

Pour une IA européenne protectrice et garante du principe de non-discrimination

AVIS ÉTABLISSANT DES RECOMMANDATIONS ET DES
PRINCIPES ESSENTIELS POUR LA FUTURE LÉGISLATION
EUROPÉENNE PORTANT SUR L'INTELLIGENCE ARTIFICIELLE

Objectif et contexte

Les recommandations suivantes constituent les composantes principales du présent avis d'Equinet portant sur la proposition législative de la Commission européenne, qui introduit une approche fondée sur les risques pour traiter un certain nombre de questions éthiques et juridiques que soulève l'intelligence artificielle. [L'analyse d'impact initiale](#) de cette proposition législative décrit le premier but spécifique de l'objectif général de la proposition de la manière suivante : « [...] a) assurer l'application effective des règles de droit européen existantes destinées à protéger la sécurité et les droits fondamentaux et à éviter les discriminations illégales en établissant des obligations de documentation nécessaires à la mise en œuvre des règles européennes tant dans le secteur public que privé. »

Le principe de non-discrimination est présenté explicitement et de manière proéminente dans la proposition de règlement de la Commission européenne comme l'une des principales préoccupations en matière de droits humains liées à l'impact des systèmes d'IA. Le rôle des organismes de promotion de l'égalité devrait s'en retrouver affecté, notamment à travers la mise en œuvre probable de nouveaux mécanismes de contrôle d'application et de conformité spécifiques à l'IA et venant s'ajouter à leurs pouvoirs en matière de non-discrimination.

Basées sur le [rapport Equinet dédié à l'IA](#) ainsi que sur nos deux contributions à la consultation de la Commission européenne portant sur l'intelligence artificielle (l'une ciblée, l'autre plus générale), les recommandations du présent avis visent à refléter la position des organismes de promotion de l'égalité sur deux questions primordiales, à savoir : 1) comment prévenir les atteintes à l'égalité dans le contexte des systèmes algorithmiques dits d'intelligence artificielle et 2) comment garantir que le rôle et les pouvoirs des organismes de promotion de l'égalité soient renforcés et effectivement pris en compte au moment où se dessine une nouvelle infrastructure de contrôle visant à faire appliquer la loi et à permettre aux individus d'obtenir réparation dans le contexte des systèmes d'IA.

Cet avis doit être lu en parallèle avec [les ambitions et les propositions d'Equinet en matière de législation européenne dédiée aux organismes de promotion de l'égalité](#). Il considère les systèmes d'IA et leur future réglementation comme faisant partie intégrante d'une discussion plus large portant sur les implications des nouvelles technologies et de la dématérialisation vis-à-vis du respect du principe de non-discrimination, et propose une approche fondée sur l'égalité, en tant qu'exigence globale et horizontale, à travers toutes ses recommandations.

Recommandations

1. Faire du respect du principe de non-discrimination une préoccupation centrale dans toute réglementation européenne dédiée à l'IA.

De par leur nature même, les systèmes d'IA fonctionnent par le biais de mécanismes d'exclusion et de différenciation, et présentent donc des risques disproportionnés pour le respect des principes d'égalité et de non-discrimination entre les individus. Lorsqu'elle s'opère aux moyens de systèmes d'IA, la violation de ce principe produit un effet néfaste extrêmement important et systémique sur la société. Cela résulte des propriétés spécifiques des technologies dites d'intelligence artificielle. Celles-ci agissent à grande échelle et de manière non transparente et imprévisible, permettant aux violations du principe de non-discrimination de se produire tout au long de leur cycle de vie : lors de leur conception, de leur déploiement, comme de leur utilisation. Le risque d'atteinte au principe d'égalité que présentent les systèmes d'IA se trouve exacerbé lorsque ces derniers sont mobilisés pour guider la prise de décision dans certains domaines essentiels de la vie sociale comme l'emploi, l'éducation, la sécurité sociale ou le maintien de l'ordre public.

Afin de reconnaître et de relever les défis spécifiques et substantiels que présentent les systèmes d'IA pour le respect des principes d'égalité et de non-discrimination, le futur règlement sur l'intelligence artificielle (*Artificial Intelligence Act* ou AIA) devrait contenir des références explicites à la nécessaire protection de ce principe en tant que préoccupation majeure au regard de l'impératif respect des droits fondamentaux. Ce besoin est d'autant plus impérieux que les systèmes d'IA éprouvent déjà les capacités du cadre juridique existant en matière de non-discrimination à protéger efficacement les individus des discriminations dites algorithmiques ; a fortiori lorsque les catégories de différenciation générées par l'algorithme ne peuvent être appréhendées par les critères protégés du droit actuel de la non-discrimination. Pour remédier à ces limitations, en reconnaissant notamment explicitement la non-discrimination comme l'un de ses principaux objectifs en matière de droits fondamentaux, le futur règlement européen sur l'IA pourrait prévoir des mécanismes permettant d'identifier et de corriger l'ensemble des pratiques discriminatoires rendues possibles par l'utilisation de systèmes d'IA.

2. Établir des mécanismes de plainte et de recours accessibles et efficaces pour les personnes concernées en cas de violation des principes d'égalité et de non-discrimination ou d'autres droits fondamentaux lorsqu'une telle violation résulte de l'utilisation de systèmes d'IA.

Dans sa version actuelle, la proposition de règlement sur l'intelligence artificielle ne prévoit pas de mécanismes de recours permettant aux citoyens de déposer plainte et d'obtenir réparation de leurs préjudices. Cela est d'autant plus problématique que le refus d'accès à la justice pour les victimes est au cœur du défi que présentent les technologies d'intelligence artificielle en matière de non-discrimination et de droits humains. Pour faire face à cette situation, le futur règlement sur l'intelligence artificielle devrait contenir des dispositions explicites et claires instituant un droit de recours individuel et collectif, mais aussi un droit à obtenir réparation en cas de préjudice causé par tout système d'IA, quel que soit son niveau

de risque. Le futur règlement devrait également prévoir des conditions destinées à faciliter l'exercice de ce droit de recours effectif.

Ces conditions devraient inclure, a minima, la reconnaissance de droits individuels aux personnes victimes de systèmes d'IA, notamment le droit de ne pas être confronté à des systèmes d'IA non conformes au futur règlement sur l'IA, ainsi que le droit de recevoir une explication intelligible et accessible de la logique sous-jacente des décisions impliquant l'utilisation de systèmes d'IA.

A cet égard, le règlement sur l'IA devrait également prévoir un mécanisme clair et accessible destiné aux organisations d'intérêt public et aux autorités compétentes en matière de droits fondamentaux à l'image des organismes de promotion de l'égalité, permettant à celles-ci de déposer des réclamations directement auprès des autorités notifiantes nationales suite à la découverte de systèmes d'IA qui menaceraient le bon respect du principe de non-discrimination ou d'autres droits fondamentaux. Suite à une telle réclamation, le texte devrait également prévoir le déclenchement d'une enquête portant sur le système d'IA mis en cause.

La participation des organismes de promotion de l'égalité, à travers l'adoption d'une obligation légale exigeant leur consultation par les autorités notifiantes, consoliderait les possibilités de recours dans le cadre du futur règlement sur l'IA. Il assurerait la complémentarité entre le mécanisme de recours spécifique à l'IA et celui applicable en matière de non-discrimination, telle que la saisine d'un organisme de promotion de l'égalité. De par leur rôle de soutien et de conseil aux victimes de discriminations, notamment par le biais du traitement des saisines et des litiges, les organismes nationaux de promotion de l'égalité sont particulièrement bien outillés pour contribuer à la réparation effective des violations du principe de non-discrimination occasionnées par des systèmes d'IA.

3. Appliquer une approche fondée sur les droits fondamentaux pour définir les notions de « préjudice » et de « risque », et non une approche tirée des régimes de sécurité des produits.

Dans sa version actuelle, l'article 7 de la proposition de règlement sur l'IA définit le risque d'incidence négative et le risque de préjudice de manière floue et insuffisamment alignée sur le droit européen existant, notamment s'agissant du droit de la non-discrimination. La classification d'un système d'IA comme étant à « haut risque » est basée exclusivement sur la législation existante en matière de sécurité des produits et, par conséquent, ne reflète pas la nature unique du préjudice qui résulte des violations de droits fondamentaux. La protection contre les produits à haut risque du point de vue de la sécurité des consommateurs n'est pas la même que la protection contre les produits à haut risque du point de vue de leur incidence négative sur les droits fondamentaux. Les garanties juridiques pour ces risques distincts (sécurité des produits vs droits fondamentaux) devraient être différenciées en conséquence par le futur règlement sur l'IA, le principe de non-discrimination et les autres droits fondamentaux recevant une protection plus stricte et renforcée à travers des exigences

obligatoires rigoureuses. Le processus d'évaluation des risques, qui conduira à la détermination des exigences applicables, devrait impliquer les groupes protégés affectés par le système d'IA envisagé. Obligatoire, leur participation devrait être explicitement intégrée dans le texte du futur règlement sur l'IA.

Toute définition et catégorisation réglementaire du terme « préjudice » dans le but de protéger le principe d'égalité et les droits humains doit être fondée sur une compréhension plus globale et plus inclusive de la notion - une compréhension qui tienne compte de la nature immatérielle, sociétale et collective du préjudice lorsqu'il survient en matière de droits humains.

4. Exiger des analyses d'impact sur l'égalité ex ante et ex post à intervalles réguliers durant l'ensemble du cycle de vie des systèmes d'IA.

Des analyses d'impact obligatoires portant sur l'égalité devraient être réalisées afin de prédire et détecter les risques que présentent les systèmes d'IA dont l'utilisation montre des risques évidents de discrimination. Celles-ci devraient être effectuées à intervalles réguliers pendant tout le cycle de vie des dispositifs d'IA : de la phase de développement à la phase de déploiement et d'utilisation. Cela permettrait une approche préventive et proactive du principe d'égalité, empêchant ainsi que des individus et/ou des communautés entières ne soient transformées en victimes par ces systèmes. Compte tenu de la définition intrinsèque du principe d'égalité et de sa portée horizontale par rapport aux autres droits fondamentaux, sa protection par le biais d'analyses d'impact obligatoires ex ante et ex post devrait impliquer des tiers et ne pas être confiée uniquement aux développeurs et aux utilisateurs de systèmes d'IA.

L'actuel règlement général sur la protection des données (RGPD) fournit également une raison impérieuse de rendre les analyses d'impact sur l'égalité obligatoires dans le contexte des systèmes d'IA. L'article 35 (3)(a) du RGPD exige en effet la réalisation d'analyses d'impact sur la protection des données (AIPD) pour tous les systèmes d'apprentissage automatique. Ces systèmes constituant l'une des principales catégories de technologies basées sur l'IA, le futur règlement européen ne saurait faire l'économie de mesures d'analyses d'impact obligatoires. L'absence de telles dispositions conduirait à une protection fragmentée et incohérente contre les risques d'atteinte au respect du principe de non-discrimination et des droits humains induits par l'IA. Pour éviter cela et garantir la sécurité juridique, le futur règlement sur l'IA devrait créer une exigence d'analyses d'impact sur l'égalité ex ante et ex post à intervalles réguliers sur l'ensemble du cycle de vie des systèmes d'IA, qui soit alignée et qui tienne compte des résultats des AIPD effectuées au titre du RGPD.

5. Assigner des "obligations d'égalité" contraignantes et opposables à tous les concepteurs et utilisateurs d'IA.

Mettre l'égalité au cœur des processus décisionnels est une obligation légale au sens des articles 8 et 10 du traité sur le fonctionnement de l'UE. Cette obligation a été établie comme une priorité dans le récent agenda politique de l'UE à travers la mise en place d'une task force sur l'égalité spécifiquement chargée d'intégrer l'égalité dans toutes les politiques de l'UE. Les « obligations en matière d'égalité »¹ sont définies comme des obligations juridiques positives de promouvoir l'égalité et de prévenir la discrimination. Celles-ci vont au-delà de l'obligation générale de s'abstenir de toute discrimination.

Le fait d'instaurer de telles obligations en matière d'IA devrait renforcer l'effet des analyses d'impact sur l'égalité en créant des obligations nouvelles pour les concepteurs et les utilisateurs de systèmes d'IA. Ceux-ci pourraient en effet s'engager à surveiller de manière proactive les incidences des systèmes d'IA qu'ils développent et/ou utilisent sur le principe d'égalité. En outre, leurs propres pratiques institutionnelles pourraient s'en retrouver considérablement améliorées puisqu'ils seraient ainsi incités à recruter et à promouvoir des équipes spécialisées diversifiées.

La proposition de règlement sur l'IA devrait se fixer pour priorité de prévenir les violations des principes d'égalité et de non-discrimination occasionnées par les systèmes d'IA. À cette fin, des « obligations d'égalité » claires devraient être imposées aux utilisateurs et aux concepteurs de ces systèmes. Celles-ci devraient s'accompagner de garanties claires assurant leur application effective sous peine de sanctions efficaces, proportionnées et dissuasives.

6. Rendre la différenciation des risques possible uniquement après une analyse obligatoire de l'impact sur le principe de non-discrimination et les autres droits humains.

Les groupes protégés affectés par les systèmes d'IA devraient être consultés et impliqués tant dans la conception, le développement et l'utilisation des analyses d'impact sur les droits

¹ Les travaux d'Equinet sur les pratiques d'intégration de l'égalité des organismes de promotion de l'égalité en Europe, à savoir « [Compendium of Good Practices in Equality Mainstreaming](#) » (2021) et « [Construire une Europe plus égale : une obligation légale ?](#) » (2016), ont identifié trois grandes catégories d'obligations légales en matière d'égalité en place en Europe : 1) Les obligations préventives sont des obligations légales imposées aux organisations (publiques et privées) de prendre des mesures pour prévenir la discrimination, le harcèlement ou le harcèlement sexuel dans l'emploi ou dans la fourniture de biens et de services. 2) Les obligations institutionnelles sont des obligations légales imposées aux organisations (publiques et privées) pour promouvoir l'égalité pour les employés ou pour les personnes accédant à leurs services. 3) Les obligations de transversalité politique exigent des autorités publiques qu'elles tiennent dûment compte de la nécessité de promouvoir l'égalité dans l'exercice de leurs fonctions, notamment en matière de législation, de budgétisation, de réglementation et d'élaboration des politiques publiques. Les obligations en matière d'égalité concernant le développement et l'utilisation des systèmes d'IA devraient englober **les trois types d'obligations - préventive** (titulaires d'obligations publiques et privées), **institutionnelle** (titulaires d'obligations publiques et privées) et **de transversalité** (titulaires d'obligations publiques au minimum).

humains. Obligatoires - tant ex ante qu'ex post - celles-ci devraient en tant que telles servir de base afin de définir les différents niveaux de risque que présentent les systèmes algorithmiques. Ce n'est effectivement qu'une fois que les risques auront été clairement identifiés et évalués à travers ces analyses que des mesures basées sur leur différenciation pourront être autorisées dans le cadre du futur règlement sur l'intelligence artificielle. Le texte pourrait ainsi permettre de moduler les approches notamment en matière de surveillance, de voies de recours et de sanctions en fonction des niveaux de risque identifiés sur la base de ces analyses d'impact récurrentes.

7. Rendre effective l'exécution des dispositions du futur règlement sur l'IA en obligeant les nouvelles autorités de surveillance nationales à consulter les organismes de promotion de l'égalité ainsi que les autres institutions compétentes en matière de droits fondamentaux

Si l'un des principaux objectifs de la proposition de règlement est de protéger contre le risque d'incidence négative des systèmes d'IA sur les droits fondamentaux, alors les nouvelles autorités de surveillance du marché, seules ou en collaboration avec d'autres régulateurs sectoriels nationaux dont l'activité ne comporte aucune composante propre au respect des droits fondamentaux, ne suffiront guère à garantir une exécution ex post effective du règlement qui soit protectrice de ces mêmes droits.

Les autorités spécialisées dans la mise en œuvre et le contrôle du respect des droits fondamentaux, dont font partie les organismes de promotion de l'égalité, sont mieux équipées pour protéger ces droits que d'autres régulateurs nationaux dont le mandat ne relève pas directement de ces questions. Par conséquent, les organismes nationaux de promotion de l'égalité devraient être consultés chaque fois que l'application du futur règlement sur l'IA concernera la protection du principe de non-discrimination.

Afin d'assurer une application cohérente et robuste du futur règlement sur l'IA dans chacune des différentes juridictions de l'UE comme entre elles, le texte devrait établir et clarifier les rôles spécifiques et la répartition institutionnelle des compétences entre les multiples autorités de contrôle nationales impliquées dans sa mise en œuvre. Cela concerne en particulier les autorités de surveillance du marché (autorités notifiantes), les autorités de protection des données, les organismes de promotion de l'égalité ainsi que les autres autorités compétentes en matière de droits fondamentaux. Dans sa version actuelle, la proposition de la Commission ne prévoit pas de structure claire pour différencier et coordonner les missions de contrôle et d'application entre les multiples autorités nationales compétentes dont l'implication est nécessaire pour assurer la mise en œuvre effective du futur règlement horizontal.

La seule mention explicite de la coopération entre les autorités nationales de contrôle et d'application existantes se limite aux deux références suivantes : 1) le pouvoir de demander

et d'accéder à toute documentation conservée en vertu du règlement ; 2) le cas échéant, la possibilité de demander aux autorités de surveillance du marché d'organiser les tests d'un système d'IA à haut risque en fournissant les moyens techniques nécessaires à leur réalisation. Ces dispositions sèment une importante confusion et risquent d'aboutir à des doubles emplois comme à d'importantes difficultés inter-institutionnelles, affaiblissant à terme l'application effective du futur règlement.

8. [Rendre obligatoire la mise en place et le financement adéquat de mécanismes de coopération entre les différents organismes impliqués dans l'application du règlement sur l'IA au niveau européen comme national.](#)

Il apparaît indispensable de se doter d'une structure de gouvernance nationale sur l'IA. Celle-ci pourrait prendre la forme d'un cadre de coopération des autorités nationales compétentes en vertu de la proposition de règlement sur l'IA afin de leur permettre de collaborer et d'acquérir l'expertise nécessaire. Les organismes de promotion de l'égalité sont des partenaires clés dans l'application de la législation nationale sur les droits humains et, à ce titre, devraient faire partie de ce cadre de gouvernance nationale.

Les systèmes d'IA ont une nature complexe et une utilisation intersectorielle. Cela signifie que la protection du principe d'égalité contre les menaces liées à l'utilisation de systèmes d'IA nécessite une collaboration active par la voie de partenariats. Par conséquent, les autorités nationales de surveillance suggérées par la proposition de l'UE sur l'IA ne sauraient travailler en vase clos.

Le règlement européen sur l'IA a donc un rôle clé à jouer pour encourager les gouvernements à mettre en place les structures interinstitutionnelles nécessaires permettant à ces partenariats et échanges d'exister sur une base régulière et durable.

Les organismes de promotion de l'égalité seront des partenaires clés dans sa mise en œuvre et, à ce titre, devront travailler de concert avec les organismes spécialisés dans la surveillance et le contrôle de l'IA comme avec les autres autorités qui participeront à l'application effective du règlement au niveau national.

Comme le soulignent [le rapport du réseau Equinet dédié à l'IA](#) ainsi que l'avis du même réseau intitulé « [Summary and Framework for Action for Equality Bodies](#) », les systèmes d'IA sont complexes et leur utilisation est intersectorielle. Cela signifie que la protection du principe d'égalité contre les menaces liées à l'IA nécessite une collaboration active et l'établissement de plusieurs partenariats d'envergure. Les organismes nationaux de promotion de l'égalité doivent en effet être en mesure de travailler efficacement avec les autorités nationales et régionales ainsi qu'avec des acteurs généralement considérés comme non traditionnels afin de préserver le principe de non-discrimination.



Ces partenariats pourraient concerner des régulateurs sectoriels comme les autorités de protection des données, les autorités de protection des consommateurs, mais aussi des acteurs du secteur privé, du monde universitaire, d'ONG de défense des droits fondamentaux à l'ère du numérique, d'organismes de standardisation ou encore des ingénieurs informaticiens spécialistes des algorithmes et des données. L'UE a un rôle clé à jouer dans la création d'un environnement institutionnel favorable à l'établissement de tels partenariats.