

Conseil de l'Europe - Instrument d'appui technique de l'Union européenne

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»¹

Résumé du rapport d'analyse de la situation actuelle et des lacunes²

Auteur.e.s du rapport : Raphaële Xenidis (sections européennes), Ine van Zeeland (section sur la Belgique), Emeline Banzuzi (section sur la Finlande) et Eduardo Leonardo dos Santos (section sur le Portugal)

Les administrations publiques européennes déploient ou expérimentent de plus en plus de systèmes d'intelligence artificielle (IA) et/ou de prise de décision automatisée (PDA) dans un large éventail de domaines tels que la police, l'application de la loi, la migration, la protection sociale, la sécurité sociale, les soins de santé, l'éducation, la fiscalité, l'emploi, etc. Si les stades de développement et les projets varient d'un pays à l'autre, la discrimination algorithmique est une préoccupation transversale importante, car les systèmes d'IA et de PDA peuvent perpétuer et amplifier des préjugés injustes. Dans les administrations publiques, le développement et le déploiement de systèmes d'IA et de PDA répondent souvent à des préoccupations telles que la rationalisation de la prise de décision et l'efficacité des dépenses publiques. Ces logiques mettent en avant certaines fonctions et utilisations de l'IA et de la PDA (par exemple, la détection des fraudes, l'optimisation de l'allocation des ressources, etc.) qui présentent des risques particuliers de discrimination. **Le respect des droits fondamentaux, y compris la non-discrimination, doit donc être une condition *sine qua non* pour la fourniture et le déploiement de systèmes d'IA et de PDA.**

En Europe, le cadre de gouvernance a évolué avec l'adoption de la **Convention-cadre du Conseil de l'Europe sur l'IA** et du **Règlement de l'UE sur l'IA** en 2024, qui visent *entre autres* à protéger les droits fondamentaux et à prévenir la discrimination dans les applications d'IA. En outre, la législation **en matière de lutte contre la discrimination et de protection des données** fournit d'importantes garanties supplémentaires. Ensemble, **ces instruments juridiques offrent des remèdes essentiels contre la discrimination algorithmique, mais plusieurs défis subsistent.**

En **Belgique**, les autorités publiques sont peu conscientes de l'impact des systèmes d'IA et de PDA sur la non-discrimination, l'égalité de traitement et les droits fondamentaux. Lorsqu'elles abordent ces questions, les autorités ont tendance à se concentrer sur les lignes directrices éthiques plutôt que sur le respect des droits fondamentaux. Se concentrer uniquement sur

¹ Le projet est cofinancé par l'Union européenne via l'instrument d'appui technique et mis en œuvre par le Conseil de l'Europe, en coopération avec la Commission européenne.

² Ce document a été réalisé avec le soutien financier de l'Union européenne et du Conseil de l'Europe. Son contenu relève de la seule responsabilité de ses auteurs. Les opinions qui y sont exprimées ne reflètent en aucun cas l'opinion officielle de l'Union européenne ou du Conseil de l'Europe.

l'éthique est insuffisant, car l'éthique n'est pas pleinement opérationnelle et manque de force exécutoire, contrairement au droit de ne pas être discriminé.

Ce manque de sensibilisation est aggravé par un manque important de transparence concernant l'utilisation des systèmes d'IA et de PDA dans la sphère publique. Bien que les autorités fédérales s'efforcent de remédier à cette situation en mettant en place un observatoire, la transparence est encore loin d'être suffisante. Par conséquent, la sensibilisation du public à l'utilisation potentielle de ces systèmes reste faible, ce qui entrave considérablement l'accès à des recours efficaces en cas de discrimination. Un aperçu structuré et accessible des systèmes d'IA ou de PDA utilisés par toutes les autorités publiques, ainsi que des politiques normalisées pour prévenir, atténuer et corriger les impacts sur les droits fondamentaux, sont nécessaires de manière urgente. La sensibilisation des différentes parties prenantes devrait être renforcée par de multiples approches.

Conformément au règlement sur l'IA, une surveillance efficace et une protection plus large des droits fondamentaux nécessiteront une coopération significative entre les autorités de surveillance du marché et les organismes de promotion de l'égalité. Compte tenu de la complexité institutionnelle en Belgique, il est urgent de conclure des accords formels pour établir cette coopération, avant la date limite d'août 2025 pour la mise en place de la structure de surveillance. Des ressources adéquates doivent être allouées aux organismes de promotion de l'égalité afin de garantir leur participation effective au cadre de surveillance, conformément aux récentes directives de l'UE sur les normes relatives aux organismes de promotion de l'égalité.

En **Finlande**, il n'existe pas de registre national commun pour les systèmes de PDA ou d'IA, ce qui entraîne un manque de connaissance globale de ces systèmes en cours d'utilisation. L'étude a révélé que dans le secteur public, l'utilisation de la PDA est perçue comme relativement élevée, tandis que l'utilisation de l'IA en est encore au stade pilote. La sensibilisation des autorités publiques aux risques de discrimination liés à l'IA reste sans doute faible. La collaboration entre les autorités publiques et les autres parties prenantes devrait permettre d'accroître la sensibilisation aux systèmes d'IA/PDA utilisés ainsi qu'aux risques de discrimination qui y sont associés, d'augmenter les ressources allouées/supplémentaires, de proposer des formations supplémentaires et d'établir une approche normalisée pour évaluer les risques que ces systèmes font peser sur les droits fondamentaux.

Au **Portugal**, l'adoption de l'IA et de la PDA par l'administration publique est relativement récente, avec des projets et des initiatives majeurs annoncés récemment, dans une tendance rendue possible par le financement provenant du plan de relance et de résilience portugais. En conséquence, le potentiel de cas de discrimination fondée sur l'IA devrait augmenter. Le Portugal pourrait profiter de l'adoption relativement récente de ces technologies dans le secteur public pour établir un cadre solide qui garantisse une compréhension globale de leur utilisation dans le secteur public tout en promouvant les meilleures pratiques.

L'adoption de systèmes d'IA dans l'administration publique est généralement considérée avec enthousiasme, du point de vue de l'innovation, étant donné qu'ils peuvent permettre d'améliorer l'efficacité et de moderniser les services publics. Toutefois, cette perspective technologique ne doit pas faire oublier les risques de discrimination et d'inégalité de genre que ces systèmes peuvent introduire. À cet égard, des inquiétudes persistent quant au rôle des organismes de défense des droits fondamentaux exclus de la liste des entités visées à l'article

77 du règlement de l'UE sur l'IA, y compris les organismes de promotion de l'égalité, dont les compétences et les ressources devraient également être réévaluées conformément aux nouvelles directives sur les organismes de promotion de l'égalité afin de s'assurer qu'ils sont suffisamment équipés pour relever les nouveaux défis.

NB : Les conclusions concernant la Belgique, la Finlande et le Portugal sont basées sur une étude documentaire et un nombre limité d'entretiens avec le personnel des autorités publiques. Elles peuvent donc ne pas refléter pleinement toutes les initiatives en cours ou les différents niveaux de sensibilisation des différentes institutions.

Les questions critiques dans l'adoption des systèmes d'IA et de PDA et les principales lacunes dans le paysage juridique existant comprennent :

- Les **discours sur la neutralité et la rentabilité** qui encouragent l'adoption de l'IA **occultent des questions importantes** concernant les **fonctions** que l'IA et les systèmes de PDA devraient remplir (par exemple, la détection des fraudes par rapport à la facilitation de l'accès à des droits et à des services publics systématiquement sous-réclamés) et le **coût réel du déploiement**, une fois que les coûts liés au respect des obligations en matière de droits fondamentaux ont été correctement internalisés.
- Un **manque généralisé de transparence et d'informations significatives sur l'utilisation de ces systèmes par les autorités publiques**, ce qui rend difficile le suivi de leur impact sur l'égalité et la discrimination. En l'absence d'obligations de déclaration, il est difficile pour les organismes de contrôle, les chercheurs et les personnes concernées d'évaluer si les décisions fondées sur l'IA ou (semi) automatisées sont conformes aux lois anti-discrimination. Bien que le règlement sur l'IA et la convention-cadre sur l'IA remédieront à ce manque de transparence, des difficultés persisteront. Par exemple, les obligations d'enregistrement sont limitées aux systèmes à haut risque et des exceptions existent.
- L'**accès à la justice** est compromis par les asymétries d'information et de pouvoir entre les déployeurs de systèmes d'IA et de PDA et les individus et groupes qui en sont affectés. Les victimes de discrimination algorithmique ont souvent du mal à identifier et à obtenir des preuves de pratiques discriminatoires, ce qui rend difficile la contestation des décisions fondées sur l'IA ou (semi-)automatisées. Si les règles de procédure existantes sur le renversement de la charge de la preuve en cas de présomption de discrimination facilitent la tâche des victimes, le seuil de preuve pourrait encore être trop élevé pour les victimes dans un contexte d'opacité algorithmique. En d'autres termes, les individus pourraient ne pas être en mesure de se prévaloir des recours juridiques existants.
- L'**application des dispositions existantes et futures** pourrait être compromise par les **lacunes juridiques et institutionnelles** existantes. Par exemple, la législation européenne en matière de lutte contre la discrimination ne reconnaît pas explicitement les **formes de discrimination** qui pourraient être particulièrement répandues dans la prise de décision automatisée, telles que la discrimination intersectionnelle et systémique. Les restrictions juridiques existantes en matière de **collecte de données sur l'égalité** pourraient également limiter les possibilités d'enquêter efficacement sur la discrimination algorithmique et de promouvoir l'égalité dans l'utilisation des systèmes

d'IA et de gestion des données d'entreprise. Si l'application des réglementations en matière d'IA repose sur la coopération institutionnelle entre les organismes de promotion de l'égalité et d'autres organismes compétents tels que les autorités de surveillance du marché, **les modèles de coopération, les canaux de participation et les mécanismes de représentation devront être améliorés et renforcés.**

- **Stratégies de réduction des risques** : les incertitudes juridiques concernant l'interprétation des définitions et du champ d'application de l'IA et des systèmes à haut risque pourraient être utilisées par les fournisseurs du secteur pour éviter d'avoir à se conformer aux exigences du règlement de l'UE sur l'IA lors de la commercialisation de leurs systèmes. Ces pratiques devraient faire l'objet d'un contrôle public efficace.
- **Normalisation** : la création de normes harmonisées en matière d'IA aidera les fournisseurs de systèmes à haut risque à se conformer aux exigences du règlement de l'UE sur l'IA, mais leur compatibilité avec la législation sur les droits fondamentaux et la non-discrimination semble incertaine. L'interprétation de ces normes à la lumière des obligations de non-discrimination et l'évaluation de leur conformité sont d'une importance capitale, mais constitueront une tâche difficile qui nécessitera un certain degré d'expertise technique.

Plusieurs mesures clés pourraient être adoptées pour faire face aux risques de discrimination dans les systèmes d'IA et de PDA et améliorer (l'application des) cadres de gouvernance :

- **Établir ou renforcer les exigences en matière de prévention, de surveillance, de sensibilisation et de formation** est essentiel pour déployer des systèmes d'IA et de PDA qui soient conformes à la législation sur les droits fondamentaux. Il pourrait s'agir, par exemple, de mesures de formation et de sensibilisation à l'intention des administrations publiques qui développent ou déploient des systèmes d'IA ou de PDA, et de règles relatives aux tests, aux audits et à la surveillance de ces systèmes dans les procédures de passation des marchés publics.
- **Étendre les obligations légales de transparence.** Par exemple, les obligations d'enregistrement dans une base de données publique pourraient être étendues au-delà des systèmes à haut risque et les exceptions existantes à la publicité devraient être appliquées de manière restrictive. Les autorités publiques qui déploient des systèmes automatisés à fort impact (par exemple, conduisant à des décisions individuelles ou permettant des pratiques de surveillance de masse) pourraient être tenues d'en informer les parlements nationaux.
- **Réviser les lois anti-discrimination pour faciliter l'accès à la justice et à des recours** efficaces. La boîte à outils conceptuelle de la loi anti-discrimination devrait être révisée et étendue pour faciliter la reconnaissance des formes de discrimination qui sont particulièrement susceptibles de survenir dans le contexte des systèmes d'IA et de PDA, telles que la discrimination intersectionnelle et systémique. Les asymétries de pouvoir et d'information devraient être traitées par des règles de procédure adaptées, en particulier celles relatives au partage de la charge de la preuve. On peut soutenir l'idée que les difficultés rencontrées par les victimes pour établir des présomptions de discrimination algorithmique devraient conduire à abaisser encore le seuil de preuve pour établir des cas *prima facie*. Les obstacles à l'accès à la justice pourraient

également être levés en révisant les règles relatives aux actions collectives, par exemple en reconnaissant les actions de groupe. La collecte de données sur l'égalité devrait également être facilitée afin de soutenir des mécanismes de recours efficaces.

- **Le renforcement des capacités, des ressources et de la participation institutionnelle des organismes de promotion de l'égalité** est essentiel pour faciliter l'application des lois anti-discrimination en ce qui concerne les systèmes d'IA et de PDA. Les nouvelles normes de l'UE relatives aux organismes de promotion de l'égalité, notamment en ce qui concerne leurs compétences et leurs ressources, devraient être transposées à la lumière de ces nouveaux défis. L'allocation des ressources devrait permettre aux organismes de promotion de l'égalité de remplir leurs nouvelles fonctions (par exemple s'ils sont désignés en vertu de l'article 77 du règlement de l'UE sur l'IA), notamment en développant une expertise appropriée. Les organismes de promotion de l'égalité et les autorités de contrôle des droits fondamentaux devraient pouvoir participer pleinement et être représentés dans le cadre de gouvernance établi par les nouveaux règlements européens sur l'IA (par exemple, le Comité européen de l'IA, le Bureau de l'IA, le groupe scientifique d'experts indépendants). Les mécanismes de coopération devraient permettre une coopération multipartite, par exemple avec les autorités de surveillance du marché.

Les organismes de promotion de l'égalité ont un rôle crucial à jouer dans le contrôle et l'application des dispositions anti-discrimination dans les utilisations d'IA et de PDA. Ainsi, les organismes de promotion de l'égalité sont encouragés à :

- **Plaider pour leur inclusion** dans la liste des autorités nationales de contrôle des droits fondamentaux désignées **en vertu de l'article 77** du règlement de l'UE sur l'IA.
- **Utiliser leur droit d'accès à l'information** (article 77(1) du règlement de l'UE sur l'IA) pour rechercher les biais potentiels et les effets discriminatoires des systèmes d'IA et de PDA.
- **Exploiter les pouvoirs de surveillance prévus par le règlement** sur l'IA, notamment en demandant et en participant aux **tests effectués** sur les systèmes d'IA à haut risque.
- **Mettre en place des mécanismes de coopération** solides entre eux, avec les autorités de surveillance du marché, les organismes possédant une expertise complémentaire (agences de protection des données, bureaux de protection des consommateurs, autorités de surveillance des médias, etc.) et les tiers concernés (chercheurs, OSC, etc.). La coopération peut les aider à relever le défi du manque d'expertise et de ressources.
- **Tirer parti de leur expertise et de leur droit de participer aux consultations publiques pour influencer les réformes et les politiques futures** afin de façonner les réglementations et les politiques futures en matière d'IA et leur application.
- **Sensibiliser à la discrimination algorithmique** au sein des administrations publiques, du secteur privé et de la population en général.

