et à la « participation » comme indicateurs de l'attribution des futures opportunités de travail. Pour plus de détails sur cette affaire, voir Purificato (2021).

3.8. L'organisation intimée peut-elle démontrer que la différence de traitement n'est pas discriminatoire ?

Dans la pratique, vous devrez probablement examiner les mêmes preuves et poser les mêmes questions que celles énoncées ci-dessus (3.1-3.8). Cependant, il est possible qu'à ce stade, des éléments de preuve supplémentaires puissent désormais être examinés ou que l'explication de l'organisation intimée sur les éléments de preuve ou le système soit plus détaillée. Cela nécessitera un examen des documents originaux et une analyse effectuée antérieurement pour s'assurer qu'ils demeurent valides à la lumière des nouveaux renseignements.

Une fois que la personne qui allègue la discrimination a établi une preuve prima facie de discrimination, la charge de la preuve est renversée et il incombe alors

à l'organisation intimée de démontrer que la différence de traitement n'est pas discriminatoire. L'organisation intimée peut procéder comme suit.

- Démontrer qu'il n'y a en fait aucune différence de traitement, parce que, par exemple, le groupe de comparaison est mal choisi ou que le fondement de la décision n'est pas le motif protégé, ou que les faits et les circonstances de l'affaire sont différents de ceux du groupe de comparaison.
- Prouver qu'il n'y avait pas de lien de causalité entre le motif de distinction interdite et la différence de traitement.
- En démontrant que, bien que la différence de traitement soit liée au motif illicite, elle peut être justifiée au titre de l'une des exceptions énumérées dans le droit de l'Union (telles que les exigences professionnelles réelles et déterminantes).
- Prouver que la mesure, la loi ou la pratique adoptée est appropriée et nécessaire pour atteindre l'objectif légitime pertinent et que les désavantages causés n'étaient pas disproportionnés par rapport aux objectifs poursuivis (justification raisonnable et objective), conformément à la Convention européenne des droits de l'homme pour la discrimination directe et indirecte et au droit de l'Union en cas de discrimination indirecte.

Si l'organisation intimée n'est pas en mesure de prouver l'un de ces éléments, sa responsabilité au titre de la discrimination sera engagée.

Par conséquent, dans le contexte des systèmes algorithmiques ou d'IA, une fois qu'une preuve prima facie est établie, cela exigerait que l'organisation intimée divulgue des informations pertinentes sur le système. Ces informations pourraient viser à démontrer que les choix effectués étaient conformes à des critères objectifs et non discriminatoires. Pour connaître la pertinence de la transmission d'informations statistiques pour réfuter la charge de la preuve, voir 3.6 approche des données statistiques .

Tableau 7. Directives relatives à la non-discrimination et le renversement de la charge de la preuve, avec des exemples de l'approche adoptée par la CJUE pour tirer des conclusions de la non-divulgation d'éléments de preuve

Directive	Considérants/ articles	Motifs protégés	Champ d'application	Exemple de cas montrant l'approche de la CJUE pour tirer des conclusions de l'omission d'une organisation intimée de divulguer des informations pertinentes ou nécessaires et/ou toute autre exigence de transparence
Directive 2000/43/CE (directive sur l'égalité I)	Considérants 15 et 21	« Race » et origine ethnique – pas nationalité et sans préjudice du traitement découlant du statut juridique des personnes ressortissantes de pays tiers et des personnes apatrides	Emploi, formation professionnelle, protection sociale, éducation, biens et services, y compris logement	Meister C-415/10 : le refus de transmettre des informations (sur les critères de recrutement, par exemple) peut donner lieu à une conclusion de discrimination
Directive 2000/78/CE (directive sur l'égalité II)	Considérant 31 Article 10	Religion ou croyance, handicap, âge, orientation sexuelle	Emploi et profession	Kelly C-104/10 : aucun droit aux données, mais le défaut de divulgation peut permettre aux tribunaux de tirer des conclusions
Directive 2006/54/CE (refonte de la directive sur l'égalité entre les hommes et les femmes)	Considérant 30 Article 19	Sexe (y compris grossesse et maternité)	Emploi et profession (y compris les régimes de sécurité sociale)	Danfoss C109/88 : lorsque le système salarial est opaque, la charge est transférée à l'employeur Enderby C-127/92 : les justifications des écarts de rémunération peuvent avoir à satisfaire au critère de justification objective Kenny C-427/11 : les données relatives aux rémunérations des entreprises comparables peuvent être exigées pour réfuter la preuve prima facie

Directive	Considérants/ articles	Motifs protégés	Champ d'application	Exemple de cas montrant l'approche de la CJUE pour tirer des conclusions de l'omission d'une organisation intimée de divulguer des informations pertinentes ou nécessaires et/ou toute autre exigence de transparence
Directive 2004/113/ CE (directive genre, biens et services)	Considérant 22 Article 9	Sexe	Biens et services	Test-achats C-236/09 Voir aussi l'interdiction spécifique de l'utilisation du sexe comme facteur dans le calcul des primes d'assurance et des services financiers connexes, examen des facteurs actuariels à l'article 5
Directive 2010/41/UE (ddirective sur le travail indépendant sans distinction de genre)		Sexe	Travailleur·ses indépendant·es, conjoint·es ou partenaires de vie de travailleur·es indépendant·es	(Affaire en cours) Loredas C-531/23 : relative à la légalité des dérogations à la tenue de registres du temps de travail
2023/970/E Directive sur l'égalité des salaires (transposition d'ici juin 2026)	Articles 5, 6 et 8-9		Emploi	Offre un droit direct à la transparence. Le chapitre II prévoit un droit à la transparence des salaires (article 5), une politique de fixation et de progression des salaires (article 6), un droit à l'information, qui doit être accessible (articles 8 et 9) et un rapport sur les écarts de rémunération (articles 9 et 10)

Existe-t-il une justification objective ?

Pour démontrer qu'une différence de traitement n'est pas discriminatoire, la partie défenderesse peut, dans un cas de discrimination indirecte[57], soutenir que la pratique est objectivement justifiée. Cette section met en évidence des points pour évaluer une justification objective avancée par une organisation intimée.

Pour justifier une différence de traitement, il faut démontrer que :

- le système algorithmique a un objectif légitime ;

57 En vertu du droit de l'Union, la justification n'est généralement pas disponible dans les cas de discrimination directe, sauf dans les cas liés à l'âge/au handicap et dans les cas d'occupation réelle.

- le système algorithmique est adéquat ou approprié pour atteindre cet objectif ;
- le système algorithmique est nécessaire pour atteindre cet objectif ;
- le système algorithmique est proportionné, c'est-à-dire que le système algorithmique est efficace pour atteindre l'objectif, qu'il n'existe aucun autre moyen d'atteindre l'objectif légitime qui conduirait à une moindre différence de traitement, et que l'objectif légitime est suffisamment important pour justifier la différence de traitement.

Exemple d'affaire de la CJUE : C-167/97 Seymour-Smith

Dans une affaire où les recours pour licenciement abusif n'étaient ouverts qu'aux personnes comptant au moins deux années d'ancienneté et où des données statistiques avaient montré que cette condition entraînait une discrimination indirecte à l'égard des femmes, il incombait à l'État membre à l'origine de cette disposition, de ce critère ou de cette pratique apparemment neutre de démontrer que cette règle répondait à un objectif légitime de sa politique sociale, que cet objectif était étranger à toute discrimination fondée sur le sexe et qu'il pouvait raisonnablement considérer que les moyens retenus étaient aptes à atteindre cet objectif. Dans cette affaire, la Cour a estimé que de « simples affirmations générales concernant l'aptitude d'une mesure déterminée à promouvoir l'embauche ne suffisent pas pour faire apparaître que l'objectif de la règle litigieuse est étranger à toute discrimination fondée sur le sexe ni à fournir des éléments permettant raisonnablement d'estimer que les moyens choisis étaient aptes à la réalisation de cet objectif » (paragaphe 76). La Cour a également relevé que les statistiques ultérieures à l'adoption de la mesure pourraient être prises en compte pour fournir un indicateur de l'incidence de la mesure sur les femmes et les hommes (paragraphe 49).

Objectif du système

Argumenter une justification raisonnable et objective commence par un but légitime. Ici, il peut être utile de noter si l'organisation intimée a mentionné des objectifs différents dans divers documents ; voir : 3.4 À quoi sert le système algorithmique ?

Prenez note des différents types d'objectifs :

- objectifs de niveau inférieur (prédire la fraude, déterminer si un·e étudiant·e est devant son ordinateur, prédire la probabilité de défaut de prêt) ;
- des objectifs de niveau supérieur (réduire au maximum les dépenses publiques consacrées aux prestations versées aux personnes qui ne remplissent pas les conditions requises, surveiller les examens en ligne, réduire au maximum le taux global de défaut de paiement).

Il peut également être nécessaire d'examiner le cadre juridique interne dans lequel le système algorithmique a été déployé. Voir Glukhin c. Russie, requête n° 11519/20 (4 juillet 2023) (ci-dessous)[58].

Adéquation ou pertinence

L'organisation intimée devrait démontrer que le système algorithmique peut atteindre l'objectif. Cela signifie que lorsque l'objectif est d'améliorer l'efficacité d'un processus décisionnel, l'organisation intimée devrait démontrer que l'algorithme a permis de réduire le temps de traitement, de diminuer les ressources nécessaires ou d'augmenter l'efficacité du processus. Dans le cas d'un système prédictif, démontrer l'adéquation pourrait signifier établir la précision prédictive du système. Il convient de noter que les plaintes en matière d'adéquation du système doivent correspondre à l'objectif qui lui est assigné. Elles devraient tenir compte à la fois d'objectifs plus larges et de plus haut niveau. Un système peut être inadapté s'il ne répond pas à l'un des objectifs suivants :

- objectifs de niveau inférieur, par exemple si le système est conçu pour optimiser quelque chose d'autre que son cas d'utilisation prévu. Par exemple, a-t-il été optimisé pour prédire les erreurs, mais est maintenant utilisé pour prédire les fraudes ? ;
- des objectifs de plus haut niveau, par exemple si l'objectif est de réduire les dépenses en prestations frauduleuses, mais que les inspections soutenues par l'algorithme sont plus coûteuses que la réduction des dépenses, alors ce système algorithmique n'est pas adapté à cet objectif spécifique.

Nécessité

- L'objectif ne peut pas être atteint (aussi efficacement) sans l'algorithme :
  - par rapport à la situation avant l'algorithme ;
  - par rapport aux approches non algorithmiques.
- L'objectif ne peut pas être atteint (aussi efficacement) avec un autre système algorithmique :
  - par rapport aux algorithmes utilisant d'autres entrées (par exemple en excluant les proxys les plus forts) ;
  - par rapport à d'autres jeux de données de développement (avec moins d'inégalités historiques ou de biais humains).

Exemple d'affaire Cour européenne des droits de l'homme : Glukhin c. Russie, requête n° 11519/20 (4 juillet 2023)

La Cour européenne des droits de l'homme a examiné l'utilisation de la technologie de reconnaissance faciale en direct (LFR) pour trouver et localiser un manifestant dans le métro russe en vertu de l'article 8 (vie privée) et de l'article 10 (liberté d'expression). Tout en reconnaissant l'importance des techniques scientifiques modernes d'investigation et d'identification (paragraphe 85) et en constatant l'absence de base juridique pour les mesures prises en raison de l'absence de

58. Voir aussi Beghal c. Royaume-Uni pour un examen plus approfondi des conditions imposées aux lois nationales pour qu'une mesure contestée soit considérée comme étant « conforme à la loi ».

limitation légale de l'utilisation de la LFR, la Cour a également évalué la « nécessité dans une société démocratique » des mesures prises, notamment à la lumière de leur « effet dissuasif » potentiel sur le droit à la liberté d'expression et de réunion (article 10). La Cour a conclu :

> 90. À la lumière de l'ensemble des considérations qui précèdent, la Cour conclut que l'utilisation d'une technologie de reconnaissance faciale très intrusive dans une situation où le requérant exerçait son droit à la liberté d'expression tel que garanti par la Convention est incompatible avec les idéaux et valeurs d'une société démocratique régie par la prééminence du droit, que la Convention est destinée à sauvegarder et à promouvoir. Le traitement des données personnelles du requérant au moyen d'une technologie de reconnaissance faciale dans le contexte d'une procédure pour infraction administrative – d'abord aux fins de l'identification de l'intéressé à partir des photographies et de la vidéo publiées sur Telegram , puis de sa localisation en vue de son arrestation alors qu'il effectuait un trajet en métro à Moscou – ne saurait être considéré comme « nécessaire dans une société démocratique ».

Proportionnalité

Dans tous les cas, la discrimination résultant de l'utilisation du système que l'organisation intimée cherche à justifier doit être proportionnée à l'objectif poursuivi. De manière générale, cela impose à la juridiction de mettre en équilibre la gravité du préjudice causé avec les autres éléments examinés dans la présente section.

Questions à envisager

- L'organisation intimée peut-elle prouver qu'elle n'a pas fait preuve de discrimination ?
- L'organisation intimée peut-elle prouver qu'elle n'a pas fait l'objet de discrimination ?
  - L'intervention humaine permet-elle de résoudre de manière satisfaisante tout problème de discrimination résiduel ? Notez que la discrimination pourrait être le résultat d'une erreur humaine plutôt que du système algorithmique.
- L'organisation intimée pourrait-elle soutenir que le mauvais élément de comparaison a été choisi ?
- L'organisation intimée peut-elle démontrer que la personne ayant déposé la plainte ne se trouve pas dans une situation semblable ou comparable à celle de son élément de comparaison ?
- Existe-t-il une autre explication durable à la différence de traitement autre que la discrimination ?
- Quelles mesures ont été prises pour éviter la discrimination (ajustements raisonnables) ?
- Le cas échéant, l'organisation intimée peut-elle démontrer que la différence de traitement n'est pas fondée sur le motif protégé, mais sur d'autres différences objectives ?
- Ces différences objectives sont-elles explicitement couvertes par les directives relatives à la non-discrimination ?

- La disposition ou la pratique en question poursuit-elle un but légitime ?
- Passez en revue la section 3.4 sur ce que le système d'IA entend faire et comment il le fait – les moyens choisis étaient-ils appropriés ?
- Les moyens employés pour atteindre l'objectif sont-ils appropriés, nécessaires et proportionnés par rapport à l'objectif recherché ?

Étape 4
Solutions

L'étape 4 détaille les différentes solutions et voies d'action dont disposent les organismes de promotion de l'égalité. Les pouvoirs et mandats sont susceptibles de différer sensiblement selon les États membres et cette étape a donc été élaborée spécifiquement pour examiner l'interrelation entre les directives relatives aux normes, qui doivent être transposées d'ici le 19 juin 2026, le règlement sur l'IA, le RGPD et tout autre droit pertinent de l'Union.

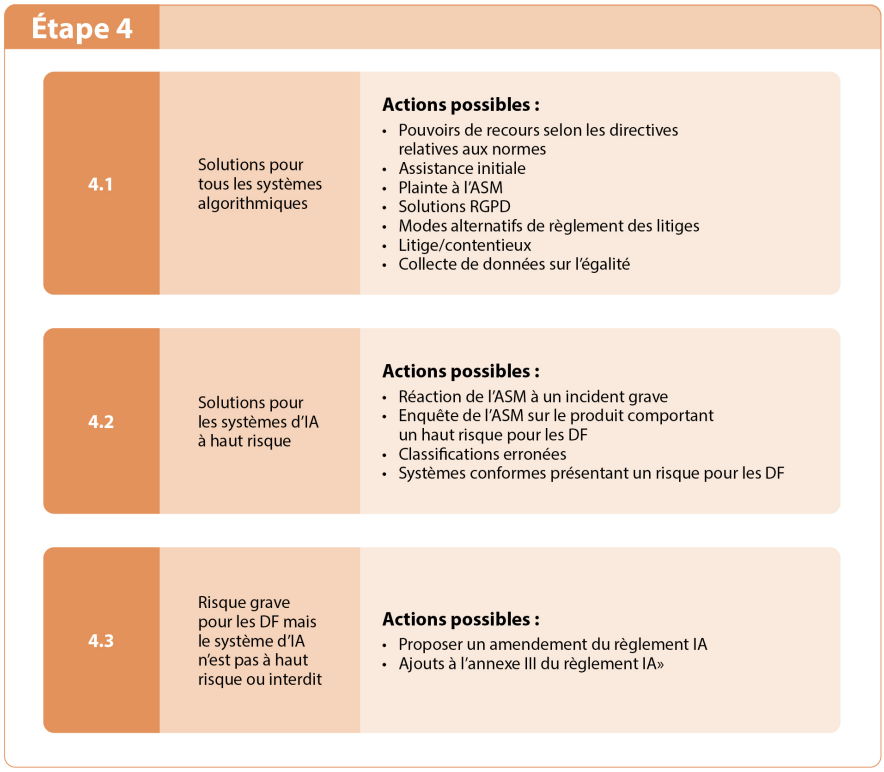
Il sera généralement nécessaire d'envisager plus d'une solution selon le problème ou l'affaire.

La section 4.1 traite de solutions pour tous les systèmes algorithmiques.

L'article 4.2 prévoit des solutions supplémentaires pour les systèmes à haut risque.

La section 4.3 décrit ce qu'un OPE devrait faire dans un cas où un système ou un cas d'utilisation présente un risque pour les droits fondamentaux, mais n'est pas classé comme étant à haut risque en vertu du règlement sur l'IA, y compris les mesures qu'une ASM peut prendre et qui nécessiteraient ou pourraient nécessiter la participation d'un OPE.

Illustration 5. Étape 4 – Solutions



En vertu des directives relatives à la non-discrimination et des directives relatives aux normes, l'OPE se voit conférer de larges pouvoirs pour mener des activités visant à prévenir la discrimination et à promouvoir l'égalité de traitement. Ces compétences comprennent la sensibilisation (article 5), l'aide aux victimes (article 6), les modes alternatifs de règlement des litiges (article 7), la conduite d'enquêtes sur une violation du principe de l'égalité de traitement (article 8), l'émission d'avis et de décisions, y compris le droit d'établir les faits (article 9) et les actions en justice, y compris le droit de présenter des observations (article 10).

Lorsqu'une discrimination a été constatée dans un système algorithmique, il convient d'envisager l'exercice de ces pouvoirs pour garantir :

- que l'organisation intimée cesse d'utiliser le système d'IA de façon discriminatoire ou, si cela n'est pas possible, cesse totalement d'utiliser le système d'IA ;
- que la connaissance de la discrimination algorithmique soit communiquée de manière appropriée aux groupes et organismes concernés ;
- que des conseils et une assistance, y compris en ce qui concerne les modes alternatifs de règlement des litiges, soient fournis aux victimes individuelles ;
- que des mesures correctives soient prises, notamment par la mise à disposition de formations, de conseils et d'une assistance ;

- qu'un débat public sur la question de la discrimination dans les systèmes algorithmiques soit engagé et que des actions positives et l'intégration de l'égalité soient encouragées parmi les entités publiques et privées ;
- que la communication avec les parties prenantes concernées, y compris les partenaires sociaux, et la promotion de l'échange de bonnes pratiques soient pérennisées.

Dans l'exercice de ces activités, les organismes de promotion de l'égalité peuvent prendre en considération des situations spécifiques de désavantage résultant d'une discrimination intersectionnelle, qui est comprise comme une discrimination fondée sur une combinaison de motifs protégés par les directives relatives à la non-discrimination.

L'OPE doit prendre en considération les outils et formats de communication appropriés pour chaque groupe cible. Il se concentre en particulier sur les groupes dont l'accès à l'information peut être entravé, par exemple en raison de leur situation économique précaire, de leur âge, de leur handicap, de leur niveau d'alphabétisation, de leur nationalité ou de leur statut de résidence, ou en raison du manque d'accès aux outils en ligne.

Ressources clés (disponibles en anglais)
- Strategic litigation handbook (Equinet 2017)
- How to use the Artificial Intelligence Act to investigate AI bias and discrimination: a guide for equality bodies (Mittelstadt 2024)

4.1. Solutions pour tous les systèmes algorithmiques

Le suivi des étapes 2 et 3 aurait dû aider à déterminer s'il existe une discrimination et d'en identifier la cause éventuelle. Les OPE devraient utiliser les informations recueillies au cours de ces phases pour suggérer des mesures correctives et adopter un plan de suppression ou de modification des pratiques, algorithmes ou systèmes contestés. Par exemple, l'OPE a peut-être constaté que le choix des paramètres d'équité était inadapté au cas d'usage considéré ou que les données n'étaient pas suffisamment représentatives.

Lorsque l'OPE décide que le risque de discrimination ou la discrimination constatée est interdite mais que celui-ci peut être corrigé, envisagez des mesures pour parvenir à :
- l'interdiction d'utiliser le système tant que la discrimination n'a pas été solutionnée ;
- toutes les mesures correctives qui devraient être prises.

Lorsque l'OPE décide que le risque de discrimination est illicite et ne peut être solutionné, il convient d'envisager de prendre l'une des mesures suivantes :
- interdiction totale de l'utilisation du système algorithmique ;
- interdiction d'utiliser le système algorithmique dans un secteur particulier ;
- interdiction d'utiliser le système algorithmique à des fins particulières.

Ces mesures pourraient être prises par les organismes de promotion de l'égalité, qui exercent leurs propres pouvoirs en matière d'avis, de recommandations et de consultations, en conjonction avec l'ASM, soit en vertu du RGPD, soit en vertu du règlement sur l'IA, soit par l'intermédiaire des juridictions.

4.1.1. Pouvoirs correctifs des organismes de promotion de l'égalité

Les organismes de promotion de l'égalité devraient d'abord examiner les pouvoirs de réparation qui leur sont conférés par le droit national, comme indiqué ci-dessus. Il pourrait s'agir de tout ou partie des mesures correctives qui doivent être formalisées par les directives relatives aux normes, telles que l'émission d'un avis ou d'une recommandation ou l'exigence d'une consultation ou de procéder à un test.

Si les directives relatives aux normes sont transposées dans les législations nationales, les OPE devraient se voir conférer des pouvoirs correctifs. Il s'agira notamment de l'article 9 (émettre un avis contraignant ou non contraignant, en fonction du contexte national) et de l'article 15 (pouvoirs de consultation et de recommandation).

Les organismes de promotion de l'égalité devraient également être habilités à fournir et documenter leur évaluation de l'affaire, y compris en établissant les faits et en concluant de manière motivée à l'existence d'une discrimination. Dans le cas de la discrimination algorithmique, cela pourrait inclure la coopération avec l'ASM.

Le cas échéant, tant les avis non contraignants que les décisions contraignantes comportent des mesures spécifiques visant à remédier à toute violation du principe de l'égalité de traitement constatée et à prévenir de nouvelles violations. L'OPE peut avoir besoin de l'assistance d'expert·es pour proposer des mesures spécifiques et, s'ils ne disposent pas de l'expertise nécessaire, ils peuvent souhaiter consulter l'ASM. Des mécanismes appropriés de suivi des avis non contraignants devraient être mis en place, comme les obligations de retours, ainsi que pour l'exécution des décisions contraignantes.

Les consultations peuvent servir à identifier l'existence de problèmes, à tester l'impact social et discriminatoire des systèmes sur les communautés touchées ou potentiellement touchées et à recueillir des preuves.

Plus généralement, les recommandations et les consultations pourraient viser à éviter que les problèmes ne se posent à l'avenir, par exemple en recommandant :

- des évaluations d'impact sur la protection des données et sur l'égalité ;
- une consultation des communautés et une AIDF dirigée par la communauté affectée ;
- l'utilisation de bacs à sable et de tests en conditions réelles.

Lectures supplémentaires

- Crowley (2023), Informing the policy agenda: equality bodies making recommendations.

4.1.2. Fournir une aide initiale à une victime

L'article 6 des directives relatives aux normes prévoit que les organismes de promotion de l'égalité fournissent une aide initiale en informant les victimes des éléments suivants :

- le cadre juridique relatif à la discrimination et aux systèmes algorithmiques, y compris les conseils ciblés sur leur situation spécifique ;
- les services offerts par l'organisme de promotion de l'égalité et les aspects procéduraux connexes ;
- les voies de recours disponibles, y compris la possibilité de poursuivre l'affaire devant les tribunaux, notamment :
  - en vertu du droit à la non-discrimination ;
  - en vertu du règlement sur l'IA – voir par exemple le droit de déposer une plainte auprès de l'ASM (article 85 du règlement sur l'IA) et le droit à une explication (article 86 du règlement sur l'IA) ;
  - en vertu du règlement sur la protection des données – par exemple, le dépôt d'une plainte auprès de l'APD (article 77 du RGPD) et le droit de faire une demande d'accès de la personne concernée (article 15 du RGPD[59].
- Les règles de confidentialité applicables et la protection des données à caractère personnel (par exemple, les politiques de protection des données de l'OPE et, le cas échéant, tout conflit entre l'obtention de preuves par l'OPE dans l'exercice d'une partie de son mandat et la possibilité de fournir ces preuves à des fins judiciaires par un particulier) ;
- la possibilité d'obtenir un soutien psychologique ou d'autres types de soutien pertinent auprès d'autres organismes ou organisations.

Les organismes de promotion de l'égalité doivent informer la personne ayant déposé la plainte, dans un délai raisonnable, de la clôture de la plainte ou de l'existence de motifs justifiant la poursuite de la plainte.

Après avoir fourni une aide initiale, les recours dont dispose une personne ayant déposé une plainte peuvent inclure de :

- déposer une plainte auprès de l'ASM (article 85 du règlement sur l'IA);
- envisager des modes alternatifs de règlement des litiges;
- plaider la discrimination devant les tribunaux nationaux.

Toute décision concernant la meilleure ligne de conduite à adopter doit être prise conjointement avec la personne ou le groupe affecté(e). Si l'OPE n'a pas la capacité, la compétence (comme les réclamations en matière de protection des données) ou les ressources nécessaires pour poursuivre la ligne de conduite recherchée par la personne concernée, l'affaire devrait être renvoyée à un autre prestataire de services juridiques compétent.

59. Un droit à une explication existe à la fois en vertu du RGPD (articles 15 et 22) et du règlement sur l'IA (article 86). À l'étape 2 et aux annexes A et B, des explications supplémentaires et des modèles pour en faire la demande sont fournis.

4.1.3. Aider une personne morale ou physique à déposer une plainte auprès de l'ASM

Lorsque l'enquête menée au titre de l'étape 3 a révélé des problèmes de conformité au titre du règlement sur l'IA, une personne physique ou morale a le droit de déposer une plainte auprès de l'ASM (article 85 du règlement sur l'IA). Cela signifie que ce droit s'étend à un organisme doté de la personnalité juridique, comme une société ou une organisation. L'OPE peut aider la personne qui dépose la plainte en lui fournissant les éléments de preuve identifiés à l'étape 3.

Ce droit est sans préjudice des droits et voies de recours existants en vertu du droit de l'Union et du droit interne et s'ajoute à ceux-ci (considérant 170). Il n'est pas limité aux systèmes d'IA à haut risque.

Toute plainte de ce type est « prise en compte » aux fins de la conduite des activités de surveillance du marché au titre du règlement (UE) 2019/1020 relatif à la surveillance du marché et à la conformité des produits. Les délais les plus susceptibles d'être fixés sont ceux de la législation nationale (voir article 11 du règlement (UE) 2019/1020).

Tout système d'IA qui ne présente pas de risque élevé relèverait du règlement sur la sécurité générale des produits (UE) 2023/988, du règlement (UE) 2023/988 et du règlement ASM (UE) 2019/1020.

> Il importe que les systèmes d'IA liés à des produits qui ne sont pas à haut risque au titre du présent règlement et qui ne sont donc pas tenus d'être conformes aux exigences énoncées pour les systèmes d'IA à haut risque soient néanmoins sûrs lorsqu'ils sont mis sur le marché ou mis en service. Pour contribuer à cet objectif, l'application du règlement (UE) 2023/988 du Parlement européen et du Conseil constituerait un filet de sécurité. (considérant 166 du règlement sur l'IA)

Le droit de déposer une plainte s'entend sans préjudice de tout autre recours administratif ou judiciaire et est donc une mesure qui devrait être prise conjointement avec tout autre recours applicable. Pris isolément, il est peu probable qu'il s'agisse de la meilleure solution pour une personne individuelle, notamment parce qu'il n'existe aucune disposition relative aux dommages et intérêts. Toutefois, il s'agit d'une mesure d'accompagnement importante et, étant donné que l'ASM dispose de pouvoirs et de recours très étendus en vertu du règlement (UE) 2019/1020, elle méritera d'être examinée dans un cas approprié. Une plainte déposée auprès de l'ASM devrait déclencher une enquête plus structurelle de la part de l'ASM et un recours plus large ayant une incidence sur la cohorte plus large de personnes touchées qu'une plainte individuelle pour discrimination. L'ASM dispose de pouvoirs étendus en vertu du règlement sur l'IA et du règlement (UE) 2019/1020 et peut effectuer des contrôles documentaires, physiques et de laboratoire (voir section 4.2).

Si la plainte est acceptée, l'ASM peut exiger qu'un opérateur économique prenne des mesures correctives, à condition que ces mesures soient appropriées et proportionnées. Si la mesure corrective n'est pas prise ou échoue, l'ASM peut interdire ou restreindre la mise à disposition d'un produit sur le marché ou ordonner le retrait ou le rappel du produit. Pour les produits présentant un risque dont les répercussions s'étendent au-delà du territoire de l'État membre de l'ASM, celle-ci est tenue d'informer immédiatement la Commission européenne et les autres

États membres de l'UE des mesures prises. Le règlement (UE) 2019/1020 prévoit de nombreuses dispositions pour la coopération transfrontalière entre les différents États membres.

4.1.4. Recours pour une personne concernée dans le cadre du RGPD ou des régimes nationaux de protection des données

En vertu du RGPD, une personne concernée (une personne physique) dispose des droits suivants en matière de données en plus du droit à une explication.

- Droit de s'opposer au traitement automatisé de ses données à caractère personnel (article 21).
- Droit de demander la rectification des données à caractère personnel inexactes la concernant et la complétion des données à caractère personnel incomplètes au moyen d'une déclaration complémentaire (article 16).
- Droit d'effacer ses données à caractère personnel sans retard injustifié lorsque l'une des conditions suivantes s'applique :
  - consent is withdrawn and no other legal basis applies;
  - they have exercised their right to object under Article 21(1);
  - the personal data have been unlawfully processed.

Elles peuvent être pertinentes lorsqu'un système algorithmique a entraîné une discrimination. Toute mesure prise peut nécessiter la participation d'un·e expert·e en protection des données ou de l'autorité chargée de la protection des données.

Lorsque les États ont adopté des régimes de protection des données en plus du RGPD ou dans des États non soumis au RGPD, il sera important de prendre en considération les lois nationales sur la protection des données.

4.1.5. Proposer des solutions alternatives de résolution des litiges conformément à la législation

En vertu de l'article 7 des directives relatives aux normes, les OPE doivent pouvoir offrir aux parties la possibilité de rechercher une solution alternative à leur litige. Ce processus peut être dirigé par l'organisme de promotion de l'égalité lui-même ou par une autre entité compétente conformément à la législation et aux pratiques nationales, auquel cas l'organisme de promotion de l'égalité peut formuler des observations à l'intention de cette entité. Ces modes alternatifs de règlement des litiges peuvent prendre différentes formes, comme la médiation ou la conciliation, conformément à la législation et aux pratique nationales. L'absence de résolution n'empêche pas les parties d'exercer leur droit d'agir dans le cadre d'une procédure judiciaire. Les États membres veillent à ce qu'il existe un délai de prescription suffisant pour garantir que les parties à un litige aient accès à une juridiction, par exemple en suspendant le délai de prescription pendant que les parties s'engagent dans un processus alternatif de règlement des litiges.

4.1.6. Non-discrimination et litiges en matière de droits humains

Les organismes de promotion de l'égalité devraient avoir le droit d'intervenir dans les procédures judiciaires en matière civile et administrative, y compris le droit de présenter des observations relatives aux directives sur la non-discrimination, conformément à la législation et aux pratiques nationales. Cela varie selon l'État.

En vertu de l'article 10 des directives relatives aux normes, le droit de l'organisme de promotion de l'égalité d'intervenir dans le cadre d'une procédure judiciaire comprend également au moins l'un des éléments suivants :

- le droit d'engager une action en justice au nom d'une ou de plusieurs victimes ;
- le droit de participer à une procédure judiciaire en faveur d'une ou de plusieurs victimes ;
- le droit d'intenter une action en justice en son nom propre, pour défendre l'intérêt public.

Ce droit comprendra les décisions d'exécution concernant les avis contraignants émis par l'OPE en vertu de l'article 9 (voir ci-dessous). Lorsqu'ils examinent s'ils doivent engager un litige stratégique, les organismes de promotion de l'égalité sont invités à se reporter aux considérations détaillées figurant dans le Manuel du litige stratégique d'Equinet (Equinet 2017: 17).

Les possibilités de recourir à un contentieux ou litige stratégique en matière de systèmes d'IA sont susceptibles d'être importantes, tant la technologie que le cadre juridique sont encore relativement récents. Le contentieux stratégique ne constituera toutefois pas nécessairement l'approche la plus appropriée dans tous les cas. Par exemple, il peut s'écouler un délai trop long avant qu'une personne concernée n'obtienne un résultat, les informations nécessaires à l'introduction d'une action susceptible d'aboutir peuvent faire défaut. Il peut être plus efficace de recourir à des procédures de mise à l'épreuve et de négociation, ou encore d'autres facteurs devront être appréciés au cas par cas.

4.1.7. Collecte de données, accès aux données sur l'égalité et établissement de rapports

L'accès aux données sur l'égalité est essentiel pour établir la discrimination par les systèmes d'IA et d'ADM. Les organismes de promotion de l'égalité peuvent tirer parti des nouveaux pouvoirs pour collecter des données sur l'occurrence sur la discrimination algorithmique de manière plus générale. Des recommandations sur les types d'informations à enregistrer ont été formulées tout au long de cette méthodologie, ce qui peut aider les OPE à compiler des données d'égalité relatives aux systèmes algorithmiques. Les statistiques recueillies pourraient être utiles pour établir un cas de discrimination prima facie devant les tribunaux, par exemple. Si l'OPE constate que des demandes d'aide sont fréquentes en rapport avec un secteur

particulier ou un cas d'utilisation particulier d'un système algorithmique, il peut exiger une enquête plus approfondie ou une priorisation.

La collecte de données sur l'égalité est insuffisante dans de nombreux pays, notamment parce qu'elle est souvent restreinte par la loi parce que les OPE sont souvent incapables de traiter des catégories sensibles de données à caractère personnel à des fins de mesure et d'atténuation des biais en vertu de l'article 9(2)(g) du RGPD et de l'article 10(5) du règlement sur l'IA. Les OPE pourrait recommander de réformer les règles et pratiques de collecte de données afin d'améliorer les connaissances existantes sur la discrimination algorithmique (voir Conseil de l'Europe 2025b).

Identifier les problèmes structurels et en rendre compte chaque année

La collecte de données devrait permettre à l'OPE d'identifier et de rendre compte des inégalités structurelles dans les systèmes algorithmiques. Lorsque l'OPE utilise une base de données pour enregistrer les détails des affaires et les conseils donnés, une nouvelle catégorie de cas devrait être ajoutée pour permettre à l'OPE de suivre l'utilisation discriminatoire réelle ou présumée des systèmes d'IA. À plus long terme, cette méthodologie pourrait être utilisée pour concevoir des champs spécifiques de la base de données afin de permettre la collecte et l'analyse de données concernant ces cas.

À titre d'inspiration, tout au long de l'étape 3 de la méthodologie, des indicateurs du type d'information qui pourrait être enregistrée dans une base de données sont identifiés à la fin de chaque section et une feuille de travail suggérée est fournie à l'annexe K.

Les problèmes structurels peuvent inclure :

- la preuve de discrimination découlant d'un type particulier de système algorithmique ou de technologie ;
- la preuve de discrimination lorsque des systèmes algorithmiques sont utilisés dans des secteurs particuliers ;
- la preuve de suspicion de discrimination dans un secteur ou un type de système algorithmique ;
- la preuve que des groupes sont affectés de manière disproportionnée par l'adoption de systèmes algorithmiques dans plusieurs secteurs ;
- les systèmes d'IA ou cas d'utilisation qui devraient être, mais ne sont pas actuellement :
  - interdits ;
  - désignés comme étant à haut risque ;
  - Des modèles à usage général avec risque systémique ;
  - listés dans l'une de ces catégories, mais présentant néanmoins un risque inacceptable de discrimination ;
- les manquements systémiques à se conformer aux dispositions pertinentes du règlement sur l'IA relatives à la non-discrimination, comme les tests de biais ou d'équité inadéquats et les pratiques inadéquates de collecte de données.

4.2. Solutions supplémentaires pour les systèmes d'IA à haut risque

4.2.1. Mesures prises par l'ASM en coopération avec l'OPE

Outre les pouvoirs conférés à l'OPE en vertu de l'article 77 du règlement sur l'IA en tant que telle, l'ASM a l'obligation de travailler avec un OPE en tant qu'autorité relevant de l'article 77 lorsqu'elle traite un « incident grave » (article 73 du règlement sur l'IA), lorsque les systèmes d'IA présentent des risques (article 79 du règlement sur l'IA) et lorsqu'elle enquête sur un système conforme qui pose un risque pour les droits fondamentaux (article 82 sur les risques pour les droits fondamentaux). Une ASM peut également avoir besoin de consulter un OPE lorsqu'elle considère un système d'IA classé à tort par un fournisseur comme étant hors haut risque (article 80 du règlement sur l'IA). Ces procédures sont discutées plus en détails ci-dessous.

Un OPE peut souhaiter renvoyer un système à l'ASM avec des preuves obtenues en vertu de cette méthodologie pour signaler un incident grave, effectuer des tests, effectuer une évaluation ou exercer des pouvoirs en vertu du règlement sur l'IA ou du Règlement relatif à l'ASM pour exiger des mesures correctives, la suspension ou le retrait du produit du marché ou d'autres mesures.

Réponse de l'ASM à une notification d'incident grave (article 73 du règlement sur l'IA)

L'ASM dispose de pouvoirs et de responsabilités importants en vertu du règlement (UE) 2019/1020 lorsqu'elle traite d'un incident grave.

L'ASM doit informer l'organisme visé à l'article 77 lorsqu'un fournisseur lui a notifié un incident grave (article 73(3) du règlement sur l'IA). Un incident grave est défini comme un incident ou un dysfonctionnement d'un système d'IA qui, par exemple, conduit directement ou indirectement à la violation d'obligations en vertu du droit de l'Union visant à protéger les droits fondamentaux, y compris le droit à la non-discrimination (article 3(49)(c) du règlement sur l'IA). Des lignes directrices seront émises par la Commission en temps voulu.

Dans de tels cas, tenez compte des points suivants.

- L'incident grave comportait-il un risque de discrimination ?
  - Dans l'affirmative, les personnes concernées devraient-elles avoir droit à un recours ?
  - Dans l'affirmative, envisagez la solution la plus efficace (par exemple, envisagez des modes alternatifs de règlement des litiges, un litige ou une enquête).
  - Si les personnes concernées ont droit à une indemnisation, veillez à ce que la solution choisie ne l'empêche pas. Veillez à respecter les délais applicables.
- Cette solution est-elle susceptible d'avoir des répercussions supplémentaires et continues pour d'autres personnes ? Par exemple, le système est-il utilisé par les banques pour déterminer la solvabilité ?

- Prenez immédiatement des mesures de sensibilisation à la problématique (article 5 des directives relatives aux normes ou par le biais de l'ASM en s'appuyant sur les réglementations de sécurité des produits).
- Examinez si le système d'IA interagit avec d'autres dispositifs/logiciels et sensibilisez à tout système affecté de la même manière.

Consignez le type de technologie et le secteur dans lesquels la question a été soulevée pour une utilisation dans les rapports de l'OPE et l'analyse des données sur l'égalité.

L'action corrective requise dépendra des problèmes identifiés. Toutefois, les informations, les étapes et les approches décrites à l'étape 3 devraient aider l'OPE à déterminer les solutions possibles

Enquête de l'ASM en vertu de l'article 79(3) du règlement sur l'IA : produit présentant un risque grave pour les droits fondamentaux, y compris en matière de discrimination

Lorsque l'ASM (seule ou en coopération avec un organisme relevant de l'article 77) constate que le système d'IA ne satisfait pas aux exigences énoncées dans le règlement, elle exige, sans retard injustifié :

- des actions correctives appropriées pour mettre le système d'IA en conformité ;
- d'informer la Commission et les autres États membres des résultats de l'évaluation et de la mesure que le fournisseur doit prendre ;
- le retrait du marché du système d'IA (article 79(5) du règlement sur l'IA) ;
- un rappel dans un délai que l'ASM peut prescrire, dans les 15 jours ouvrables ou comme le prévoit la législation de l'Union (à ce stade, la législation pertinente devra être vérifiée) ;
- que l'ASM informe l'organisme notifié concerné.

Procédure de l'ASM pour traiter les systèmes d'IA classés par le fournisseur comme ne présentant pas de haut risque en application de l'annexe III (articles 6(4) et 80 du règlement sur l'IA)

Si un système d'IA a été classé à tort dans le cadre de la procédure d'auto-évaluation (article 6(4) et procédure de l'article 80), l'ASM doit procéder à une évaluation pour déterminer si l'article 6(3), s'applique.

Dans ce cas, il est possible qu'il n'y ait pas d'évaluation du risque fondamental pertinente et qu'il n'y ait pas de prise en compte de la discrimination. Les questions examinées sont de savoir si le système d'IA :

- effectue une tâche procédurale précise ;
- améliore le résultat d'une activité précédemment réalisée par un être humain ;
- détecte les schémas décisionnels ou les écarts par rapport aux schémas décisionnels antérieurs ;

- remplit une tâche préparatoire à une évaluation.

Toutefois, il est possible que même ces systèmes présentent des risques pour les droits fondamentaux et la non-discrimination. L'OPE peut renvoyer un système à l'ASM pour déterminer si le système d'IA a été mal classifié et fournir les preuves sur lesquelles fonder ce renvoi.

Systèmes d'IA à haut risque conformes au règlement sur l'IA qui présentent un risque pour la santé ou la sécurité des personnes, pour les droits fondamentaux ou pour d'autres aspects de la protection de l'intérêt public (article 82 du règlement sur l'IA)

L'ASM a des pouvoirs à l'égard d'un système conforme qui présente un risque en vertu de l'article 82 du règlement sur l'IA. Avant de prendre ces mesures, l'ASM doit avoir fait ce qui suit :

- avoir effectué une évaluation en vertu de l'article 79 du règlement sur l'IA ; et
- avoir consulté l'organisme compétent relevant de l'article 77.

L'opérateur sera tenu de prendre « toutes les mesures appropriées pour s'assurer que le système d'IA concerné, lorsqu'il est mis sur le marché ou mis en service, ne présente plus ce risque ».

Bien que cela ne soit pas explicitement indiqué du règlement sur l'IA, l'OPE devrait être consulté au moins lorsque l'ASM cherche à déterminer si les risques pour les droits fondamentaux et/ou l'égalité ont été éliminés avec succès.

4.3. Que faire lorsqu'un système ou un cas d'utilisation présente un risque pour les droits fondamentaux mais n'est pas désigné comme étant « à haut risque » ou « interdit »

Cette situation peut survenir lorsque :

- un fournisseur a auto-évalué son système algorithmique comme n'étant pas un système d'IA conformément à la définition de l'article 3 et qu'il échappe, de ce fait, au champ d'application du règlement ;
- un système algorithmique qui devrait être interdit en raison du risque qu'il présente ne figure pas sur la liste des systèmes interdits figurant à l'article 5 du règlement sur l'IA ;
- un système algorithmique qui devrait être classé comme présentant un risque élevé n'est pas inscrit à l'annexe III du règlement sur l'IA parce que ni le secteur ni le cas d'utilisation n'y sont mentionnés ;
- un fournisseur a appliqué la dérogation prévue à l'article 6(3) à un système d'IA qui, en l'absence de cette dérogation, serait classé comme étant à haut risque.

Remarque : les déployeurs ont l'obligation de surveiller leurs systèmes à haut risque et d'informer l'ASM lorsqu'ils ont des raisons de considérer que l'utilisation du système d'IA conformément aux instructions peut entraîner un risque grave pour les droits fondamentaux (article 26 conjointement avec l'article 79 du règlement sur l'IA).

Remarque : si un système d'IA est un « produit présentant un risque » au sens de l'article 3 du point 19 du règlement (UE) 2019/1020, l'ASM est tenue d'évaluer le système et, lorsque des risques pour les droits fondamentaux sont constatés, de coopérer avec les autorités compétentes relevant de l'article 77. Les solutions pour un produit de ce type en vertu du règlement sur l'IA sont limitées aux systèmes conformes à haut risque. D'autres solutions peuvent être prévues en vertu du règlement relatif à l'ASM et de la législation générale sur la sécurité des produits, du droit des contrats ou de la responsabilité délictuelle, du droit à la non-discrimination ou du droit relatif aux droits humains.

Pour les personnes ou les groupes affectés, le fait qu'un système ne soit pas désigné comme étant à haut risque ou interdit ne signifie pas qu'il ne peut pas causer de discrimination ou entraîner ou être utilisé d'une manière qui viole les droits fondamentaux. L'OPE dispose de tous les recours habituels en cas de discrimination en vertu des directives relatives à la non-discrimination et du droit national.

4.3.1. Demander une modification du règlement sur l'IA

Si l'OPE estime qu'une pratique de l'IA contrevient fondamentalement aux valeurs de l'Union, à savoir le respect de la dignité humaine, de la liberté, de l'égalité, de la démocratie, de l'État de droit et des droits fondamentaux consacrés dans la Charte, des démarches devraient être faites auprès de l'ASM et de la Commission afin de demander une modification du règlement sur l'IA. Le règlement sur l'IA prévoit une procédure permettant de modifier ses dispositions dans différentes circonstances.

L'article 112 impose à la Commission d'évaluer chaque année la nécessité de modifier la liste figurant à l'annexe III (systèmes d'IA à haut risque) et à l'article 5 (pratiques interdites en matière d'IA) et de présenter les conclusions de cette évaluation au Parlement européen. Il est donc important de faire part à la Commission de toute préoccupation concernant les systèmes d'IA qui devraient être interdits ou les systèmes d'IA qui présentent un risque élevé dans des domaines ou des cas d'utilisation qui ne figurent pas à l'annexe III du règlement sur l'IA, conformément aux procédures prévues par ce dernier.

Pendant les cinq premières années du règlement sur l'IA (jusqu'au 1er août 2029 et automatiquement prorogée par la suite, sauf opposition), la Commission a le pouvoir d'adopter des actes délégués dans les circonstances suivantes[60].

Article 6(6)	Pour ajouter, modifier ou supprimer des conditions d'auto-évaluation comme n'étant pas à haut risque
Article 7	Modifier l'annexe III en ajoutant ou en modifiant des cas d'utilisation (mais pas des zones) et en retirant des systèmes de la liste

60 La Commission dispose de pouvoirs délégués en ce qui concerne la GPAI, mais cela n'entre pas dans le cadre du présent document.

Article 11(3)	Modifier la documentation technique requise par l'annexe IV (en particulier à la lumière du progrès technique)
Article 47(5)	Mettre à jour le contenu de la déclaration UE de conformité figurant à l'annexe V

Bien que la Commission ne soit pas habilitée à adopter un acte délégué pour modifier les rubriques de l'annexe III, elle est tenue d'évaluer la nécessité de modifier de telles rubriques existantes ou d'ajouter de nouvelles rubriques de l'annexe III, de modifier les listes de systèmes nécessitant une transparence accrue et de modifier le système de surveillance/gouvernance avant le 2 août 2028. La Commission est ensuite tenue de soumettre son évaluation au Parlement européen et au Conseil de l'Union européenne.

Le Bureau de l'IA élaborera une méthodologie pour l'évaluation de ces critères.

4.3.2. Système utilisé dans un secteur à haut risque avec un cas d'utilisation non enregistré à l'annexe III du règlement sur l'IA

Si un « cas d'utilisation » particulier ne figure pas à l'annexe III du règlement sur l'IA, la Commission a le pouvoir, en vertu de l'article 7, d'amender le règlement en ajoutant ou en modifiant des cas d'utilisation lorsque :

- le système est destiné à être utilisé dans une zone visée à l'annexe III ; et
- l'incidence négative sur les droits fondamentaux (y compris la discrimination) est équivalente ou supérieure au risque de préjudice ou d'incidence négative présenté par les systèmes à haut risque figurant déjà à l'annexe III.

Par conséquent, si la Commission estime qu'un cas d'utilisation répondant à ces critères a un effet négatif sur les droits fondamentaux mais ne figure pas sur les listes de l'annexe III, envisagez de le signaler à la Commission pour qu'il soit ajouté à la liste en vertu de l'article 7 du règlement sur l'IA.

Les facteurs à prendre en considération par la Commission sont énoncés à l'article 7(2) du règlement sur l'IA et devraient être pris en compte lors de l'établissement d'un rapport :

- (a) la finalité prévue du système d'IA ;
- (b) la mesure dans laquelle un système d'IA a été utilisé ou est susceptible d'être utilisé ;
- (c) la nature et la quantité de données traitées et utilisées par le système d'IA, en particulier si des catégories particulières de données à caractère personnel sont traitées ;
- (d) la mesure dans laquelle le système d'IA agit de manière autonome et la possibilité pour une personne de passer outre une décision ou des recommandations susceptibles de causer un préjudice potentiel ;
- (e) la mesure dans laquelle l'utilisation d'un système d'IA a déjà causé un préjudice à la santé et à la sécurité, a eu une incidence négative sur les droits fondamentaux ou a suscité des préoccupations importantes quant à la probabilité d'un tel préjudice ou de cette incidence négative, comme en

témoignent, par exemple, les rapports ou allégations documentées présentés aux autorités nationales compétentes ou, le cas échéant, d'autres rapports ;

- (f) l'étendue potentielle d'un tel préjudice ou d'un tel impact négatif, en particulier en termes d'intensité et de capacité à affecter plusieurs personnes ou à affecter de manière disproportionnée un groupe particulier de personnes ;
- (g) 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, notamment pour des raisons pratiques ou juridiques en conséquence desquelles il n'est pas raisonnablement possible de s'affranchir de ces résultats ;
- (h) la mesure dans laquelle il y a un déséquilibre des pouvoirs, où les personnes ayant potentiellement subi un préjudice ou une incidence négative se trouvent dans une situation de vulnérabilité par rapport au déployeur d'un système d'IA, notamment en raison du statut, de l'autorité, de connaissances, de circonstances économiques ou sociales ou de l'âge ;
- (i) la mesure dans laquelle les résultats obtenus avec un système d'IA sont facilement corrigibles ou réversibles, compte tenu des solutions techniques disponibles pour les corriger ou les inverser, de sorte que les résultats ayant une incidence négative sur la santé, la sécurité ou les droits fondamentaux ne sont pas considérés comme facilement corrigibles ou réversibles ;
- (j) l'ampleur et la probabilité d'avantages du déploiement du système d'IA pour des individus, des groupes ou la société dans son ensemble, y compris les améliorations possibles de la sécurité des produits ;
- (k) la mesure dans laquelle le droit de l'Union en vigueur prévoit :
  - (i) des mesures de réparation efficaces en rapport avec les risques présentés par un système d'IA, à l'exclusion des demandes de dommages et intérêts ;
  - (ii) des mesures efficaces visant à prévenir ou à réduire sensiblement ces risques.

Références

Littérature

Adams-Prassl J., Binns R. et Kelly-Lyth A. (2022), Directly discriminatory algorithms, The Modern Law Review, 86(1), pp. 144-175. Disponible en anglais en suivant ce lien : https://doi.org/10.1111/1468-2230.12759 (consulté le 17 octobre 2025).

Ahmed M. (2020), UK Passport photo Checker shows bias against dark-skin Women, BBC News. Disponible en anglais en suivant ce lien : www.bbc.co.uk/news/technology-54349538.amp (consulté le 13 octobre 2025).

Allen R. et Masters D. (2019), Artificial intelligence: the right to protection from discrimination caused by algorithms, machine learning and automated decision-making, Journal of the Academy of European Law. DOI: 10.1007/s12027-019-00582-w (consulté le 19 juin 2025).

Allen R. et Masters D. (2020), Regulating for an equal AI: a new role for equality bodies, Equinet. Disponible en anglais en suivant ce lien : https://equineteurope.org/publications/regulating-for-an-equal-ai-a-new-role-for-equality-bodies/?__cf_chl_f_tk=3KwuXAuyY1rQSy6KKjf6WQaiw8Fzq0tVZpxUk2K39Qs-1782916549-1.0.1.1-nTAmM4flODMeHv8UOjuY4hye5_yL66T80oAESRgxorA (consulté le 20 octobre 2025).

Algorithm Audit (2024a), Preventing prejudice. Disponible en anglais en suivant ce lien : https://algorithmaudit.eu/algoprudence/cases/aa202401_preventing-prejudice/ (consulté le 11 octobre 2025).

Algorithm Audit (2024b), Addendum to preventing prejudice. Disponible en anglais en suivant ce lien : https://algorithmaudit.eu/algoprudence/cases/aa202402_preventing-prejudice_addendum/ (consulté le 1er octobre 2025)

Algorithm Watch (2022), Automated decision-making systems and discrimination – Understanding causes, recognizing cases, supporting those affected. Disponible en anglais en suivant ce lien : https://algorithmwatch.org/en/autocheck/ (consulté le 29 septembre 2025).

Amnesty International (2021), Xenophobic machines: discrimination through unregulated use of algorithms in the Dutch childcare benefits scandal. Disponible en anglais en suivant ce lien : https://www.amnesty.org/en/documents/eur35/4686/2021/en/ (consulté le 13 octobre 2025).

Autorité de protection des données norvégienne (2023), Evaluation of the Norwegian Data Protection Authority's regulatory sandbox for artificial intelligence. Disponible en anglais en suivant ce lien : https://www.datatilsynet.no/en/news/aktuelle-nyheter-2023/the-sandbox-has-been-evaluated/ (consulté le 17 octobre 2025).

Autoriteit Persoonsgegevens (2024), Final recommendation on supervision of AI: sector and centrally coordinated. Disponible en anglais en suivant ce lien : www.autoriteitpersoonsgegevens.nl/en/current/final-recommendation-on-supervision-of-ai-sector-and-centrally-coordinated (consulté le 1er octobre 2025).

Barocas S., Hardt M. and Narayanan A. (2023), Fairness and machine learning: limitations and opportunities, MIT Press, Cambridge, MA.

Barros Vale S. and Zanfir-Fortuna G. (2022), Automated decision-making under the GDPR: practical cases from courts and data protection authorities, Future of Privacy Forum. Disponible en anglais en suivant ce lien : https://fpf.org/blog/fpf-report-automated-decision-making-under-the-gdpr-a-comprehensive-case-law-analysis/ (consulté le 1er octobre 2025).

Big Brother Watch (2025), Suspicion by design. Disponible en anglais en suivant ce lien : https://bigbrotherwatch.org.uk/campaigns/welfare-data-watch/ (consulté le 10 novembre 2025).

Bird K. A., Castleman B. L. et Song Y. (2024), Are algorithms biased in education? Exploring racial bias in predicting community college student success, Journal of Policy Analysis and Management, Vol. 44 (2), disponible en anglais en suivant ce lien : https://onlinelibrary.wiley.com/doi/full/10.1002/pam.22569 (consulté le 17 octobre 2025).

Boscarol F. (2022), Costly birthplace: discriminating insurance practice, Algorithm Watch. Disponible en anglais en suivant ce lien : https://algorithmwatch.org/en/discriminating-insurance/ (consulté le 13 octobre 2025).

Ceross A. (2025), Understanding artificial intelligence: concepts, limitations, and risks, Equinet. Disponible en anglais en suivant ce lien : https://equineteurope.org/publications/understanding-artificial-intelligence-concepts-limitations-and-risks/ (consulté le 1er octobre 2025).

College voor de Rechten van de Mens (2023), Oordeelnummer 2023-111. Disponible en néerlandais en suivant ce lien : https://oordelen.mensenrechten.nl/oordeel/2023-111#kop (consulté le 14 octobre 2025).

College voor de Rechten van de Mens (CRM) (2025), Toetsingskader risicoprofilering. Disponible en néerlandais en suivant ce lien : www.mensenrechten.nl/actueel/nieuws/2025/01/28/nieuw-toetsingskader-tegen-discriminatie-door-risicoprofilering (consulté le 29 septembre 2025).

Conseil de l'Europe et Agence des droits fondamentaux de l'Union européenne (2018), Manuel de droit européen en matière de non-discrimination, 2e éd., Office des publications de l'Union européenne. ISBN : 978-92-871-9851-8. Disponible en suivant ce lien : https://fra.europa.eu/fr/publication/2018/manuel-de-droit-europeen-en-matiere-de-non-discrimination-edition-2018 (consulté le 1er juillet 2026).

Conseil de l'Europe (2025a), La protection juridique contre la discrimination algorithmique en Europe : cadres actuels et lacunes subsistantes. Disponible en suivant ce lien : https://rm.coe.int/legal-protection-fra-web/48802a38d9 (consulté le 4 juin 2026).

Conseil de l'Europe (2025b), Les 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.

Disponible en suivant ce lien : https://rm.coe.int/policy-guidelines-fra-web/48802a38d2 (consulté le 4 juin 2026).

Conseil de l'Europe (2025c), Méthodologie et modèle 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). Disponible en suivant ce lien : https://rm.coe.int/prems-002826-fra-2006-huderia-text-web-a4/48802ba7b2 (consulté le 19 juin 2025).

Crenshaw K. (1991), Mapping the margins: intersectionality, identity politics, and violence against women of color, Stanford Law Review, 43(6), pp. 1241-1299. Disponible en anglais en suivant ce lien : www.jstor.org/stable/1229039 (consulté le 1er octobre 2025).

Crenshaw K. (2015), Why intersectionality can't wait, The Washington Post. Disponible en anglais en suivant ce lien : www.washingtonpost.com/news/in-theory/wp/2015/09/24/why-intersectionality-cant-wait/ (consulté le 1er octobre 2025).

Crowley N. (2023), Informing the policy agenda: equality bodies making recommendations, Equinet. Disponible en anglais en suivant ce lien : https://equineteurope.org/informing-the-policy-agenda-equality-bodies-making-recommendations/ (consulté le 1er octobre 2025).

Dastin J. (2018), Amazon scraps secret AI recruiting tool that showed bias against women, Reuters. Disponible en anglais en suivant ce lien : https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-airecruiting-tool%20that-showed-bias-against-women-idUSKCN1MK08G/ (consulté le 10 novembre 2025).

Dimitrova D. et De Hert P. (2024), The LED's right of access to one's data: loopholes on proper access, legality review and data protection authority accountability, New Journal of European Criminal Law, 15(1), pp. 12-32. Disponible en anglais en suivant ce lien : https://doi.org/10.1177/20322844231214484 (consulté le 17 octobre 2025).

Equinet (2017), Strategic litigation handbook. Disponible en anglais en suivant ce lien : https://equineteurope.org/publications/strategic-litigation-handbook/ (consulté le 29 septembre 2025).

Equinet (2022), Equality Bodies working on cases without an identifiable victim: Actio popularis. Disponible en anglais en suivant ce lien : https://equineteurope.org/equality-bodies-working-on-cases-without-an-identifiable-victim-actio-popularis/ (consulté le 29 septembre 2025).

Equinet (2025), Webinar series: equality by design, deliberation and oversight. Disponible en anglais en suivant ce lien : https://equineteurope.org/webinar-series-equality-by-design-empowering-equality-bodies-and-civil-society-on-the-interface-between-technical-standards-and-legal-requirements/ (consulté le 1er octobre 2025).

Erdogan I. (2023), Transparency in the face of new technologies: information rights under the law enforcement directive, Jean Monnet Network on EU Law Enforcement Working Paper Series. Disponible en anglais en suivant ce lien : https://jmn-eulen.nl/papers/events-on-ai/ (consulté le 17 octobre 2025).

Erdogan I. (2024), Diving into the iceberg: establishing transparency in AI for law enforcement, European Papers, 9(3), pp. 956-977. Disponible en anglais en suivant ce lien : www.europeanpapers.eu/e-journal/diving-into-iceberg-establishing-transparency-ai-for-law-enforcement (consulté le 17 octobre 2025).

Eticas Foundation (2024), Automating (in)justice? An adversarial audit of riscanvi. Disponible en anglais en suivant ce lien : https://www.eticasfoundation.org/automating-injustice-an-adversarial-audit-of-riscanvi/ (consulté le 13 octobre 2025).

European Center for Not-for-Profit Law et Danish Institute for Human Rights (2025), A guide to fundamental rights impact assessments (FRIA). Disponible en anglais en suivant ce lien : https://ecnl.org/publications/guide-fundamental-rights-impact-assessments-fria (consulté le 30 avril 2026).

Commission européenne (2019), Equality law review: European network of legal experts in gender equality and non-discrimination, Office des publications de l'Union européenne. Disponible en anglais en suivant ce lien : https://op.europa.eu/s/AiAN (consulté le 19 juin 2026).

Commission européenne (2021), Orientations pour améliorer la collecte et l'utilisation des données relatives à l'égalité. Disponible en suivant ce lien : https://op.europa.eu/fr/publication-detail/-/publication/a3d2cd88-0eba-11ec-b771-01aa75ed71a1 (consulté le 17 octobre 2025).

Commission européenne (2025), Soutenir la mise en œuvre de la législation sur l'IA au moyen de lignes directrices claires. Disponible en suivant ce lien : https://digital-strategy.ec.europa.eu/fr/news/supporting-implementation-ai-act-clear-guidelines (consulté le 1er mai 2025).

Cour européenne des droits de l'homme (2025), Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention. Disponible en suivant ce lien : https://ks.echr.coe.int/documents/d/echr-ks/guide_art_14_art_1_protocol_12_eng (consulté le 10 novembre 2025).

Farkas L. et O'Farrell O. (2015), Reversing the burden of proof: practical dilemmas at the European and national level, Office des publications de l'Union européenne. Disponible en anglais en suivant ce lien : https://op.europa.eu/en/publication-detail/-/publication/a763ee82-b93c-4df9-ab8c-626a660c9da8/language-en (consulté le 10 novembre 2025).

FRA (2025), Assessing high-risk artificial intelligence: fundamental rights risks. Disponible en anglais en suivant ce lien : Assessing High-risk Artificial Intelligence: Fundamental Rights Risks – European Union Agency for Fundamental Rights (consulté le 4 juin 2026).

Fredman S. (2016), Intersectional discrimination in EU gender equality and non-discrimination law, European Equality Law Network, Office des publications de l'Union européenne. Disponible en anglais en suivant ce lien : https://op.europa.eu/en/publication-detail/-/publication/d73a9221-b7c3-40f6-8414-8a48a2157a2f/language-en (consulté le 17 octobre 2025).

González Fuster G. (2020), Artificial intelligence and law enforcement. Impact on fundamental rights. Parlement européen. Disponible en anglais en suivant ce lien : https://op.europa.eu/s/AiyD (consulté le 12 novembre 2025).

Groupe de travail sur la protection des données Article 29 (2017a), Guidelines on data protection impact assessment (DPIA) and determining whether processing is likely to result in a high risk for the purposes of Regulation 2016/679. Disponible en anglais en suivant ce lien : https://ec.europa.eu/newsroom/article29/items/611236/en (consulté le 8 novembre 2025).

Groupe de travail sur la protection des données Article 29 (2017b), Opinion on some key issues of the law enforcement directive (EU 2016/680) – wp258, Commission européenne. Disponible en anglais en suivant ce lien : https://ec.europa.eu/newsroom/article29/items/610178/en (consulté le 17 octobre 2025).

Groupe de travail sur la protection des données Article 29 (2018a), Guidelines on automated individual decision-making and profiling for the purposes of Regulation 2016/679 (WP251rev.01, 17/EN), Commission européenne. Disponible en anglais en suivant ce lien : https://ec.europa.eu/newsroom/article29/items/612053/en (consulté le 10 novembre 2025).

Groupe de travail sur la protection des données Article 29 (2018b), Lignes directrices sur la transparence au sens du règlement (UE) 2016/679, Commission européenne. Disponible en suivant ce lien : https://www.edpb.europa.eu/documents/guideline/article-29-working-party-guidelines-on-transparency-under-regulation-2016679_fr (consulté le 10 novembre 2025).

Howard E. (2024), Intersectional discrimination and EU law: time to revisit Parris, Sage Journals. Disponible en anglais en suivant ce lien : https://journals.sagepub.com/doi/10.1177/13582291241285336 (consulté le 1er mai 2026).

Ilieva S. (2024), Handbook on identifying and using equality data in legal casework, Equinet. Disponible en anglais en suivant ce lien : https://equineteurope.org/publications/handbook-on-identifying-and-using-equality-data-in-legal-casework/ (consulté le 29 septembre 2025).

Information Commissioner's Office et Equality and Human Rights Commission (2023), Memorandum of understanding between the Information Commissioner and the Equality and Human Rights Commission. Disponible en anglais en suivant ce lien : https://ico.org.uk/about-the-ico/our-information/working-with-other-bodies/ (consulté le 1er octobre 2025).

Institut de recherches – Digital Human Rights Center (2025), Enlightenment 4.0 – human understanding of AI decision-making: report on the theoretical basis of the right to explanation of individual decision-making under Article 86 AI Act, Ministère fédéral du Travail et des Affaires sociales, de la Santé et de la Protection des consommateurs, République d'Autriche. Disponible en anglais en suivant ce lien : https://researchinstitute.at/en/enlightenment-4-0-project-results-published-in-english/; Website of the Ministry of Social Affairs (consulté le 29 septembre 2025).

Lighthouse Reports (2023), How we investigated France's mass profiling machine. Disponible en anglais en suivant le lien www.lighthousereports.com/methodology/how-we-investigated-frances-mass-profiling-machine/ (consulté le 19 octobre 2025).

Lighthouse Reports (2024), How we investigated Sweden's suspicion machine. Disponible en anglais en suivant ce lien : www.lighthousereports.com/methodology/sweden-ai-methodology/ (consulté le 20 octobre 2025).

Lighthouse Reports (2025), How we investigated Amsterdam's attempt to build a 'fair' fraud detection model. Disponible en anglais en suivant ce lien : www.lighthousereports.com/methodology/amsterdam-fairness/ (consulté le 19 juin 2026).

Mahieu R. L., Asghari H. et van Eeten M. (2018), Collectively exercising the right of access: individual effort, societal effect, Internet Policy Review, 7(3). Disponible en anglais en suivant ce lien : https://doi.org/10.14763/2018.3.927 (consulté le 10 novembre 2025).

Makkonen T. (2007), Quantifier les discriminations - Collecte de données et droit européen sur l'égalité, Office des publications de l'Union européenne. Disponible en suivant ce lien : https://op.europa.eu/fr/publication-detail/-/publication/7d20295d-212c-4acb-bd9f-6f67f4c7ce67 (consulté le 10 novembre 2025).

Makkonen T. (2016), European handbook on equality data: 2016 revision, Office des publications de l'Union européenne. Disponible en anglais en suivant ce lien : https://ec.europa.eu/newsroom/just/document.cfm%3Faction%3Ddisplay%26doc_id%3D43205&ved=2ahUKEwiysrmm-eeQAxXiwAIHHUBOH9MQFnoECBgQAQ&usg=AOvVaw1aG86yvlJ51QDEGbA_FR7L (consulté le 10 novembre 2025).

Mantelero A. (2024), The fundamental rights impact assessment (FRIA) in the AI Act: roots, legal obligations and key elements for a model template, Computer Law & Security Review, 54, p. 106020. Disponible en anglais en suivant ce lien : www.sciencedirect.com/science/article/pii/S0267364924000864 (consulté le10 novembre 2025).

Mittelstadt B., Wachter, S. and Russell, C. (2023), The Unfairness of Fair Machine Learning: Levelling down and strict egalitarianism by default. Disponible en anglais en suivant ce lien : https://arxiv.org/abs/2302.02404 (consulté le 17 juillet 2026).

Mittelstadt B. (2024), How to use the Artificial Intelligence Act to investigate AI bias and discrimination: a guide for equality bodies, Equinet. Disponible en anglais en suivant ce lien : https://equineteurope.org/publications/how-to-use-the-artificial-intelligence-act-to-investigate-ai-bias-and-discrimination-a-guide-for-equality-bodies/ (consulté le 1er octobre 2025).

Nations Unies (2011), Principes directeurs relatifs aux entreprises et aux droits de l'homme – Mise en œuvre du cadre de référence « protéger, respecter et réparer » des Nations Unies, Nations Unies, New York et Genève. Disponible en suivant ce lien : https://digitallibrary.un.org/record/720245?v=pdf&ln=fr (consulté le 1er octobre 2025).

Parviainen H. (2024), Challenges of direct discrimination in algorithmic recruitment: insuperable or not?*, International Journal of Comparative Labour Law and Industrial

Relations, 40, 4 décembre 2024, pp. 437-466. Disponible en anglais en suivant ce lien : https://doi.org/10.54648/ijcl2024017 (consulté le 17 octobre 2025).

Projet JuLIA (2024), Handbook: artificial intelligence, judicial decision-making and fundamental rights, JuLIA 101046631, Commission européenne (DG JUST – JUST-2021-JTRA). Disponible en anglais en suivant ce lien : https://www.julia-project.eu/resources (consulté le 1er octobre 2025).

Purificato I. (2021), Behind the scenes of Deliveroo's algorithm: in the blindness of 'frank' its discriminatory potential, Italian Labour Law e-Journal, 14(1), pp. 169–194. Disponible en anglais en suivant ce lien : doi:10.6092/issn.1561-8048/12990 (consulté le 22 juin 2026).

PWC (2024), Onderzoek misbruik uitwonendenbeurs. Disponible en néerlandais en suivant ce lien : www.tweedekamer.nl/kamerstukken/detail?id=2024D07564&did=2024D07564 (consulté le 12 novembre 2025).

Schot E. and Davidson D. (2025), Onbetrouwbaar algoritme bestempelt jongeren als toekomstig crimineel, Follow the money. Disponible en néerlandais en suivant ce lien : www.ftm.nl/artikelen/algoritme-bestempelt-jongeren-onterecht-als-toekomstig-crimineel?share=UommUbfeHZLpB371b2sHbWwyodKB4uUhl7QNw2Jm6lIWXJmeq7VOeP7i73ixLPA%3D (consulté le 13 octobre 2025).

Skiotytė G. et Sadauskaitė A. (2026), A Digital Omnibus: identifying interlinks and possible overlaps between different legal acts in the field of digital legislation to streamline tech rules, Parlement européen, Département de l'économie et de la croissance, Bruxelles. Disponible en anglais en suivant ce lien : https://www.europarl.europa.eu/thinktank/en/document/ECTI_STU(2026)772641 (consulté le 30 avril 2026).

Smith E. et Vogell H. (2022), How your shadow credit score could decide whether you get an apartment, ProPublica. Disponible en anglais en suivant ce lien : www.propublica.org/article/how-your-shadow-credit-score-could-decide-whether-you-get-an-apartment (consulté le 17 octobre 2025).

Soler Garrido J., De Nigris S., Bassani E., Sanchez I., Tatjana E., Antoine-Alexandre A. et Boulangé T. (2024), Harmonised standards for the European AI Act, Joint Research Centre, Commission européenne, disponible en anglais en suivant le lien suivant : https://publications.jrc.ec.europa.eu/repository/handle/JRC139430 (consulté le 17 octobre 2025).

The Guardian (2021), What happened when a 'wildly irrational' algorithm made crucial healthcare decisions. Disponible en anglais en suivant ce lien : www.theguardian.com/us-news/2021/jul/02/algorithm-crucial-healthcare-decisions (consulté le 17 octobre 2025).

Vogiatzoglou P. et Marquiene T. (2022), Assessment of the implementation of the Law Enforcement Directive, Parlement européen. Disponible en anglais en suivant ce lien : https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU(2022)740209 (consulté le 17 octobre 2025).

Wachter, Mittelstadt et Russell, 2020, Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. Computer Law & Security

Review, Tome 41, juillet 2021. Disponible en anglais en suivant ce lien : https://www.sciencedirect.com/science/article/abs/pii/S0267364921000406 (consulté le 1er juillet 2026).

Weerts H., Xenidis R., Tarissan F., Olsen H. P. et Pechenizkiy M. (2023), Algorithmic unfairness through the lens of EU non-discrimination law: or why the law is not a decision tree, Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, pp. 805-816.

Weerts H., Kelly-Lyth A., Binns R. et Adams-Prassl J. (2024), Unlawful proxy discrimination: a framework for challenging inherently discriminatory algorithms, The 2024 ACM Conference on Fairness, Accountability, and Transparency, pp. 1850-1860. doi:10,1145/3630106,3659010.

Xenidis R. (2021), Tuning EU equality law to algorithmic discrimination: three pathways to resilience, Maastricht Journal of European and Comparative Law, 27(6), pp. 736-758.

Zuiderveen Borgesius F. (2018), Discrimination, intelligence artificielle et décisions algorithmiques, Direction générale de la Démocratie, Conseil de l'Europe. Disponible en suivant ce lien : https://rm.coe.int/etude-sur-discrimination-intelligence-artificielle-et-decisions-algori/1680925d84 (consulté le 10 novembre 2025).

Législation

Conseil de l'Europe (1950), Convention européenne des droits de l'homme, Rome, 4 novembre 1950, Série des traités européens n° 5.

Conseil de l'Europe (2000), Protocole n° 12 à la Convention de sauvegarde des droits de l'homme et libertés fondamentales, Rome, 4 novembre 2000, Série des traités européens n° 177.

Conseil de l'Europe (2024), 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, Série des traités du Conseil de l'Europe n° 225.

Décision 2007/533/JAI du Conseil du 12 juin 2007 sur l'établissement, le fonctionnement et l'utilisation du système d'information Schengen de deuxième génération (SIS II), Journal officiel de l'Union européenne, L 205.

Directive 76/207/CEE du Conseil du 9 février 1976 relative à la mise en œuvre du principe de l'égalité de traitement entre hommes et femmes en ce qui concerne l'accès à l'emploi, à la formation et à la promotion professionnelles, et les conditions de travail, Journal officiel des Communautés européennes, JO, L 39, 14 février 1976, pp. 40-42.

Directive 97/80/CE du Conseil du 15 décembre 1997 relative à la charge de la preuve dans les cas de discrimination fondée sur le sexe, Journal officiel des Communautés européennes, JO, L 014, 20 janvier 1998, pp. 6-8.

Directive 2000/43/CE du Conseil du 29 juin 2000 relative à la mise en œuvre du principe de l'égalité de traitement entre les personnes sans distinction de race ou

d'origine ethnique, Journal officiel des Communautés européennes, JO, L 180, 19 juillet 2000, pp. 22-26.

Directive 2000/78/CE du Conseil du 27 novembre 2000 portant création d'un cadre général en faveur de l'égalité de traitement en matière d'emploi et de travail, Journal officiel des Communautés européennes, JO, L 303, 2 décembre 2000, pp. 16-22.

Directive 2004/113/CE du Conseil 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, Journal officiel de l'Union européenne, JO, L 373, 21 décembre 2004, pp. 37-43.

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). Journal officiel de l'Union européenne, JO, L 204, 26 juillet 2006, pp. 23-36.

Directive (UE) 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 (texte présentant de l'intérêt pour l'EEE), JO, L 132, 17 mai 2023, pp. 21-44.

Directive (UE) 2024/1385 du Parlement européen et du Conseil du 14 mai 2024 sur la lutte contre la violence à l'égard des femmes et la violence domestique, JO L, 2024/1385, 24 mai 2024.

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, Journal officiel de l'Union européenne, JO L, 2024/1499, 2024/1500, 29 mai 2024.

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, Journal officiel de l'Union européenne, JO, L, 2024/1500, 29 mai 2024.

Lignes directrices de la Commission européenne sur les pratiques interdites en matière d'intelligence artificielle, telles que définies par le Règlement (UE) 2024/1689) (règlement sur l'IA) C(2025) 5052 final, 29 juillet 2025.

Loi suédoise de 2008 sur la discrimination : 567

Proposition de règlement du Parlement européen et du Conseil modifiant les règlements (UE) 2016/679, (UE) 2018/1724, (UE) 2018/1725, (UE) 2023/2854 ainsi que les directives 2002/58/CE, (UE) 2022/2555 et (UE) 2022/2557 en ce qui concerne

la simplification du cadre législatif numérique, et abrogeant les règlements (UE) 2018/1807, (UE) 2019/1150 et (UE) 2022/868 ainsi que la directive (UE) 2019/1024 (règlement Omnibus numérique), COM/2025/837 final, https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:52025PC0837.

Règlement (CE) 767/2008 du Parlement européen et du Conseil du 9 juillet 2008 concernant le système d'information sur les visas (VIS) et l'échange de données entre les États membres sur les visas de court séjour (règlement VIS), Journal officiel de l'Union européenne, JO, L 218, 13 août 2008, pp. 60-81.

Règlement (UE) n° 603/2013 du Parlement européen et du Conseil du 26 juin 2013 relatif à la création d'Eurodac pour la comparaison des empreintes digitales aux fins de l'application efficace du règlement (UE) n°604/2013 établissant les critères et mécanismes de détermination de l'État membre responsable de l'examen d'une demande de protection internationale introduite dans l'un des États membres par un ressortissant de pays tiers ou un apatride et relatif aux demandes de comparaison avec les données d'Eurodac présentées par les autorités répressives des États membres et Europol à des fins répressives, et modifiant le règlement (UE) n°1077/2011 portant création d'une agence européenne pour la gestion opérationnelle des systèmes d'information à grande échelle au sein de l'espace de liberté, de sécurité et de justice (refonte), Journal officiel de l'Union européenne, JO, L 180, 29 juin 2013, pp. 1-30.

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, et abrogeant la directive 95/46/CE (règlement général sur la protection des données, RGPD) (texte présentant de l'intérêt pour l'EEE), Journal officiel de l'Union européenne, JO, L 119, 4 mai 2016, pp. 1-88.

Règlement (UE) 2019/1020 du Parlement européen et du Conseil du 20 juin 2019 sur la surveillance du marché et la conformité des produits, et modifiant la directive 2004/42/CE et les règlements (CE) n° 765/2008 et (UE) n° 305/2011 (texte présentant de l'intérêt pour l'EEE.), Journal officiel de l'Union européenne, JO, L 169, 25 juin 2019, pp. 1-44.

Règlement (UE) 2023/988 du Parlement européen et du Conseil du 10 mai 2023 relatif à la sécurité générale des produits, modifiant le règlement (UE) no 1025/2012 du Parlement européen et du Conseil et la directive (UE) 2020/1828 du Parlement européen et du Conseil, et abrogeant la directive 2001/95/CE du Parlement européen et du Conseil et la directive 87/357/CEE du Conseil (texte présentant de l'intérêt pour l'EEE), Journal officiel de l'Union européenne, JO, L 135, 23 mai 2023, pp. 1-51.

Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle et modifiant les règlements (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 et (UE) 2019/2144 et les directives 2014/90/UE, (UE) 2016/797 et (UE) 2020/1828 (règlement sur l'intelligence artificielle) (texte présentant de l'intérêt pour l'EEE), Journal officiel de l'Union européenne, JO, L, 2024/1689, 12 juillet 2024.

Règlement (UE) 2026/1744 du Parlement européen et du Conseil du 8 juillet 2026 modifiant les règlements (UE) 2024/1689, (UE) 2018/1139 et (UE) 2023/1230 en

ce qui concerne la simplification de la mise en œuvre des règles harmonisées concernant l'intelligence artificielle (train de mesures omnibus numérique sur l'IA), disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=OJ:L_202601744 (accès le 24 juillet 2026).

Résolution législative du Parlement européen du 16 juin 2026 sur la proposition de règlement du Parlement européen et du Conseil modifiant les règlements (UE) 2024/1689 et (UE) 2018/1139 en ce qui concerne la simplification de la mise en œuvre des règles harmonisées concernant l'intelligence artificielle (train de mesures omnibus numérique sur l'IA). Disponible en suivant ce lien : https://www.europarl.europa.eu/doceo/document/TA-10-2026-0198_FR.html, consulté le 25 juin 2026.

Jurisprudence

Andersen c. Région de Syddanmark, (2010), Cour de justice de l'Union européenne, Affaire C-499/08. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62008CJ0499 (consulté le 1er juillet 2026).

Association Belge des Consommateurs Test-Achats ASBL et al. c. Conseil des ministres (2011), Cour de justice de l'Union européenne, Affaire C-236/09; ECLI :UE :C:2011:100. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62009CJ0236 (consulté le 12 novembre 2025).

Beghal c. le Royaume-Uni, Requête n° 4755/16 (22 août 2016), Cour européenne des droits de l'homme, pour une discussion plus approfondie sur les exigences imposées aux lois nationales pour qu'une mesure contestée soit considérée comme « conforme à la loi », disponible en anglais en suivant ce lien : https://hudoc.echr.coe.int/eng/?i=001-166724.

CHEZ Razpredelenie Bulgaria AD c. Komisia za zashtita ot diskriminatsia (2015), Cour de justice de l'Union européenne, Affaire C-83 ; ECLI :UE :C:2015:480. Disponible en suivant ce lien https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62014CJ0083 (consulté le 10 novembre 2025).

CK c. Dun & Bradstreet Austria GmbH et Magistradt der Stadt Wien (2025), Cour de justice de l'Union européenne, Affaire C-203 ; ECLI :UE :C:2025:117. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62022CJ0203 (consulté le 7 octobre 2025).

David L. Parris c. Trinity College Dublin et al. (2016), Cour de justice de l'Union européenne , Affaire C-232 ; ECLI :C:2010:674. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62015CJ0443 (consulté le 1er mai 2026).

Dekker c. Stichting Vormingscentrum voor Jong Volwassenen (VJV-Centrum) Plus (1990), Cour de justice de l'Union européenne, Affaire C-177/88, ECLI :UE :C:1990:383. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:61988CJ0177 (consulté le 7 octobre 2025).

D.H. et al. c. République tchèque (2007), Cour européenne des droits de l'homme, Requête n° 57325/00. Disponible en suivant ce lien : https://hudoc.echr.coe.int/fre?i=001-83258 (consulté le 1er octobre 2025).

Dita Danosa c. LKB Līzings SIA (2010), Cour de justice de l'Union européenne, Affaire C-232/09 ; ECLI :UE :C:2010:674. Disponible en suivant ce lien : eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62009CJ0232 (consulté le 20 octobre 2025).

Enderby c. Frenchay Health Authority et Secrétaire d'État à la Santé (1993), Cour de justice de l'Union européenne , Affaire C-127/92;ECLI :UE :C:1993:313. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:61992CJ0127 (consulté le 1er octobre 2025).

Évaluation de la solvabilité, autorité, discrimination multiple directe, genre, langue, âge, lieu de résidence, motifs financiers, amende conditionnelle (2018), Tribunal national de la non-discrimination et de l'égalité de Finlande, session plénière (vote), 216/2017. Disponible en anglais en suivant ce lien : https://www.yvtltk.fi/en/decisions/multiple-discrimination-in-the-assessment-of-creditworthiness/ (consulté le 13 octobre 2025).

Filcams Cgil Bologna et al. c. (…) S.R.L. (2020), Tribunal de Bologne, ordonnance n° 2949/2019. Disponible en anglais en suivant ce lien : https://apps.eurofound.europa.eu/platformeconomydb/italian-court-rules-that-the-deliveroo-algorithm-discriminates-workers-106121.

Frédéric Hay c. Crédit Agricole mutuel de Charente Maritime et des Deux-Sèvres (2012), Cour de justice de l'Union européenne, Affaire C-267-12 ; ECLI :UE :C:2013:823. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62012CJ0267 (consulté le 12 novembre 2025).

Galina Meister c. Speech Design Carrier Systems GmbH (2012), Cour de justice de l'Union européenne, Affaire C-415/10 ; ECLI :UE :C:2012:217. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62010CJ0415 (consulté le 12 novembre 2025).

Glukhin c. Russie (2023), Cour européenne des droits de l'homme, Requête n° 11519/20. Disponible en suivant ce lien : https://hudoc.echr.coe.int/eng?i=001-225817 (consulté le 7 octobre 2025).

Handels- og Kontorfunktionærernes Forbund I Danmark c. Dansk Arbejdsgiverforening, agissant au nom de Danfoss (1989), Cour de justice de l'Union européenne, Affaire C109/99 ; ECLI :UE :C:1989:383. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:61988CJ0109 (consulté le 12 novembre 2025).

Helga Nimz c. Freie und Hansestadt Hamburg (1991), Cour de justice de l'Union européenne, Affaire C-184/89 ; ECLI :UE :C:1991:50. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:61989CJ0184 (consulté le 12 novembre 2025).

Hoogendijk c. Pays-Bas (2005), Cour européenne des droits de l'homme, Requête n° 58641/00. Disponible en anglais en suivant ce lien : https://hudoc.echr.coe.int/eng?i=001-68064 (consulté le 1er octobre 2025).

Inge Nolte c. Landesversicherungsanstalt Hannover (1995), Cour de justice de l'Union européenne, Affaire C-317-93 ; ECLI :UE :C:1995:438. Disponible en suivant ce lien

: https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:61993CJ0317 (consulté le 12 novembre 2025).

Ingeniørforeningen i Danmark c. Région de Syddanmark (2010), Cour de justice de l'Union européenne, Affaire C-499/08 ; ECLI :UE :C:2010:600. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62008CJ0499 (consulté le 1er octobre 2025).

Ingrid Rinner-Kühn c. FWW Spezial-Gebäudereinigung GmbH & Co. KG (1989), Cour de justice de l'Union européenne, Affaire C-171/88 ; ECLI :UE :C:1989:328. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:61989CJ0184 (consulté le 12 novembre 2025).

IX v. Wabe eV et MH Müller Handels GmbH c. MJ (2021), Cour de justice de l'Union européenne, Affaires C-804/18 et C-341/19 ; ECLI :UE :C:2021:144. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62019CC0341 (consulté le 7 octobre 2025).

James c. Eastleigh Borough Council (1990), Chambre des Lords, 2 AC 751. Disponible en anglais en suivant ce lien : https://www.bailii.org/uk/cases/UKHL/1990/6.html (consulté le 12 novembre 2025).

Jyske Finans c. Ligebehanlingsnævnet (2017), Cour de justice de l'Union européenne, C-668/15 ; ECLI :UE :C:2017:278. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62015CJ0668 (consulté le 10 novembre 2025).

Kelly c. National University of Ireland (University College, Dublin) (2011), Cour de justice de l'Union européenne, Affaire C-104/10 ; ECLI :UE :C:2011:506. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62010CJ0104 (consulté le 10 novembre 2025).

La Quadrature du Net et al. c. Premier ministre et Ministère de la Culture (2024), Cour de justice de l'Union européenne, Affaire C-470/21 ; ECLI :UE :C:2024:370. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62021CJ0470 (consulté le 14 novembre 2025).

(Loredas) HJ c. US et MU (2024), Cour de justice de l'Union européenne, Affaire C-531-23 ; ECLI :UE :C:2024:1050. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62023CJ0531 (consulté le 12 novembre 2025).

M. A. De Weerd, née Roks, et al. c. Bestuur van de Bedrijfsvereniging voor de Gezondheid, Geestelijke en Maatschappelijke Belangen et al. (1994), Cour de justice de l'Union européenne, Affaire C-343/92 ; ECLI :UE :C:1994:71. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:61992CJ0343 (consulté le 12 novembre 2025).

Manjang c. Uber Eats UK Ltd et al. (2021), Tribunal du travail, Affaire n° 3206212/2021. Disponible en anglais en suivant ce lien : www.gov.uk/employment-tribunal-decisions/mr-p-e-manjang-v-uber-eats-uk-ltd-and-others-3206212-slash-2021 (consulté le 7 octobre 2025).

Maria Kowalska c. Freie und Hansestadt Hamburg (1990), Cour de justice de l'Union européenne, C-33-89 ; ECLI :UE :C:1990:265. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:61989CJ0033 (consulté le 12 novembre 2025).

Minoo Schuch-Ghannadan c. Medizinische Universität Wien (2019), Cour de justice de l'Union européenne, C-274/18 ; ECLI :UE :C:2019:828. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62018CJ0274 (consulté le 10 novembre 2025).

N.B. c. Slovaquie (2010), Cour européenne des droits de l'homme, Requête n° 29518/10, jugement du 12 juin 2012. Disponible en anglais en suivant ce lien : https://hudoc.echr.coe.int/eng?i=001-111427 (consulté le 12 novembre 2025).

Oršuš et al. c. Croatie (2010), Cour européenne des droits de l'homme, Requête n° 15766/03, jugement du 16 mars 2010. Disponible en suivant ce lien : https://hudoc.echr.coe.int/eng?i=001-97690 (consulté le 12 novembre 2025).

PublicResourceOrg, Inc. et Right to Know CLG c. Commission européenne (2024), Cour de justice de l'Union européenne, Affaire C-588/21 ; ECLI :UE :C:2024:201. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62021CJ0588 (consulté le 17 octobre 2025).

R (sur la requête de Bridges) c. Chief Constable of South Wales Police (2020), Cour d'appel, Affaire EWCA Civ 1058. Disponible en anglais en suivant ce lien : www.bailii.org/ew/cases/EWCA/Civ/2020/1058.html (consulté le 1er octobre 2025).

R c. Secrétaire d'État à l'Emploi, au nom de Seymour-Smith et Perez (1999), Cour de justice de l'Union européenne, Affaire C-167/97 ; ECLI :UE :C:1999:60. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:61997CJ0167 (consulté le 1er octobre 2025).

Rechtbank Overijssel (2024) Indirecte discriminatie door DUO bij selectie van uitwonende studenten voor controle. Het verkregen bewijs wordt uitgesloten, Affaire : ECLI:NL:RBOVE :2024:5627. Disponible en néerlandais en suivant ce lien : ECLI :NL :RBOVE :2024 :5627, Rechtbank Overijssel, ak_23_1340 en ak_23_1341 (consulté le 1er juillet 2026).

(SCHUFA Holding) OQ c. Land Hessen (2023), Cour de justice de l'Union européenne, Affaire C-634/21 ; ECLI :UE :C:2023:957. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62021CJ0634 (consulté le 7 octobre 2025).

(Opinion de l'AG Pikamäe dans SCHUFA Holding) Opinion de l'Avocat général Pikamäe OQ c. Land Hessen (2023), Cour de justice de l'Union européenne, Affaire C-634/21 ; ECLI :UE :C:2023:220. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62021CC0634 (consulté le 7 octobre 2025).

S et Marper c. Royaume-Uni, [GC] (2008), Cour européenne des droits de l'homme, Requêtes n° 30562/04 et 30566/04, Jugement du 4 décembre 2008. Disponible en suivant ce lien : https://hudoc.echr.coe.int/fre?i=001-90052 (consulté le 12 novembre 2025).

Sylvie Beghal c. Royaume-Uni (2016), Cour européenne des droits de l'homme, Requête n° 4755/16, Décision du 14 janvier 2016. Disponible en anglais en suivant ce lien : https://hudoc.echr.coe.int/eng/#{%22itemid%22:[%22001-166724%22]} (consulté le 12 novembre 2025).

Tadao Maruko c. Versorgungsanstalt der deutschen Bühnen (2008), Cour de justice de l'Union européenne, C-267/06 ; ECLI :UE :C:2008:179. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62006CJ0267 (consulté le 12 novembre 2025).

Verli Hariev Belov c. CHEZ Elektro Balgaria AD et al. (2013), Cour de justice de l'Union européenne, Affaire C-394/11 ; ECLI :UE :C:2013:48. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62011CJ0394 (consulté le 17 octobre 2025).

VL c. Szpital Kliniczny im. dra J. Babińskiego Samodzielny Publiczny Zakład Opieki Zdrowotnej w Krakowie (2021), Cour de justice de l'Union européenne, Affaire C-16/19 ; ECLI :UE :C:2021:64, para. 48. Disponible en suivant ce lien : https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62019CJ0016 (consulté le 7 octobre 2025).

Annexe A

Modèle de demande d'accès de personne concernée (DAPC)

Remarque : une DAPC peut être rédigée en termes généraux ou de manière plus précise, chaque approche présentant ses propres avantages et ses propres risques.

Les demandes spécifiques peuvent accroître la probabilité de recevoir les informations précises recherchées, ce qui réduit le besoin de correspondance de suivi et accélère le processus. Pourtant, dans la pratique, les organisations interrogées ne fournissent souvent une réponse individualisée qu'après une demande de suivi. Des demandes définies avec précision à un stade précoce peuvent représenter une charge pour la personne concernée, qui peut ne pas connaître les activités de traitement ou la terminologie pertinentes, augmentant le risque de manque d'informations et permettant des interprétations restrictives par les responsables du traitement.

Dans le cadre de la proposition de règlement de l'Omnibus numérique[61], ce point sera formalisé. Elle stipule que les DAPC devraient être « aussi spécifiques que possible », tandis que « les demandes trop vagues et indifférenciées devraient également être considérées comme excessives », suggérant que les demandes générales et indifférenciées peuvent être rejetées comme étant excessives.

Ce modèle ne sert qu'à titre d'exemple et doit être adapté au cas particulier.

Nom, adresse de la personne concernée]
[option : numéro de téléphone et adresse e-mail de la personne concernée]

[Date de la demande]

[Informations concernant la personne responsable du traitement – adresse de la société][62]

61 Considérant 35, article 3, proposition de règlement du Parlement européen et du Conseil modifiant les règlements (UE) 2016/679, (UE) 2018/1724, (UE) 2018/1725, (UE) 2023/2854 et les directives 2002/58/ce, (UE) 2022/2555 et (UE) 2022/2557 en ce qui concerne la simplification du cadre législatif numérique, et abrogeant les règlements (UE) 2018/1807, (UE) 2019/1150, (UE) 2022/868 et la directive (UE) 2019/1024 (règlement Omnibus numérique)

62. À envoyer à la personne responsable du traitement : leurs coordonnées se trouvent généralement dans la politique de confidentialité de l'organisation ou dans la notice légale sur son site Web. Dans les cas impliquant un système d'IA, la personne responsable du traitement des données sera généralement le déployeur du système (l'organisation qui l'utilise, c'est-à-dire l'entité avec laquelle la personne concernée a interagi lors du processus concerné, comme une banque, une compagnie d'assurance, une université ou un employeur). L'organisation à but non lucratif Datenanfragen.de e.V., qui gère le site web demandetesdonnees.fr – Utilise ton droit à la vie privée. (disponible en plusieurs autres langues) tient à jour une base de données d'entreprises avec les informations de contact pertinentes pour les demandes liées au respect de la vie privée de nombreuses entreprises.

Demande d'accès de la personne concernée en vertu de l'article 15 du RGPD

À qui de droit,

J'exerce par la présente mes droits en vertu de **l'article 15 du Règlement général sur la protection des données (RGPD)** et demande confirmation que vous traitez ou non des données à caractère personnel me concernant.

Si vous traitez mes données à caractère personnel, veuillez me fournir les informations suivantes conformément à l'article 15 du RGPD.

1. Accès à mes données à caractère personnel à des fins X

- Une copie de toutes les données à caractère personnel me concernant qui ont été utilisées ou prises en compte aux fins suivantes :

  [insérer les finalités centrales à votre cas ici ; par exemple, ouvrir ou gérer un compte bancaire à mon nom / conclure, renouveler ou prolonger mon contrat de téléphonie mobile / traiter ou évaluer mon éligibilité à [insérer l'avantage spécifique] / traiter ou évaluer ma candidature / déterminer mon accès aux comptes en ligne, applications ou abonnements numériques / m'attribuer des trajets ou des ordres de livraison sur la plateforme, etc.]

2. Informations sur les finalités de traitement supplémentaires

- Des informations sur toute autre finalité pour laquelle mes données à caractère personnel ont été ou sont traitées, au-delà de celles énumérées ci-dessus ;
- Une description des catégories de données à caractère personnel concernées.

3. Divulgations à des tiers

- Les destinataires ou catégories de destinataires auxquels mes données à caractère personnel ont été divulguées, ainsi que les copies des données partagées et l'objectif du partage, ainsi que la ou les dates de divulgation.

4. Prise de décision automatisée et profilage

- Des informations indiquant si une prise de décision automatisée, visée à l'article 22, paragraphes 1 et 4 du RGPD, a été prise en relation avec mes données à caractère personnel ;
- Dans l'affirmative, veuillez fournir des informations utiles sur la logique sous-jacente, y compris les catégories de données à caractère personnel, les procédures et les principes utilisés pour obtenir un résultat spécifique ainsi que les conséquences envisagées d'un tel traitement pour moi.

5. Détails du profilage

- Des informations indiquant si un profilage a été effectué en utilisant mes données à caractère personnel ;
- Si oui, veuillez me fournir le contenu de ce profil et expliquer comment il a été créé.

Si vous transférez mes données à caractère personnel vers un pays tiers ou une organisation internationale, je demande à être informé des garanties appropriées conformément à l'article 46 du RGPD concernant le transfert.

Veuillez mettre à ma disposition les données à caractère personnel me concernant, que j'ai demandées, dans un format structuré, couramment utilisé et lisible par machine, conformément à l'article 20, paragraphe 1, du RGPD.

Ma demande inclut explicitement tous les autres services et sociétés dont vous êtes le responsable du traitement tel que défini par l'article 4, paragraphe 7, du RGPD.

Comme le prévoit l'article 12, paragraphe 3, du RGPD, vous devez me fournir les informations demandées sans retard injustifié et en tout état de cause, dans un délai d'un mois à compter de la réception de la demande. Si une prolongation de deux mois au maximum est nécessaire, veuillez m'en informer.

Selon l'article 15(3) du RGPD, vous devez répondre à cette demande sans que cela ne m'entraîne de frais.

[À titre de preuve de mon identité, je joins une copie de ma carte d'identité/mon passeport/permis de conduire, etc.]

ou

[Veuillez entrer vos données d'identification ici. Ces données incluent souvent des informations comme votre nom, votre date de naissance, votre adresse, votre adresse e-mail, etc.]

Si vous ne répondez pas à ma demande dans le délai imparti, je me réserve le droit d'intenter une action en justice contre vous et de déposer une plainte auprès de l'autorité de surveillance compétente.

Merci d'avance.

Veuillez agréer l'expression de mes salutations distinguées.

[Entrez votre nom ici]

Annexe B

Modèle de demande d'explication en vertu de l'article 86 du règlement sur l'IA

Le modèle suivant intègre une version développée par l'Institut de recherche – Digital Human Rights Center (2025b), telle que publiée sur le site Web du ministère autrichien des Affaires sociales (disponible en allemand).

[Nom, adresse de la personne affectée]

[option : numéro de téléphone et adresse e-mail de la personne affectée]

[Date de la demande]

[Informations concernant le déployeur des systèmes d'IA[63]]

Demande d'explication en vertu de l'article 86 du règlement sur l'IA

À qui de droit,

Je fais référence à la décision qui me concerne [description de la décision, par exemple le rejet d'une demande de prêt, l'évaluation des risques reçue, l'attribution d'une cote de crédit], que j'ai reçue le [date de réception de la décision].

Cette décision [ ] produit des effets juridiques à mon encontre, car elle [expliquez brièvement l'impact de la décision, par exemple le rejet d'une demande de prêt, le refus d'accès à certains services, l'augmentation des prix des assurances, etc.].

[ou] m'affecte considérablement et négativement, car elle [donnez les raisons de la défaillance, comme la discrimination fondée sur des critères de différenciation, un ciblage des annonces de paris sportifs pour des personnes ayant une dépendance au jeu].

Je fais valoir mon droit à une explication conformément à l'article 86 du règlement (UE) 2024/1689 (règlement sur l'IA) et demande une explication claire et significative

63 À envoyer au déployeur du système d'IA : l'organisation qui l'utilise – c'est-à-dire l'entité avec laquelle la personne concernée a interagi au cours du processus en question, comme une banque, une assurance, une université ou un employeur. S'il s'agit d'une autorité publique (ou d'une entité agissant en son nom), trouvez leurs coordonnées dans la partie C de la base de données de l'UE. Dans tous les autres cas, recherchez les coordonnées du déployeur dans l'interface du système d'IA, la politique de confidentialité de l'organisation ou la notice légale sur son site Web. Si un déployeur n'est pas explicitement indiqué, utilisez les coordonnées du responsable du traitement, qui sera généralement également le déployeur du système.

du rôle du système d'IA dans le processus décisionnel et des principaux éléments de la décision prise.

À cette fin, je demande les informations suivantes.

1. Rôle du système d'IA dans le processus décisionnel

- Dans quelle mesure et de quelle manière le système d'IA utilisé a-t-il contribué au processus décisionnel ?

2. Principaux éléments de la décision prise

- Quelles ont été les principales raisons de la décision, y compris les principaux critères d'influence, les données d'entrée et les décisions préliminaires qui ont influencé la décision prise ?
- Quelles autres implications la décision a-t-elle pour moi, et comment ces effets ont-ils été évalués ?
- Comment les critères qui ont influencé la décision ont-ils été pondérés et pourquoi ?

3. Examen humain et fondement [non obligatoire]

- La décision a-t-elle été prise par le système d'IA de manière entièrement automatisée ou a-t-elle été examinée par un être humain ? Si c'est le cas, quel aspect a pris cet examen exactement et qui/quel service était responsable ?
- [en option le cas échéant] Sur quelle base contractuelle ou juridique la décision est-elle fondée ?

[À titre de preuve de mon identité, je joins une copie de ma carte d'identité/mon passeport/permis de conduire, etc.]

ou

[Entrez vos données d'identification ici. Elles comprennent souvent des informations comme votre nom, votre date de naissance, votre adresse postale, votre adresse e-mail, etc.]

Je vous demande formellement de me fournir les informations demandées dans un délai de [x mois] [par écrit], gratuitement et sous une forme compréhensible. Si vous avez des questions, n'hésitez pas à me contacter [pour les demandes au format papier, utilisez les coordonnées fournies ci-dessus].

Je vous remercie par avance d'avoir traité ma demande et j'attends avec impatience de recevoir votre réponse.

Veuillez agréer l'expression de mes salutations distinguées.

[Nom de la personne concernée]

Annexe C

Exemples d'utilisation des algorithmes et de l'IA

Cette annexe comprend des exemples dans lesquels les algorithmes et les systèmes d'IA sont couramment utilisés. Elle est classée par secteur, de sorte que vous puissiez facilement déterminer l'implication algorithmique probable. Pour cette liste, nous avons utilisé les ressources développées par Allen and Masters (2020) et Algorithm Watch (2022), à propos desquelles vous trouverez plus de détails dans la liste de références.

Secteur	Exemple
Banque	Évaluation des risques-client·es
	Automatisation dans la procédure d'acceptation des client·es
	Détection de fraude ou lutte contre le blanchiment d'argent
Services de protection de l'enfance et de la jeunesse	Prévoir les futurs besoins en services de protection sociale pour un·e enfant
Éducation	Admission
	Automatisation de la répartition des étudiant·es entre écoles, universités ou programmes
	Surveillance automatisée
	Prédire ou détecter l'absentéisme scolaire et l'abandon scolaire précoce
	Détection des problèmes d'apprentissage
Emploi	Sélection de candidature ou sélection de CV
	Agences pour l'emploi prédisant la probabilité de trouver du travail
	Attribution des offres d'emploi aux personnes demandeuses d'emploi
	Affectation des personnes demandeuses d'emploi à une formation ou un coaching
	Analyse des entretiens (de départ)
Santé	Prédiction des risques de maladie
	Systèmes de diagnostic
	Prédiction du nombre de personnes âgées nécessitant des soins supplémentaires

Secteur	Exemple
Logement	Vérification automatisée de locataires potentiels
	Formulaires en ligne
	Vérification de crédit
	Annonces d'appartements sur des plateformes en ligne : quelles publicités sont affichées à qui ? (Placement automatisé)
Immigration et gestion des contrôles aux frontières	Reconnaissance faciale des personnes sur une liste de surveillance
	Automatisation de l'acceptation d'une demande de visa
	Examen automatisé des documents de voyage
Assurance	Formulaires en ligne
	Algorithmes de tarification
	Automatisation dans la procédure d'acceptation des client·es
	Traitement automatique des plaintes
	Détection de fraude
Procédure judiciaire	Aide à la rédaction de lois
Justice	Détermination de la peine suggérée en fonction des données
Police	Prévision de la perpétration de crimes dans certains quartiers ou zones (services de police prédictifs géographiques)
	Prédiction d'un comportement désordonné ou violent
	Outils d'évaluation des risques ou de profilage utilisés pour prédire le risque futur de commettre un crime, y compris la récidive, ou plus généralement pour identifier les personnes « dangereuses » ou « pertinentes » (services de police prédictifs individuels)
	Évaluation ou estimation de la probabilité qu'une personne soit l'auteur·e d'un crime existant (services de police prédictifs individuels)
	Surveillance par caméra pour détecter une activité criminelle
Allocations sociales	Formulaires en ligne
	Prédiction des fraudes ou des erreurs
	Automatisation du traitement des prestations
Impôts	Prédiction des fraudes ou des erreurs

Annexe D

Exemples de discrimination algorithmique ou de risque de discrimination

Secteur ou technologie	Algorithme de description	Description de la pratique discriminatoire	Source
Banque	« Notation » ou évaluation de la solvabilité	Le système d'ADM était soumis à des règles qui l'ont amené à décider qu'une « femme de 40 ans » n'était pas solvable.	Algorithm Watch (2022) (p. 6) (cas similaire en Finlande, disponible en anglais ici)
Éducation	Algorithme de prédiction de l'achèvement d'une formation ou d'un diplôme	Les groupes minoritaires sous-représentés se sont vu attribuer systématiquement un risque plus élevé, ainsi qu'un score de précision plus faible.	Bird, Castleman et Song (2024)
Technologie de reconnaissance faciale	Technologie capable d'identifier ou de faire correspondre des images faciales	Il existe des différences de précision entre les différentes données démographiques, comme l'âge, le genre et la couleur de peau. Ces différences pourraient conduire à des cas de discrimination. L'organisme de promotion de l'égalité néerlandais a déclaré que les organisations devaient prouver que leur technologie n'était pas discriminatoire, inversant ainsi la charge de la preuve.	College voor de Rechten van de Mens (2023)
Santé	Algorithme attribuant les « soins nécessaires » aux patient·es	Système de points arbitraires, conduisant à une répartition différente des heures dans des cas similaires.	The Guardian (2021)

Secteur ou technologie	Algorithme de description	Description de la pratique discriminatoire	Source
Logement	Algorithmes de filtrage des locataires, attribution d'un score de risque	Sélection des locataires fondée sur des caractéristiques dont le lien avec le risque n'a pas été prouvé.	Smith et Vogell (2022)
Assurance	Algorithme de tarification	La discrimination directe en tant que citoyenneté/ nationalité a eu un impact sur le prix de l'assurance automobile dans des profils identiques par ailleurs.	Boscarol (2022)
Gestion de l'immigration et du contrôle aux frontières	Révision de documents de voyage, technologie de reconnaissance faciale	Discrimination directe car la technologie a donné un résultat erroné plus souvent pour les images avec peau plus foncée que pour les images avec une peau plus claire.	Ahmed (2020)
Système judiciaire/ protection de l'enfance et de la jeunesse	Évaluation du risque de récidive	Certaines nationalités ont conduit à un score de risque plus élevé, augmentant ainsi les risques de poursuites.	Schot et Davidson (2025)
Police/ application de la loi	Évaluation du risque de récidive	Le système d'IA n'avait ni transparence, ni explicabilité, ni robustesse. Des scores de risque incohérents et plus élevés pour les personnes sans « facteurs de stabilité sociale », principalement des individus plus jeunes.	Eticas Foundation (2024)
Allocations sociales	Détection de demandes incorrectes ou frauduleuses de prestations de garde d'enfants	Le profilage racial, où les candidatures de personnes ressortissantes non néerlandaises se voyaient systématiquement attribuer des scores de risque plus élevés.	Amnesty International (2021)

Secteur ou technologie	Algorithme de description	Description de la pratique discriminatoire	Source
Allocations sociales	Système d'assistance sociale exacerbant l'exclusion	Discrimination indirecte par un critère apparemment neutre qui désavantage la démographie protégée.	Amnesty International (2023)
Détection de fraude sociale	Système de notation des risques de fraude utilisé pour identifier les bénéficiaires de prestations sociales en vue d'un examen plus approfondi	Le modèle a signalé de manière disproportionnée les femmes, les personnes migrantes, les personnes à faible revenu et les personnes n'ayant pas suivi d'études superieures comme étant suspectes.	Lighthouse Reports (2024)
Détection de fraude sociale	Système de notation des risques de fraude utilisé pour identifier les bénéficiaires de prestations sociales en vue d'un examen approfondi (CNAF)	Le système a été contesté pour avoir des effets discriminatoires directs et indirects sur les femmes, les personnes handicapées et les personnes vivant dans la pauvreté.	La Quadrature du Net (2024)

Annexe E
Proxys

Du point de vue de la science/analyse des données, des variables indirectes remplacent d'autres variables, par exemple si celles-ci ne peuvent être observées ou mesurées directement. Les variables proxys permettent de tirer des conclusions sur certaines caractéristiques sans enregistrer directement ces caractéristiques. Par exemple, « 30 ans d'expérience de travail » indique que la personne doit avoir au moins 45 ans (Algorithm Watch 2022).

Dans le contexte du droit à la non-discrimination et des systèmes algorithmiques/d'IA, les « proxys » désignent des caractéristiques, des propriétés ou des combinaisons de celles-ci qui ne sont pas elles-mêmes listées parmi les motifs de discrimination explicitement interdits dans le cadre juridique applicable, mais qui sont corrélées à ces motifs de telle sorte que leur utilisation peut conduire à une discrimination interdite.

Deux types de proxy peuvent être identifiés à cet égard.

a. Les proxys statistiques : caractéristiques (ou combinaisons de ces caractéristiques) en corrélation avec des motifs protégés par la loi et dont l'utilisation peut désavantager de manière disproportionnée les personnes appartenant à un groupe protégé. Ces proxys sont généralement associés à une discrimination indirecte. Par exemple, dans la plupart des pays européens, il existe une corrélation entre le niveau d'études et la « race » ou l'origine ethnique.
b. Les proxys inextricablement liés : une caractéristique (ou des combinaisons de ces caractéristiques) qui sont inextricablement liées à un attribut protégé, de sorte que leur utilisation reflète une pratique intrinsèquement discriminatoire. Ces proxys sont généralement associés à une discrimination directe. Par exemple, la grossesse a été considérée comme inextricablement liée au sexe (voir tableau 8 ci-dessous).

Bien qu'il reste à déterminer à quel point la correspondance entre un proxy et un groupe protégé doit être étroite pour constituer une discrimination directe, un lien inextricable n'exige pas une correspondance complète (100 %) (Parvianen 2024). Adams-Prassl, Binns et Kelly-Lyth ont estimé qu'un « degré suffisant de correspondance » était nécessaire (2023), tandis que Weerts et al. (2024) ont proposé que le proxy et le motif protégé soient « au moins fortement associés et peut-être même déterministiquement liés ». Weerts et al. notent la pertinence de la question « mais pour » : une personne vivant dans un pays où l'âge de la retraite est fondé sur le sexe pourrait raisonnablement prétendre qu'elle aurait atteint l'âge de la retraite si elle n'était pas de ce sexe ; un couple de même sexe vivant dans un pays où le mariage entre personnes de même sexe est interdit pourrait facilement affirmer qu'il serait marié s'il n'était pas de cette orientation sexuelle ; une femme pourrait raisonnablement faire valoir qu'elle ne serait pas enceinte si elle n'était pas de ce sexe ; et une femme musulmane pourrait faire valoir qu'elle ne porterait pas de voile si elle n'était pas de ce sexe et de cette religion.

L'existence d'un lien inextricable dans une affaire donnée nécessite une appréciation juridique au cas par cas, à travers laquelle des critères contextuels (plutôt que des seuils statistiques précis) prennent une importance particulière, conformément à l'approche de la CJUE dans les affaires de discrimination (voir, par exemple, IX/WABE EV et MH Müller Handels GmbH/MJ, point 78, où l'accent a été mis sur le type de signes en cause). Reportez-vous au tableau ci-dessous pour des exemples de différents types de proxy pour les caractéristiques protégées.

Tableau 8 : Exemples de proxys connus (tels que mentionnés, entre autres, dans Algorithm Watch 2022, Weerts et al. 2024, CRM 2025)

Caractéristique	Proxy pour	Type de proxy
30 ans d'expérience professionnelle (Algorithm Watch 2022)	L'individu a plus de 45 ans	Inextricablement lié
Grossesse (Dekker c. Stichting voor Jong Volwassenen (VJV-Centrum) Plus, 1990)	Sexe	Inextricablement lié
État civil (Tadao Maruko c. Versorgungsanstalt der deutschen Bühnen, 2008 ; Frédéric Hay c. Crédit Agricole mutuel de Charente Maritime et des Deux-Sèvres, 2012)	Orientation sexuelle, dans les pays qui n'autorisent pas le mariage homosexuel	Inextricablement lié, dans les pays qui n'autorisent pas le mariage homosexuel
Âge de la retraite fondé sur le sexe (James v. Eastleigh Borough Council, 1990, §10)	Sexe	Inextricablement lié
Langue ou niveau de compétence linguistique	« Race »	Proxy statistique ou inextricablement lié
Alphabétisation	« Race » (selon le pays, ce pourrait également être un proxy pour l'âge et le sexe)	Proxy statistique
Niveau d'éducation	« Race », origine ethnique	Proxy statistique
Code postal ou quartier	« Race » (malgré une ségrégation des quartiers)	Proxy statistique
Revenus	« Race », sexe	Proxy statistique
Nombre de personnes dans le ménage ou la famille proche	« Race » (sexe)	Proxy statistique
Lieu de naissance	« Race »	Proxy statistique

Caractéristique	Proxy pour	Type de proxy
Percevoir des allocations sociales essentielles	« Race »	Proxy statistique
Religion	« Race »	Proxy statistique
Noms comprenant, terminant ou commençant par une combinaison de lettres spécifique	« Race »	Proxy statistique
Absence du travail/annulations tardives/disponibilité réduite du travail (Filcams Cgil Bologna et al. c. Deliveroo Italia S.R.L, 2020)	Croyance (appartenance syndicale), sexe ou état de santé	Proxy statistique
Vêtements et couvre-chefs (port du voile, par exemple) (IX c. WABE et MH Müller Handels, 2021)	Religion et croyance	Proxy statistique (se reportant à tout signe de croyance politique, philosophique ou religieuse, par exemple) ou inextricablement lié (se reportant explicitement à des signes visibles et de taille conséquente, par exemple)

Annexe F

Liste des registres d'IA et d'algorithmes en Europe

Portée géographique	Titre	Champ d'application	Lien	Date de mise à jour
Belgique	BoSA AI Observatory	Liste de tous les projets et applications connus en IA utilisés par les institutions et agences fédéraux	https://bosa.belgium.be/fr/AI4Belgium/observatoire#anchor-2AI4Belgium/observatoire#anchor-2	
Base de données de l'UE pour les systèmes d'IA à haut risque - à paraître		Systèmes d'IA à haut risque		
Chaque pays de l'UE – à venir		Systèmes d'IA à haut risque liés à des infrastructures critiques		
UE	Affaires en matière d'IA sélectionnées dans le secteur public	IA dans le secteur public	Joint Research Centre Data Catalogue – Selected AI cases in the public sector (JRC129301), Commission européenne (disponible en anglais)	31 décembre 2021

Portée géographique	Titre	Champ d'application	Lien	Date de mise à jour
UE	Public Sector Tech Watch	Observatoire dédié à la surveillance, à l'analyse et à la diffusion de l'utilisation des technologies émergentes	https://interoperable-europe.ec.europa.eu/collection/public-sector-tech-watch (disponible en anglais)	
Allemagne			Use Case Platform – Applied AI Institute (disponible en anglais)	
Helsinki, Finlande	Les systèmes d'intelligence artificielle de la ville d'Helsinki	IA dans le secteur public	AI Register | City of Helsinki AI Register (disponible en anglais)	Mises à jour constantes
Pays-Bas	Le registre des algorithmes du gouvernement néerlandais	Systèmes algorithmiques dans le secteur public	The Algorithm Register of the Dutch Government	Mises à jour constantes
Royaume-Uni	Plateforme d'enregistrement des normes de transparence algorithmique	Gouvernement britannique (norme volontaire)	Find out how algorithmic tools are used in public organisations – GOV.UK (disponible en anglais)	Mises à jour constantes
Royaume-Uni	Registre des organisations non gouvernementales : Projet de loi publique sur le suivi de registre automatisé du gouvernement (Tracking Automated Government ou TAG)		TAG Register (disponible en anglais)	Mises à jour constantes

Annexe G

Directives relatives à la non-discrimination, motifs protégés et champ d'application

Directive	Motifs protégés	Champ d'application	Standards directive
Directive 2000/43/CE (directive sur l'égalité I)	« Race » et origine ethnique – pas nationalité et sans préjudice du traitement découlant du statut juridique des personnes ressortissantes de pays non-membres de l'UE et des personnes apatrides	Emploi, formation professionnelle, protection sociale, éducation, biens et services, y compris logement	2024/1499
Directive 2000/78/CE (directive sur l'égalité II)	Religion or belief, disability, age, sexual orientation	Employment and occupation	2024/1499
Directive 2004/113/CE (directive sur l'égalité entre les hommes et les femmes dans le domaine des biens et services)	Sexe	Biens et services	2024/1499
Directive 2006/54/CE (refonte de la directive sur l'égalité entre les hommes et les femmes)	Sexe (y compris grossesse et maternité)	Emploi et profession (y compris les régimes de sécurité sociale)	2024/1500

Directive	Motifs protégés	Champ d'application	Standards directive
Directive 2010/41/UE (directive sur le travail indépendant sans distinction de genre)	Sexe	Travailleur·es indépendant·es, conjoint·es ou partenaires de vie de travailleur·es indépendant·es	2024/1500
Directive 2023/970/UE sur l'égalité des salaires (transposition en juin 2026)		Emploi	2024/1500 (sans préjudice des dispositions plus spécifiques de la directive elle-même)

Annexe H

Autres lignes directrices et règlements potentiellement pertinents

Conseil de l'Europe

Non-discrimination

Recommandation CM/Rec(2019)1 du Comité des Ministres aux États membres sur la prévention et la lutte contre le sexisme : https://search.coe.int/cm?i=09125948802339be.

Recommandation CM/Rec(2026)1 du Comité des Ministres aux États membres sur l'égalité et l'intelligence artificielle : https://search.coe.int/cm?i=09125948802acd80.

Recommandation CM/Rec(2020)1 du Comité des Ministres aux États membres sur les impacts des systèmes algorithmiques sur les droits de l'homme : https://search.coe.int/cm?i=09125948802662bf.

Études

Bartoletti, Ivana et Xenidis, Raphaële, 2023, Étude sur l'impact des systèmes d'intelligence artificielle, leur potentiel de promotion de l'égalité, y compris l'égalité de genre, et les risques qu'ils peuvent entraîner en matière de non-discrimination, GEC/CDADI. Disponible en suivant ce lien : https://rm.coe.int/prems-107623-fra-2530-etude-sur-l-impact-de-ai-a5-web/1680ac99e2.

Prévenir les discriminations résultant de l'utilisation de l'intelligence artificielle – Rapport de l'Assemblée parlementaire du Conseil de l'Europe, Commission sur l'égalité et la non-discrimination ; Rapporteur : M. Christophe Lacroix, (Belgique, SOC), 29 septembre 2020 : https://pace.coe.int/fr/files/28715/html.

Shrishak, Kris et Pénicaud, Soizic, 2025, La protection juridique contre la discrimination algorithmique en Europe : cadres actuels et lacunes subsistantes. Conseil de l'Europe. Disponible en suivant ce lien : https://rm.coe.int/legal-protection-fra-web/48802a38d9 (consulté le 1er juillet 2026)

Xenidis, Raphaële, 2025, Les 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, Conseil de l'Europe. Disponible en suivant ce lien : https://rm.coe.int/policy-guidelines-fra-web/48802a38d2 (consulté le 1er juillet 2026).

Zuiderveen Borgesius F, 2018, Discrimination, intelligence artificielle et décisions algorithmiques, ECRI, Disponible en suivant ce lien : https://rm.coe.int/etude-sur-discrimination-intelligence-artificielle-et-decisions-algori/1680925d84.

Genre

Recommandation CM/Rec(2026)2 du Comité des Ministres sur l'obligation de rendre des comptes en matière de violence à l'égard des femmes et des filles facilitée par la technologie : https://search.coe.int/cm?i=09125948802acd84.

Justice et application de la loi

Outil d'évaluation pour l'opérationnalisation de la Charte éthique européenne sur l'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement – CEPEJ(2023)16Final : https://rm.coe.int/cepej-2023-16final-operationalisation-de-la-charte-ethique-ia-fr/1680adcc9d.

Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement – CEPEJ(2018)14 : https://rm.coe.int/charte-ethique-fr-pour-publication-4-decembre-2018/16808f699b.

Lignes directrices sur la reconnaissance faciale, Comité consultatif de la Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel, Convention 108, 28 janvier 2021 : https://edoc.coe.int/fr/intelligence-artificielle/9749-lignes-directrices-sur-la-reconnaissance-faciale.html.

Sécurité sociale

Decl(17/03/2021)2 - Déclaration du Comité des Ministres sur les risques de la prise de décision assistée par ordinateur ou reposant sur l'intelligence artificielle dans le domaine du filet de sécurité sociale. Disponible en suivant ce lien : https://search.coe.int/cm?i=09125948802520f5.

Éducation

Havinga, Beth, Holmes, Wayne et Persson Jen (2024), Regulating the Use of Artificial Intelligence Systems in Education: Preparatory study on the development of a legal instrument. Disponible en anglais en suivant ce lien : https://www.coe.int/en/web/education/-/regulating-the-use-of-artificial-intelligence-systems-in-education

Recommandation sur l'impact de l'intelligence artificielle (IA) sur le secteur de l'éducation, adoptée par la Conférence des organisations internationales non gouvernementales – CONF-AG(2023)REC4. Disponible en anglais en suivant ce lien : https://rm.coe.int/conf-ag-2023-rec4-draft-recommendation-on-the-impact-of-artificial-int/1680ac8e25

Migration, immigration et asile

Recommandation CM/Rec(2022)17 du Comité des Ministres aux États membres sur la protection des droits des femmes et des filles migrantes, réfugiées et demandeuses d'asile (paragraphes 5-8 et 22-25 en particulier) : https://search.coe.int/cm?i=0900001680a69408.

Voir également : Dossier politique IT Flows, disponible en anglais : https://bura.brunel.ac.uk/handle/2438/27982.

Annexe I

Pratiques d'IA à haut risque

Domaine à haut risque	Cas d'utilisation
1. Biométrie	Systèmes d'identification biométrique à distance, à l'exclusion de la vérification
	Catégorisation biométrique de motifs sensibles ou protégés, ou conclusion de caractéristiques sur la base de ces motifs
	Reconnaissance des émotions
2. Infrastructure critique	Composants de sécurité dans la gestion et l'exploitation d'infrastructures critiques
3. Éducation et formation professionnelle	Déterminer l'accès ou l'admission, ou assigner des individus au niveau d'éducation
	Évaluer les résultats de l'apprentissage et/ou piloter le processus d'apprentissage dans le cadre de la formation éducative ou professionnelle
	Évaluer le niveau d'éducation approprié
	Surveiller ou détecter les comportements interdits des élèves pendant les tests
4. Emploi, gestion de la main d'œuvre et accès à l'emploi indépendant	Recrutement ou sélection de personnes physiques, offres d'emploi ciblées ou évaluation de candidat·es/candidatures
	Prise de décisions concernant les relations de travail (comme une promotion), la répartition des tâches et la surveillance ou l'évaluation du rendement
5. Accès aux services essentiels	Évaluation de l'admissibilité aux prestations et services publics
	Évaluation de la solvabilité ou établissement de la cote de crédit, hors fraude financière
	Évaluation des risques et tarification en assurance vie et maladie
	Évaluation et/ou classification des appels d'urgence pour l'envoi, ou établissement de la priorité dans l'envoi de la première intervention d'urgence ou le triage des soins de santé

Domaine à haut risque	Cas d'utilisation
6. Mise en application de la loi	Évaluation du risque d'être victime d'une infraction pénale
	Polygraphe ou outils similaires
	Évaluation de la fiabilité des preuves
	Évaluation du risque de (re)délinquance, ou évaluation des traits de personnalité ou du comportement criminel
	Aide à la détection, aux enquêtes ou aux poursuites pénales
7. Migration, asile et gestion du contrôle aux frontières	Polygraphe ou outils similaires
	Évaluation des risques en matière de sécurité, de migration irrégulière ou de santé posés en cas d'entrée sur le territoire d'un État membre
	Examen des demandes d'asile, de visa, de permis de séjour ou de plaintes relatives à l'éligibilité des demandes de statut
	Détection, reconnaissance ou identification des personnes physiques, à l'exception de la vérification (des documents de voyage)
8. Administration de la justice et des processus démocratiques	Aider les autorités judiciaires à effectuer des recherches et à interpréter les faits et le droit, et à appliquer le droit à des faits concrets
	Systèmes destinés à influencer le comportement électoral lors d'élections et/ou de référendums

Remarque : la Commission a publié un projet de lignes directrices sur la classification des systèmes d'IA à haut risque, en consultation jusqu'en juin 2026[64]

64. Pour en savoir plus, veuillez cliquer ici : La Commission sollicite un retour d'information sur le projet de lignes directrices pour la classification des systèmes d'intelligence artificielle à haut risque, consulté le 5 juin 2026.

Annexe J

Pratiques interdites en matière d'IA

Tableau 9. Liste des pratiques interdites concernant l'utilisation des systèmes d'IA

Pratique interdite	Articles du règlement sur l'IA	Considérant du règlement sur l'IA
Manipuler ou tromper intentionnellement, déformer de façon importante le comportement ou nuire à la capacité de prendre une décision éclairée, vraisemblablement susceptible de causer un préjudice important, y compris la production de contenus sexuels et intimes non consensuels ou de matériel pédopornographique.	5(1)(a)	Considérant 29
Exploiter toute vulnérabilité due à l'âge, à un handicap ou à la situation socio-économique, avec pour objectif ou pour effet de fausser sensiblement le comportement, raisonnablement susceptible de causer un préjudice important.	5(1)(a); 5(1)(ba) & (bb), 5(1a) and (1b)	Considérant 29
Classer ou évaluer une ou plusieurs personnes physiques sur une certaine période en fonction de leur comportement social ou de leurs caractéristiques personnelles, avec un score social conduisant à un traitement préjudiciable ou défavorable dans : ▸ un contexte non lié à celui initialement collecté ; ▸ de manière injustifiée ou disproportionnée.	5(1)(c)	Considérant 31
Évaluation du risque ou prévision de la commission d'une infraction pénale, fondée uniquement sur le profilage d'une personne physique.	5(1)(d)	Considérant 42
Capture non ciblée d'images faciales à partir de CCTV ou d'internet, pour développer ou créer une base de données de reconnaissance faciale.	5(1)(e)	Considérant 44
Reconnaissance des émotions en milieu de travail, sauf raisons médicales et de sécurité.	5(1)(f)	Recital 44

Pratique interdite	Articles du règlement sur l'IA	Considérant du règlement sur l'IA
Catégorisation biométrique de personnes individuelles pour déduire la « race », l'opinion politique, l'adhésion à un syndicat, les croyances religieuses ou philosophiques, la vie sexuelle ou l'orientation sexuelle.	5(1)(g)	Considérant 30
Identification biométrique à distance « en temps réel » dans des lieux accessibles au public à des fins de mise en application de la loi, des exceptions existent.	5(1)(h)	Considérants 32, 33, 38, 39, 40 et 41

Tableau 10. Les pratiques d'IA interdites peuvent être liées à l'IA à haut risque

Systèmes d'IA à haut risque (article 6(2), annexe III du règlement sur l'IA	Pratique interdite (article 5 du règlement sur l'IA)	Exemples de cas d'utilisation et d'explication interdits et potentiellement interdits
1) Biométrie a) Systèmes d'identification biométrique à distance (Remote biometric identification ou RBI)	(h) Identification biométrique à distance « en temps réel » dans les espaces publics par les forces de l'ordre, sauf en cas d'utilisation pour : i. la recherche d'une victime d'enlèvement ; ii. la prévention d'une menace spécifique, substantielle et imminente ; iii. l'identification d'une personne pour des infractions spécifiques (voir annexe II).	**Éléments clés de l'utilisation de la RBI :** **i) « en temps réel »** **ii) dans les espaces publics** **iii) par les forces de l'ordre ou en leur nom** Le considérant 32 précise que l'utilisation de systèmes RBI 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 » **Exemple : la police utilise la vidéosurveillance par reconnaissance faciale en direct pour identifier et arrêter une personne d'intérêt.** Après avoir identifié une personne manifestant de manière pacifique, la police de Moscou (Glukhin c. Russie, 2024) a utilisé des images de vidéosurveillance pour identifier et localiser un individu à l'aide d'une identification biométrique en temps réel.

Systèmes d'IA à haut risque (article 6(2), annexe III du règlement sur l'IA	Pratique interdite (article 5 du règlement sur l'IA)	Exemples de cas d'utilisation et d'explication interdits et potentiellement interdits
1a) systèmes d'identification biométrique à distance (RBI)	(e) capture non ciblée d'images faciales pour créer ou développer une base de données	**Éléments clés : méthode de création ou d'expansion d'une base de données de reconnaissance par capture non ciblée plutôt que par l'identification.** C'est interdit car la capture non ciblée fonctionne comme un « aspirateur », comme le prévoit l'article 5(1)(e), en collectant sans distinction des images faciales sans cibler des individus spécifiques. Le considérant 43 ajoute à la justification fondée sur le « sentiment de surveillance de masse » et la possibilité de violation des droits fondamentaux. **Exemple : les autorités de contrôle aux frontières établissent une base de données de reconnaissance faciale par capture de masse** Les autorités de contrôle aux frontières pourraient capturer des images faciales à partir des réseaux sociaux, d'articles de presse et d'images de vidéosurveillance sans cibler spécifiquement les individus. Si elles sont ensuite ajoutées à une base de données pour l'identification biométrique, ce serait interdit en vertu de l'article 5(1)(e).

Systèmes d'IA à haut risque (article 6(2), annexe III du règlement sur l'IA	Pratique interdite (article 5 du règlement sur l'IA)	Exemples de cas d'utilisation et d'explication interdits et potentiellement interdits
1b) Catégorisation biométrique des caractéristiques protégées (ou conclusion de celle-ci)	(g) Catégorisation biométrique qui déduit la « race », les opinions politiques, l'adhésion à un syndicat, les croyances religieuses ou philosophiques, la vie sexuelle ou l'orientation sexuelle	**Éléments clés : l'objectif spécifique de déduire les caractéristiques protégées énumérées (« race », opinions politiques, appartenance syndicale, religion, vie sexuelle ou orientation sexuelle) sur la base de données biométriques.** La catégorisation biométrique est déjà immédiatement considérée comme présentant un risque élevé, l'article 5(1)(g) interdisant absolument l'utilisation de données biométriques pour déduire ces attributs spécifiques. **Exemple : un système d'IA qui analyse les images des réseaux sociaux d'une personne pour déduire son orientation sexuelle à partir de ses données biométriques afin de lui envoyer de la publicité ciblée.**
1c) Reconnaissance des émotions	(f) Reconnaissance des émotions sur le lieu de travail et dans les établissements d'enseignement (exception : raisons médicales et de sécurité)	**Éléments clés : déduire des émotions ou des intentions dans le contexte d'un milieu de travail et de l'éducation où les exemptions de justification médicale ou de sécurité ne s'appliquent pas.** L'article 5(1)(f) interdit l'utilisation de la pratique de reconnaissance des émotions à haut risque dans ces contextes spécifiques. En outre, le considérant 44 met en lumière les préoccupations quant au fondement scientifique de ces systèmes d'IA qui, combinés en particulier aux déséquilibres de pouvoir dans un contexte professionnel ou éducatif, pourraient avoir des effets préjudiciables.

Systèmes d'IA à haut risque (article 6(2), annexe III du règlement sur l'IA	Pratique interdite (article 5 du règlement sur l'IA)	Exemples de cas d'utilisation et d'explication interdits et potentiellement interdits
		Remarque : la reconnaissance des états physiques, par exemple si une personne conduisant un bus est fatigué ou non, n'est pas interdite ici. **Exemple : un système d'IA qui déduit si les employé·es quittent leur lieu de travail « heureux », « en colère » ou « tristes ».** Un·e employeur·e essayant de mesurer les émotions des employés de son restaurant à partir d'images de vidéosurveillance.
3) Éducation et formation professionnelle a) Déterminer l'accès/l'admission	(a) Manipulation, tromperie ou altération d'une décision éclairée, causant un préjudice	**Éléments clés : utilisation d'une technique de manipulation ou trompeuse qui déforme sensiblement (ou est raisonnablement susceptible de déformer) le comportement d'une personne.** L'article 5(1)(a) interdit les techniques qui ont pour objet ou pour effet de fausser le comportement par des techniques qui nuisent à la prise de décisions éclairées et causent un préjudice significatif. Remarque : l'intention n'est pas nécessaire, l'effet suffit. **Exemple : une université employant un chatbot dans la détermination des admissions, avec des comptes à rebours ou des messages fabriqués.** Un chatbot déployant des techniques autrement connues sous le nom de motifs sombres, comme « il ne reste que 3 places disponibles ! » serait interdite en vertu de cet article.

Systèmes d'IA à haut risque (article 6(2), annexe III du règlement sur l'IA	Pratique interdite (article 5 du règlement sur l'IA)	Exemples de cas d'utilisation et d'explication interdits et potentiellement interdits
3) Éducation et formation professionnelle b) Évaluation des résultats d'apprentissage et orientation du processus d'apprentissage c) Évaluation du niveau d'éducation (approprié) d) Surveillance/ détection des comportements interdits des élèves	f) Reconnaissance des émotions sur le lieu de travail et dans les établissements d'enseignement (exception : raisons médicales et de sécurité)	**Élément clé : la déduction des états émotionnels ou des intentions des élèves, et l'influence de ceux-ci sur leur processus d'apprentissage.** Les systèmes d'IA peuvent être utilisés pour évaluer les élèves ; cependant, il est interdit de reconnaître leurs émotions ou leurs intentions. En outre, le considérant 44 met en lumière les préoccupations quant au fondement scientifique de ces systèmes d'IA qui, combinés en particulier aux déséquilibres de pouvoir dans un contexte professionnel ou éducatif, pourraient avoir des effets préjudiciables. **Exemple : une université utilise une plateforme d'apprentissage personnalisée qui prend en compte les émotions des étudiant·es sur leurs webcams lors de conférences en ligne pour orienter le processus d'apprentissage.**
4) Gestion de l'emploi et des travailleurs et travailleuses a) Recrutement (publicité ciblée) ou sélection (évaluation des candidatures)	(a) Manipulation, tromperie ou altération d'une décision éclairée, causant un préjudice (b) Exploitation d'une vulnérabilité	**Éléments clés : utilisation d'une technique de manipulation ou trompeuse qui déforme sensiblement (ou est raisonnablement susceptible de déformer) le comportement d'une personne** L'article 5(1)(a) interdit les techniques qui ont pour objet ou pour effet de fausser le comportement par des techniques qui nuisent à la prise de décisions éclairées et causent un préjudice significatif. Remarque : l'intention n'est pas nécessaire, l'effet suffit. **Exemple : une plateforme de recrutement par IA avec une urgence fabriquée et une manipulation trompeuse pour signer un contrat.**

Systèmes d'IA à haut risque (article 6(2), annexe III du règlement sur l'IA	Pratique interdite (article 5 du règlement sur l'IA)	Exemples de cas d'utilisation et d'explication interdits et potentiellement interdits
		Une agence de recrutement déployant un système d'IA avec des salaires/avantages exagérés ou une urgence entravant une décision éclairée serait interdite en vertu de cet article.
4) Gestion de l'emploi et des travailleurs et travailleuses b) Décisions affectant les relations de travail, telles que la cessation d'emploi ou l'attribution de tâches	(a) Manipulation, tromperie ou altération d'une décision éclairée, causant un préjudice (b) Exploitation d'une vulnérabilité	**Éléments clés : utilisation d'une technique de manipulation ou trompeuse qui déforme sensiblement (ou est raisonnablement susceptible de déformer) le comportement d'une personne.** L'article 5(1)(a) interdit les techniques qui ont pour objet ou pour effet de fausser le comportement par des techniques qui nuisent à la prise de décisions éclairées et causent un préjudice significatif. Remarque : l'intention n'est pas nécessaire, l'effet suffit. **Exemple : une entreprise de logistique qui déploie l'IA pour allouer les tâches de livraison et surveiller les performances des chauffeur·ses.** Il existe un scénario réalisable où, grâce à un tel système, le comportement des conducteurs et conductrices pourrait être faussé en travaillant plus longtemps que les heures légales, ce qui aurait des répercussions sur la santé et la sécurité et causerait des dommages importants. Un tel système serait donc interdit.

Systèmes d'IA à haut risque (article 6(2), annexe III du règlement sur l'IA	Pratique interdite (article 5 du règlement sur l'IA)	Exemples de cas d'utilisation et d'explication interdits et potentiellement interdits
4) Gestion de l'emploi et des travailleur·ses b) Décisions affectant les relations de travail, comme la cessation d'emploi ou l'attribution de tâches	(f) Reconnaissance des émotions et/ou des intentions sur le lieu de travail et dans les établissements d'enseignement	**Élément clé : la déduction des états émotionnels ou des intentions des travailleurs·ses, et l'influence de ceux-ci sur leur travail** Des systèmes d'IA peuvent être utilisés pour aider à la prise de décision. Toutefois, il est interdit de fonder toute assistance sur l'utilisation de la reconnaissance des émotions. En outre, le considérant 44 met en lumière les préoccupations quant au fondement scientifique de ces systèmes d'IA qui, combinés en particulier aux déséquilibres de pouvoir dans un contexte professionnel ou éducatif, pourraient avoir des effets préjudiciables. **Exemple : un centre d'appels qui déploie l'IA pour surveiller les travailleur·ses utilise leur état émotionnel pour évaluer leurs performances.** Dans l'exemple, si un système d'IA devait déduire les états émotionnels d'un·e salarié·e en fonction d'enregistrements vocaux et en tenir compte pour l'évaluation des performances, il s'agirait d'une pratique interdite.

Systèmes d'IA à haut risque (article 6(2), annexe III du règlement sur l'IA	Pratique interdite (article 5 du règlement sur l'IA)	Exemples de cas d'utilisation et d'explication interdits et potentiellement interdits
5) Accès à des services et prestations essentiels a) Évaluation de l'éligibilité d'une personne physique	c) Notation sociale conduisant à un traitement défavorable : i) dans un contexte différent des données collectées et/ou ii) injustifié ou disproportionné	**Éléments clés : traitement défavorable fondé sur leur comportement social dans un contexte non lié et/ou disproportionné dans son effet.** Si un système d'IA est utilisé pour évaluer l'admissibilité, il se peut qu'il ne s'agisse pas de données provenant d'un contexte différent. **Exemple : les autorités publiques déploient un système d'IA qui utilise des données non liées pour allouer des avantages.** Il est important, dans ces cas interdits, que le contexte dans lequel les données sont collectées diffère sensiblement du contexte dans lequel elles sont utilisées.
5b) Évaluation de la solvabilité	c) Notation sociale conduisant à un traitement défavorable : i) dans un contexte différent des données collectées et/ou ii) injustifié ou disproportionné	**Éléments clés : traitement défavorable fondé sur leur comportement social dans un contexte non lié et/ou disproportionné dans son effet.** Si un système d'IA est utilisé pour évaluer le crédit d'une personne, il peut le faire en fonction de l'historique financier, mais pas d'autres contextes non reliés. **Exemple : les institutions financières déploient un système d'IA pour évaluer les prêts, mais aussi pour accéder à d'autres bases de données non financières.** Si une banque devait également analyser les connexions aux réseaux sociaux ou tenir compte de l'adresse de la personne pour décider si le prêt doit être approuvé ou non, cela pourrait être interdit en vertu de l'article 5(1)(c).

Systèmes d'IA à haut risque (article 6(2), annexe III du règlement sur l'IA	Pratique interdite (article 5 du règlement sur l'IA)	Exemples de cas d'utilisation et d'explication interdits et potentiellement interdits
5c) Évaluation des risques en ce qui concerne la tarification de l'assurance maladie ou de l'assurance vie	c) Notation sociale conduisant à un traitement défavorable : i) dans un contexte différent des données collectées et/ou ii) injustifié ou disproportionné	**Éléments clés : traitement défavorable fondé sur leur comportement social dans un contexte non lié et/ou disproportionné dans son effet** Si un système d'IA est utilisé pour établir le prix de l'assurance-vie, il peut le faire en fonction des antécédents médicaux, mais pas d'autres contextes non liés. **Exemple : une assurance déploie un système d'IA pour évaluer les prix en fonction d'informations provenant d'un contexte non lié.** Si une assurance devait analyser les connexions et le comportement sur les réseaux sociaux, cela pourrait être interdit en vertu de l'article 5(1)(c).
6) Mise en application de la loi a) Évaluation du risque de victimisation b) Polygraphes c) Évaluation de la fiabilité des preuves d) Évaluation du risque de (re) délinquance, ou évaluation des caractéristiques e) Profilage lors de la détection, de l'enquête ou des poursuites	d) Évaluation du risque de criminalité uniquement fondée sur le profilage	**Éléments clés : l'évaluation est fondée sur les traits et les caractéristiques de la personnalité et aucun fait objectif.** **Remarque : si un système appuie une décision humaine, une exception peut exister.** L'article 5(1)(d) interdit l'évaluation des risques pour les infractions pénales fondée uniquement sur le profilage des traits de personnalité sans fait objectif vérifiables liés à une activité criminelle. **Exemple : système de prévision policière, profilant les individus en fonction de leurs données démographiques.** Si le service de police devait prédire la probabilité d'une activité criminelle sur le fondement de caractéristiques telles que l'âge, la nationalité ou le voisinage, il s'agirait d'un système interdit en vertu de l'article 5(1)(d).

Systèmes d'IA à haut risque (article 6(2), annexe III du règlement sur l'IA	Pratique interdite (article 5 du règlement sur l'IA)	Exemples de cas d'utilisation et d'explication interdits et potentiellement interdits
6) Mise en application de la loi e) Profilage lors de la détection, de l'enquête ou de poursuites	(h) Identification biométrique à distance en temps réel dans les espaces publics destinés à la mise en application de la loi, sauf si elle est utilisée pour : i. la recherche d'une victime d'enlèvement ii. la prévention d'une menace spécifique, substantielle et imminente ; iii. l'identification d'une personne pour des infractions spécifiques (voir annexe II).	**Éléments clés -l'utilisation de la RBI :** **i) « en temps réel »** **ii) dans les espaces publics** **iii) par les forces de l'ordre ou en leur nom** Le considérant 32 précise que l'utilisation de systèmes RBI 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 ». **Exemple : la police utilise la vidéosurveillance par reconnaissance faciale en direct pour identifier et arrêter une personne d'intérêt.** Après avoir identifié une personne manifestant de manière pacifique, la police de Moscou (Glukhin c. Russie, 2024) a utilisé des images de vidéosurveillance pour identifier et localiser un individu à l'aide d'une identification biométrique en temps réel.

Systèmes d'IA à haut risque (article 6(2), annexe III du règlement sur l'IA	Pratique interdite (article 5 du règlement sur l'IA)	Exemples de cas d'utilisation et d'explication interdits et potentiellement interdits
7) Migration, asile et contrôles aux frontières b) Évaluation des risques liés à la sécurité, à l'irrégularité ou à la santé c) Évaluation de l'asile/de visa/ de permis, de l'éligibilité et de plaintes	d) Évaluation du risque de criminalité uniquement fondée sur le profilage	**Éléments clés : l'évaluation est fondée sur les traits de personnalité et aucun fait objectif.** **Remarque : si un système appuie une décision humaine, des exceptions peuvent exister.** L'article 5(1)(d) interdit l'évaluation des risques pour les infractions pénales fondée uniquement sur le profilage des traits de personnalité sans fait objectif vérifiable lié à une activité criminelle. **Exemple : les autorités de contrôle aux frontières déploient un système d'IA qui fonde le profilage des risques présentés par les individus sur leurs données démographiques.** Si les autorités de contrôle aux frontières devaient prédire le risque de délinquance sur le fondement de caractéristiques telles que la nationalité ou la région, il pourrait s'agir d'un système interdit en vertu de l'article 5(1)(d).
7) Migration, asile et contrôles aux frontières (b) Évaluation des risques en matière de sécurité, d'irrégularité ou de santé	(h) Identification biométrique à distance en temps réel dans les espaces publics destinés à la mise en application de la loi, sauf si elle est utilisée pour : i. la recherche d'une victime d'enlèvement ; ii. la prévention d'une menace spécifique, substantielle et imminente ; iii. l'identification d'une personne pour des infractions spécifiques (voir annexe II)..	**Éléments clés - l'utilisation de la RBI :** **i) « en temps réel »** **ii) dans les espaces publics** **iii) par les forces de l'ordre ou en leur nom** Le considérant 32 précise que l'utilisation de systèmes RBI 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 ». **Exemple : l'utilisation de la vidéosurveillance par reconnaissance faciale en direct à la frontière pour identifier les personnes d'intérêt.**

Systèmes d'IA à haut risque (article 6(2), annexe III du règlement sur l'IA	Pratique interdite (article 5 du règlement sur l'IA)	Exemples de cas d'utilisation et d'explication interdits et potentiellement interdits
7) Migration, asile et contrôle aux frontières (d) Détection, reconnaissance ou identification des personnes physiques (exception pour la vérification des documents de voyage)	(e) capture non ciblée d'images faciales	**Éléments clés : méthode de création ou d'expansion d'une base de données de reconnaissance par capture non ciblée plutôt que par l'identification.** C'est interdit car la capture non ciblée fonctionne comme un « aspirateur », comme le prévoit l'article 5(1)(e), en collectant sans distinction des images faciales sans cibler des individus spécifiques. Le considérant 43 ajoute à la justification fondée sur le « sentiment de surveillance de masse » et la possibilité de violation des droits fondamentaux. **Exemple : les autorités de contrôle aux frontières établissent une base de données de reconnaissance faciale par capture de masse.** Les autorités de contrôle aux frontières pourraient capturer des images faciales à partir des réseaux sociaux, d'articles de presse et d'images de vidéosurveillance sans cibler spécifiquement les individus. Si elles devaient ensuite être ajoutées à une base de données pour l'identification biométrique à la frontière, ce serait interdit en vertu de l'article 5(1)(e).

Annexe K

Feuille d'enregistrement pour l'étape 3

Cette feuille a été conçue comme un exemple de la façon dont une personne en charge d'un dossier suivant les étapes de la section 3 pourrait consigner ses constats. Les exemples sont fictifs pour donner une idée des types d'informations qui pourraient être enregistrées. L'enregistrement réel devrait être beaucoup plus détaillé. La dernière colonne vise à permettre un examen de la preuve à la suite de toute divulgation par l'organisation intimée.

Étape		Presumption of discrimination	Following respondent's disclosure
3.1. Risques connus et présumés **3.1.1 Type de système**	Indiquez si le système est un système à haut risque ou s'il est destiné à être utilisé dans l'un des secteurs ou domaines identifiés comme présentant un risque de discrimination dans le règlement sur l'IA.	**Exemples** **Secteur :** **Cas d'utilisation :**	
3.1.2 Risques connus de discrimination	Consignez le domaine dans lequel il existe un risque connu, le type de technologie et toute autre information et preuve pertinentes.	**Exemples** **Secteur :** application de la loi **Technologie :** reconnaissance faciale en direct, biométrie **Source:** OSC ou recherche universitaire, règlement sur l'IA, article 5 (systèmes interdits), article 6 (haut risque) et annexe III.	
3.2 Caractéristiques protégées et proxy	Enregistrez l'utilisation directe des caractéristiques protégées. Où sont-elles utilisées et comment ?	**Caractéristiques protégées** **Mode d'utilisation :** données d'entrée, par exemple	

Étape		Presumption of discrimination	Following respondent's disclosure
Approfondir l'analyse de la conception technique du système et l'influence des proxys (suspectés) sur les résultats du système. **Le fait qu'un proxy soit utilisé en tant qu'entrée ne suffit pas en soi comme preuve de discrimination. La prochaine étape consiste à évaluer les preuves pour comprendre si l'utilisation de ces entrées mène à un traitement inégal.**	Consignez toute utilisation identifiée de caractéristiques protégées ou de proxys qui sont utilisés ou dont on croit qu'ils ont une incidence sur les résultats du système. Notez également pour chaque proxy pourquoi l'OPE considère qu'il s'agit de proxy.	**Exemples** **Proxy:** par exemple, un code postal en tant que proxy pour la « race », la grossesse en tant que proxy pour le sexe. **Preuves et arguments**	
3.3 : Analyses des risques, y compris AIDF	Consignez tout risque de discrimination identifié par le fournisseur ou le déployeur et toute mesure identifiée pour atténuer ce risque.	**Exemples** **Risque :** discrimination raciale due à des biais historiques **En cas d'identification :** analyse des risques **Mesure d'atténuation :** technique de débiaisement	

Étape		Presumption of discrimination	Following respondent's disclosure
	Examinez et notez s'il y a des défaillances claires et évidentes dans l'évaluation des risques, l'AIPD ou l'AIDF.	L'analyse des risques fait référence à la discrimination mais pas à la « race ». AIPD : ne tient pas compte de la « race » dans les données de catégorie spéciale. AIDF : ne tient pas compte de la discrimination raciale ni de l'article 14 de la Convention européenne des droits de l'homme.	
	Prenez en compte et enregistrez l'existence et l'impact de tout être humain dans la boucle. Qui est la personne concernée ? La personne concernée a-t-elle la compétence, la formation et l'autorité nécessaires pour outrepasser une décision prise par un système d'IA ? La personne concernée a-t-elle été formée pour comprendre et interpréter le résultat du système et reconnaître les cas où des biais peuvent survenir et où une intervention est nécessaire ? Existe-t-il des lignes directrices ou des protocoles destinés aux prises de décision pour aider à comprendre comment utiliser le système d'IA ou interpréter les résultats ?	Personne de contact : Rôle : examine les décisions, mais n'examine pas les données qui ont mené à la décision. Demandez si c'est suffisant ?	

Étape		Presumption of discrimination	Following respondent's disclosure
3.4 : Objectif du système	Consignez l'objectif du système et ce qu'il optimise. Enregistrez les proxys et les caractéristiques protégées qui influencent le résultat du système et, le cas échéant, les mesures de leur poids/influence sur le résultat	Objectif du système : Conçu pour optimiser : Caractéristiques protégées qui influencent le système : Proxys qui influencent le système : Indicateurs de poids/influence :	
3.5 Entraînement, tests et données de validation	Consignez tout problème général lié aux données, y compris les problèmes liés à la qualité des données.	Par exemple, l'échantillon de données est très petit, ancien ou se rapporte à une zone géographique différente.	
	Questions relatives aux données concernant les inégalités et la discrimination.		
	Problèmes liés à l'insuffisance de la documentation.		
	Préoccupations concernant les mesures d'atténuation.		
	Argumentation quant à la manière dont les problèmes de données et d'atténuation conduiraient à un traitement inégal dans le cas présent.		
	Consignez tout indicateur de non-conformité aux dispositions du règlement sur l'IA relatives aux données.		

Étape		Presumption of discrimination	Following respondent's disclosure
3.6 Tests et évaluation	Les écarts de résultat ou de traitement révélés par les statistiques.		
	Problèmes avec les statistiques rapportées (par exemple, une métrique incorrecte ou des problèmes de données).		
	Lacunes identifiées dans les tests.		
Conclusion	Examinez les renseignements et consignez le fondement sur lequel une preuve de discrimination prima facie peut être invoquée ou conclure qu'elle ne peut pas l'être.		
	Examinez toutes les étapes après la divulgation afin de déterminer si l'organisation intimée a réussi à réfuter la présomption ou à justifier la discrimination.		

L'intelligence artificielle (IA) transforme nos sociétés et a des répercussions sur les droits fondamentaux et l'égalité. Ces systèmes peuvent renforcer et engendrer des discriminations, car ils sont souvent conçus pour établir des distinctions entre les personnes : en les classant, en les hiérarchisant ou en les ciblant. Combinés aux inégalités historiques et aux préjugés humains ancrés dans les données, les règles et les choix de conception des systèmes algorithmiques, ces schémas discriminatoires risquent d'être reproduits et amplifiés s'ils ne sont pas identifiés et atténués de manière proactive.

Cette méthodologie fournit des orientations étape par étape aux organismes de promotion de l'égalité et autres organisations de défense des droits humains en Europe afin d'évaluer les cas de discrimination liés à l'IA ou aux systèmes algorithmiques. Elle s'adresse aux chargé·es de dossiers, conseiller·ères et juristes des entités chargées d'enquêter sur les plaintes pour discrimination impliquant, ou soupçonnées d'impliquer, l'IA ou la prise de décision automatisée.

Cette méthodologie s'appuie sur le cadre réglementaire européen en matière d'IA, notamment sur le règlement sur IA de l'UE et la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle, les droits de l'homme, la démocratie et l'État de droit, ainsi que sur la législation anti-discrimination, et fait référence à d'autres textes législatifs, tels que ceux relatifs à la protection des données, à des exemples de cas concrets et à la jurisprudence, le cas échéant.

UNIA Together for equality

YHDENVERTAISUUSVALTUUTETTU
NON-DISCRIMINATION OMBUDSMAN

CIG COMISSÃO PARA A CIDADANIA E A IGUALDADE DE GÉNERO

PREMS 114926

FR

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.

http://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 oeuvre de la Convention dans les Etats membres.

www.coe.int

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