

BIEN PRÉPARER L'AVENIR

L'INTELLIGENCE ARTIFICIELLE ET LES DROITS FONDAMENTAUX

RÉSUMÉ



3

Principales conclusions et avis de la FRA

5

La sauvegarde des droits fondamentaux:
champ d'application, analyses d'impact
et responsabilité

10

Non-discrimination, protection des données
et accès à la justice: trois thématiques
horizontales

© Agence des droits fondamentaux de l'Union européenne, 2021.

Reproduction autorisée, moyennant mention de la source

Toute utilisation ou reproduction de photos ou de tout autre matériel dont l'Agence des droits fondamentaux de l'Union européenne ne possède pas les droits d'auteur requiert l'autorisation préalable des titulaires des droits en question.

Ni l'Agence des droits fondamentaux de l'Union européenne ni aucune personne agissant au nom de l'Agence n'est responsable de l'usage qui pourrait être fait des informations données ci-après.

Luxembourg: Office des publications de l'Union européenne, 2021.

Print	ISBN 978-92-9461-226-7	doi:10.2811/088	TK-04-20-657-FR-C
PDF	ISBN 978-92-9461-233-5	doi:10.2811/7549	TK-04-20-657-FR-N

Crédits photo:

Couverture: © HQUALITY/Adobe Stock
Page 1: © Mykola Mazuryk/AdobeStock
Page 3: © Mimi Potter/AdobeStock
Page 4: © Sikov/Adobe Stock
Page 7: © Gorodenkoff/Adobe Stock
Page 9: © VideoFlow/Adobe Stock
Page 11: © Monsitj/Adobe Stock

«L'intelligence artificielle est une excellente chose, mais nous devons apprendre à l'utiliser.»

(Entreprise privée, Espagne)

«Le risque est de trop se fier à la machine.»

(Administration publique, France)

L'intelligence artificielle (IA) est de plus en plus utilisée dans les secteurs privé et public, ce qui a une incidence sur la vie quotidienne. D'aucuns associent l'IA à la fin du contrôle des machines par l'homme. D'autres y voient la technologie qui aidera l'humanité à relever certains de ses défis les plus urgents. Si aucune de ces interprétations n'est exacte, il ne fait aucun doute que les préoccupations concernant l'impact de l'IA sur les droits fondamentaux ne cessent de croître, appelant à un examen minutieux de son utilisation par les acteurs des droits humains.

Le rapport de l'Agence des droits fondamentaux de l'Union européenne (FRA) intitulé *Bien préparer l'avenir: l'intelligence artificielle et les droits fondamentaux* offre un aperçu de l'utilisation qui est faite aujourd'hui des technologies liées à l'IA au sein de l'Union, et analyse ses implications sur les droits fondamentaux. Il aborde principalement des cas concrets d'utilisation dans quatre domaines essentiels: les prestations sociales, la police prédictive, les services de soins de santé et la publicité ciblée.

Le présent résumé décrit les principales idées exposées dans le rapport.

Définition de l'IA

Il n'existe aucune définition de l'IA qui soit universellement acceptée. L'IA ne désigne pas des applications concrètes, mais reflète plutôt des évolutions technologiques récentes qui englobent tout un éventail de technologies.

Dans le cadre de ses travaux de recherche, la FRA n'a pas appliqué une définition stricte de l'IA aux cas d'utilisation qu'elle présente dans son rapport principal. Aux fins des entretiens, elle a retenu une définition large de l'IA s'appuyant en partie sur celle fournie par le groupe d'experts de haut niveau sur l'intelligence artificielle de la Commission européenne:

«L'intelligence artificielle (IA) fait référence à des systèmes qui affichent un comportement intelligent en analysant leur environnement et en prenant des mesures — avec un certain degré d'autonomie — pour atteindre des objectifs spécifiques. Les systèmes basés sur l'IA peuvent reposer exclusivement sur des logiciels et être utilisés dans le monde virtuel (par exemple les assistants vocaux, les logiciels d'analyse d'images, les moteurs de recherche, les systèmes de reconnaissance vocale et faciale) ou l'IA peut être intégrée dans des produits matériels (par exemple les robots avancés, les voitures autonomes, les drones ou les applications de l'internet des objets).»

Cette **définition initiale** du groupe d'experts de haut niveau sur l'intelligence artificielle a fait l'objet d'une discussion plus approfondie au sein du groupe. Voir la publication du groupe d'experts de haut niveau sur l'intelligence artificielle (2019), «**Une définition de l'IA: fonctions principales et disciplines**».

Quels étaient les domaines couverts par la recherche?

La FRA a mené une recherche sur le terrain dans cinq États membres de l'Union européenne (UE): l'Estonie, l'Espagne, la France, les Pays-Bas et la Finlande. Elle a recueilli des informations auprès des personnes impliquées dans la conception et l'utilisation de systèmes d'IA dans des secteurs privés et publics essentiels, afin de mieux connaître la manière dont elles traitent les questions pertinentes relatives aux droits fondamentaux.

Cette recherche, qui reposait sur 91 entretiens individuels, a permis de recueillir des informations sur:

- la finalité et l'application pratique des technologies d'IA;
- les évaluations menées lors de l'utilisation de l'IA ainsi que le cadre juridique et les mécanismes de contrôle applicables;
- la connaissance des questions relatives aux droits fondamentaux et des éventuelles garanties mises en place; et
- les projets pour l'avenir.

Par ailleurs, dix experts spécialisés dans la surveillance ou l'observation des violations potentielles des droits fondamentaux dans le cadre de l'utilisation de l'IA, et notamment des acteurs de la société civile, des avocats et des organes de contrôle, ont été interrogés.

Pour accéder à une description plus détaillée de la méthodologie de recherche et des questions posées lors des entretiens, se reporter à l'annexe 1 du rapport principal, consultable sur le [site web de la FRA](#).

«Ce qui importe le plus, c'est de traiter les dossiers de manière plus efficace. Il s'agit d'utiliser les effectifs, c'est-à-dire les personnes qui traitent les dossiers, de la manière la plus efficace possible.»
(Administration publique, Pays-Bas)

«Lorsque nous avons testé le système, nous n'avons pas vraiment regardé les aspects juridiques, mais plutôt la rentabilité du système.»
(Entreprise privée, Estonie)

Travaux de la FRA sur l'IA, les mégadonnées et les droits fondamentaux

Le rapport de la FRA sur l'IA et les droits fondamentaux constitue la principale publication découlant du [projet de la FRA sur l'intelligence artificielle, les mégadonnées et les droits fondamentaux](#). Ledit projet vise à évaluer les aspects positifs et négatifs, pour les droits fondamentaux, des nouvelles technologies, y compris l'IA et les mégadonnées.

Le rapport s'appuie sur les conclusions d'un certain nombre de travaux antérieurs:

- *Facial recognition technology: fundamental rights considerations in the context of law enforcement* [Technologie de reconnaissance faciale: considérations relatives aux droits fondamentaux dans le cadre du maintien de l'ordre] (2019, en anglais): ce document présente et analyse les défis en matière de droits fondamentaux liés à la mise en place, par les pouvoirs publics, de la technologie de reconnaissance faciale en temps réel à des fins répressives. Il présente aussi succinctement les mesures à adopter pour éviter les violations des droits.
- *Data quality and artificial intelligence — mitigating bias and error to protect fundamental rights* [Qualité des données et intelligence artificielle — atténuer les biais et les erreurs pour protéger les droits fondamentaux] (2019, en anglais): ce document met en exergue l'importance de prendre conscience des problèmes liés à la mauvaise qualité des données et de les éviter.
- *#BigData: Discrimination in data-supported decision making* [Mégadonnées: la discrimination dans la prise de décision étayée par des données] (2018, en anglais): ce document de réflexion traite de la manière dont peut survenir une telle discrimination et propose d'éventuelles solutions.

Dans le cadre de ce projet, la FRA envisage également la possibilité d'étudier des exemples concrets de défis en matière de droits fondamentaux liés à l'utilisation d'algorithmes pour la prise de décision, par le biais d'expériences en ligne ou d'études de simulation.



Cadre juridique

Le cadre général des droits fondamentaux (*) qui s'applique à l'utilisation de l'IA dans l'UE est composé de la charte des droits fondamentaux de l'Union européenne (la Charte) et de la convention européenne des droits de l'homme.

Ce cadre est complété par de nombreux autres instruments du Conseil de l'Europe et d'organismes internationaux se référant aux droits humains. Il s'agit notamment de la déclaration universelle des droits de l'homme de 1948 et des principales conventions des Nations unies relatives aux droits de l'homme (**).

Le droit européen dérivé propre à chaque secteur, notamment l'acquis de l'UE en matière de protection des données et la législation européenne en matière de non-discrimination, contribue également à la sauvegarde des droits fondamentaux dans le contexte de l'IA. Enfin, la législation nationale des États membres de l'UE s'applique également.

(*) *Pour de plus amples informations, voir FRA (2012), **Donner corps aux droits: le paysage des droits fondamentaux dans l'Union européenne**, Luxembourg, Office des publications de l'Union européenne.*

(**) *Ces conventions majeures incluent: le pacte international relatif aux droits civils et politiques de 1966, le pacte international relatif aux droits économiques, sociaux et culturels de 1966, la convention internationale sur l'élimination de toutes les formes de discrimination raciale de 1965, la convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes de 1979, la convention contre la torture de 1984, la convention relative aux droits de l'enfant de 1989, la convention relative aux droits des personnes handicapées de 2006, et la convention internationale pour la protection de toutes les personnes contre les disparitions forcées de 2006.*

*Pour de plus amples informations sur le cadre juridique international universel relatif aux droits de l'homme, y compris leurs mécanismes d'application, se reporter, par exemple, à De Schutter, O. (2015), *International Human Rights Law: Cases, Materials, Commentary*, Cambridge, Cambridge University Press, 2^e édition.*

LA SAUVEGARDE DES DROITS FONDAMENTAUX: CHAMP D'APPLICATION, ANALYSES D'IMPACT ET RESPONSABILITÉ

Prendre en considération l'intégralité du champ d'application des droits fondamentaux au regard de l'IA

Le recours aux systèmes d'IA a un impact sur un large éventail de droits fondamentaux, quel que soit le domaine d'application concerné. Parmi ces droits figurent notamment, mais pas seulement, le respect de la vie privée, la protection des données, la non-discrimination et l'accès à la justice.

La charte des droits fondamentaux de l'UE (la Charte) est devenue juridiquement contraignante en décembre 2009 et possède la même valeur juridique que les traités de l'UE. Elle regroupe, au sein d'un même texte, les droits civils, politiques, économiques et sociaux. Conformément à l'article 51, paragraphe 1, de la Charte, les institutions, organes et organismes de l'Union sont tenus de respecter tous les droits énoncés dans la Charte. Les États membres de l'UE doivent faire de même lorsqu'ils mettent en œuvre le droit de l'Union. Cela vaut pour l'IA comme pour n'importe quel autre domaine.

Le travail mené sur le terrain dans le cadre de cette recherche montre que des systèmes très divers sont utilisés sous l'appellation d'IA. Les technologies analysées impliquent différents niveaux d'automatisation et de complexité. Elles varient également en termes d'échelle et d'impact potentiel sur les personnes.

Les conclusions de la FRA montrent que, quel que soit le domaine d'application concerné, l'utilisation de systèmes d'IA a des implications sur une large palette de droits fondamentaux. Parmi ces droits figurent notamment, mais pas seulement, le respect de la vie privée, la protection des données, la non-discrimination et l'accès à la justice. Néanmoins, ainsi qu'il ressort des entretiens, au moment d'aborder les conséquences de l'IA au regard des droits fondamentaux, l'étendue de l'impact de l'IA se limite souvent à certains droits spécifiques.

Or, un éventail de droits plus large doit être pris en considération dans le cadre de l'utilisation de l'IA, en fonction de la technologie et du domaine d'utilisation. Outre les droits relatifs à la protection de la vie privée et des données, à l'égalité et à la non-discrimination, ainsi qu'à l'accès à la justice, d'autres droits pourraient être pris en considération. On peut citer, par exemple, le droit à la dignité humaine, le droit à la sécurité sociale et à l'assistance sociale, le droit à une bonne administration (qui concerne tout particulièrement le secteur public) et la protection des consommateurs (particulièrement importante pour les entreprises). Selon le contexte dans lequel l'IA est utilisée, il convient de prendre en considération les autres droits protégés par la Charte.



AVIS 1 DE LA FRA

Lorsqu'ils introduisent de nouvelles politiques et adoptent de nouvelles lois sur l'IA, le législateur européen et les États membres, agissant dans le cadre du droit européen, doivent veiller à ce que le respect de l'ensemble des droits fondamentaux, tels que consacrés par la Charte et les traités de l'UE, soit pris en considération. Les politiques et les législations pertinentes doivent être assorties de garanties spécifiques en matière de droits fondamentaux.

Ce faisant, l'UE et ses États membres devraient se fonder sur des preuves solides de l'impact de l'IA sur les droits fondamentaux de façon à garantir que toute restriction de certains droits fondamentaux respecte les principes de nécessité et de proportionnalité.

Des garanties pertinentes doivent être prévues par la loi afin de protéger efficacement contre toute atteinte arbitraire aux droits fondamentaux et d'offrir une sécurité juridique tant aux développeurs qu'aux utilisateurs de l'IA. Des programmes volontaires de respect et de sauvegarde des droits fondamentaux dans le cadre du développement et de l'utilisation de l'IA peuvent contribuer à atténuer les violations des droits. Conformément aux exigences minimales en matière de clarté juridique – en tant que principe fondamental de l'état de droit et condition préalable à la garantie des droits fondamentaux –, le législateur doit faire preuve de la diligence nécessaire dans la définition du champ d'application d'une telle législation sur l'IA.

Compte tenu de la diversité des technologies regroupées sous le terme d'IA et du manque de connaissances sur l'étendue réelle de son impact potentiel sur les droits fondamentaux, il pourrait être nécessaire d'évaluer régulièrement la définition juridique des termes liés à l'IA.



AVIS 2 DE LA FRA

Le législateur européen devrait envisager de rendre obligatoire la réalisation d'analyses d'impact couvrant la totalité des droits fondamentaux. Ces analyses devraient couvrir les secteurs privé et public et être mises en œuvre en amont de l'utilisation d'un système d'IA. Les analyses d'impact devraient tenir compte de la nature et du champ d'application variables des technologies d'IA, notamment du niveau d'automatisation et de complexité, ainsi que du préjudice potentiel. Elles devraient inclure des exigences de contrôle élémentaires pouvant également contribuer à sensibiliser aux implications potentielles pour les droits fondamentaux.

Les analyses d'impact devraient s'inspirer des bonnes pratiques établies dans d'autres domaines et être réitérées régulièrement pendant le déploiement, le cas échéant. Ces analyses devraient être menées en toute transparence. Leurs résultats et recommandations devraient, dans la mesure du possible, être rendus publics. En vue de faciliter la procédure d'analyse d'impact, les entreprises et l'administration publique devraient être tenues de recueillir les informations nécessaires à une évaluation approfondie de l'impact potentiel sur les droits fondamentaux.

L'UE et les États membres devraient envisager des actions ciblées pour soutenir ceux qui développent, utilisent ou envisagent d'utiliser des systèmes d'IA, de façon à garantir le respect effectif de leurs obligations en matière d'analyse d'impact sur les droits fondamentaux. Ces actions pourraient notamment passer par des financements, des lignes directrices, des formations ou des actions de sensibilisation. Elles devraient cibler plus particulièrement, mais pas exclusivement, le secteur privé.

L'UE et les États membres devraient envisager d'utiliser les outils existants, tels que les listes de contrôle ou les outils d'autoévaluation, développés à l'échelle européenne et internationale. Parmi ces outils figurent ceux élaborés par le groupe d'experts de haut niveau sur l'intelligence artificielle de l'UE.

Utiliser des analyses d'impact efficaces pour prévenir les effets négatifs

Les analyses d'impact antérieures portent principalement sur des questions techniques. Elles abordent rarement les conséquences potentielles sur les droits fondamentaux. Cela s'explique par le manque de connaissances sur la manière dont l'IA affecte ces droits.

Le déploiement de systèmes d'IA affecte un large éventail de droits fondamentaux, quel que soit le domaine d'application concerné. Conformément à l'article 51, paragraphe 1, de la Charte, les États membres de l'UE sont tenus de respecter tous les droits énoncés dans la Charte lorsqu'ils mettent en œuvre le droit de l'Union. Conformément aux normes internationales existantes (notamment les principes directeurs des Nations unies relatifs aux entreprises et aux droits de l'homme), les entreprises devraient mettre en place «[u]ne procédure de diligence raisonnable en matière de droits de l'homme pour identifier leurs incidences sur les droits de l'homme, prévenir ces incidences et en atténuer les effets, et rendre compte de la manière dont elles y remédient» (principes 15 et 17). Ces principes s'appliquent indépendamment de la taille et du secteur d'activité des entreprises et concernent dès lors celles qui utilisent l'IA.

Soucieuse de respecter ses engagements vis-à-vis des principes directeurs des Nations unies, l'UE a adopté plusieurs actes législatifs portant sur des instruments sectoriels spécifiques, notamment dans le cadre des obligations de diligence raisonnable en matière de droits humains. Des discussions ont actuellement lieu pour proposer un nouveau droit dérivé de l'UE. Ce droit obligerait les entreprises à faire preuve de diligence raisonnable au regard des conséquences potentielles de leurs activités et de leurs chaînes d'approvisionnement sur les droits humains et sur l'environnement. Cette législation serait probablement intersectorielle et prévoirait des sanctions en cas de non-respect, ce qui devrait concerner également l'utilisation de l'IA. Se reporter au récent rapport de la FRA intitulé *Business and human rights – access to remedy* [Entreprises et droits de l'homme – amélioration de l'accès aux voies de recours], qui appelle à améliorer les règles horizontales de diligence en matière de droits humains dans les entreprises établies dans l'UE.

Les analyses d'impact constituent un outil important aussi bien pour les entreprises que pour l'administration publique afin d'atténuer le potentiel impact négatif de leurs activités sur les droits fondamentaux. Dans certains secteurs, la législation de l'UE prévoit certaines formes d'analyses d'impact, notamment des analyses d'impact relatives à la protection des données en vertu du règlement général sur la protection des données (RGPD). De nombreuses personnes interrogées ont déclaré avoir procédé à une analyse d'impact relative à la protection des données, comme le prévoit la loi. Ces analyses prennent cependant différentes formes. Qui plus est, les analyses préalables, lorsqu'elles sont menées, se concentrent principalement sur les aspects techniques. Elles abordent rarement les impacts potentiels sur les droits fondamentaux. D'après certaines personnes interrogées, les analyses d'impact relatives aux droits fondamentaux ne sont pas réalisées dès lors qu'un système d'IA n'a pas, ou ne semble pas avoir, d'incidence négative sur les droits fondamentaux.

La recherche montre que les connaissances des personnes interrogées sur les droits fondamentaux (en dehors des droits à la protection des données et, dans une moindre mesure, à la non-discrimination) sont limitées. La majorité d'entre elles reconnaissent toutefois que le recours à l'IA a une incidence sur les droits fondamentaux. Certaines personnes interrogées indiquent que leurs systèmes n'ont pas d'incidence sur les droits fondamentaux, ce qui, dans une certaine mesure, a trait à la nature des tâches pour lesquelles les systèmes d'IA sont utilisés.

Toutes les personnes interrogées sont conscientes des problématiques associées à la protection des données. La plupart des répondants reconnaissent également que la discrimination pourrait, de manière générale, constituer un problème lors de l'utilisation de l'IA. Toutefois, la signification et l'applicabilité exactes des droits associés à la protection des données et à la non-discrimination restent floues pour bon nombre des personnes interrogées.

Les conclusions de la recherche font état de différences entre le secteur privé et le secteur public. Ainsi, les personnes interrogées issues du secteur privé connaissent souvent moins bien les différents droits fondamentaux qui sont susceptibles d'être affectés. Les questions relatives à la protection des données sont connues du secteur privé. Toutefois, d'autres droits, tels que le droit à la non-discrimination ou à l'accès à la justice, sont moins bien connus des représentants d'entreprises ayant recours à l'IA. Certains étaient parfaitement conscients des problèmes potentiels. D'autres, en revanche, ont déclaré qu'il incombait à leurs clients de vérifier les questions relatives aux droits fondamentaux.



AVIS 3 DE LA FRA

L'UE et les États membres devraient veiller à ce que des systèmes de responsabilisation efficaces soient mis en place pour contrôler et, le cas échéant, traiter efficacement tout impact négatif des systèmes d'IA sur les droits fondamentaux. Outre les analyses d'impact sur les droits fondamentaux (se reporter à l'avis 2 de la FRA), ils devraient envisager d'instaurer des garanties spécifiques pour s'assurer de l'efficacité du programme de responsabilisation. Cela pourrait passer par l'obligation légale de mettre à disposition suffisamment d'informations pour permettre une analyse de l'impact des systèmes d'IA sur les droits fondamentaux. Cela permettrait aux organes compétents d'exercer un contrôle externe et d'assurer une surveillance des droits humains.

L'UE et les États membres devraient également faire un meilleur usage des structures d'experts existantes en matière de contrôle afin de protéger les droits fondamentaux lors de l'utilisation de l'IA. Ces structures incluent notamment des autorités chargées de la protection des données, des organismes de promotion de l'égalité, des institutions nationales de défense des droits de l'homme, des médiateurs et des organismes de protection des consommateurs.

Des ressources supplémentaires devraient être affectées à la mise en place de systèmes de responsabilisation efficaces en renforçant les compétences et en diversifiant le personnel travaillant pour les organes de contrôle. Cela leur permettrait de traiter des questions complexes liées au développement et à l'utilisation de l'IA.

De la même façon, les organes appropriés devraient être dotés de ressources, de pouvoirs et, plus important encore, de connaissances suffisantes pour prévenir et évaluer les violations des droits fondamentaux et soutenir efficacement ceux dont les droits fondamentaux sont affectés par l'IA.

Faciliter la coopération entre les organismes appropriés à l'échelle nationale et européenne peut contribuer au partage d'expertise et d'expérience. L'établissement d'un dialogue avec d'autres acteurs ayant une expertise pertinente (tels que les organisations spécialisées de la société civile) peut également se révéler utile. Lors de la mise en œuvre de ces actions au niveau national, les États membres devraient envisager d'utiliser les mécanismes de financement de l'UE disponibles.

Garantir un contrôle efficace et une responsabilité globale

Les entreprises et les administrations publiques qui développent et utilisent l'IA sont en contact avec divers organes chargés de superviser les systèmes liés à l'IA dans le cadre de leurs mandats et secteurs respectifs. Ces organes incluent les autorités chargées de la protection des données. Cependant, les utilisateurs de l'IA ne savent pas toujours quels sont les organes chargés du contrôle des systèmes d'IA.

Conformément à des normes internationales bien établies en matière de droits de l'homme [par exemple l'article 1^{er} de la convention européenne des droits de l'homme (CEDH) et l'article 51 de la Charte], les États sont tenus de protéger les droits et libertés des personnes. Pour s'y conformer efficacement, les États doivent notamment mettre en place des mécanismes de contrôle et d'exécution efficaces. Cela vaut également pour l'IA.

S'agissant de la surveillance, les conclusions mettent l'accent sur le rôle important des organes spécialisés qui ont été créés dans certains secteurs spécifiques et qui, dans le cadre de leur mandat, sont également responsables du contrôle de l'IA. Il s'agit, par exemple, de la surveillance dans le domaine bancaire ou des autorités chargées de la protection des données. Diverses instances de ce type pourraient jouer un rôle dans le contrôle de l'IA du point de vue des droits fondamentaux. Cependant, pour de nombreux répondants issus du secteur privé comme du secteur public, les responsabilités de ces organes en matière de contrôle de l'IA restent floues.

L'utilisation de l'IA par les administrations publiques est parfois contrôlée dans le cadre des audits réguliers dont elles font l'objet. Les entreprises privées de secteurs spécifiques disposent également d'organes de contrôle spécialisés, par exemple dans le domaine de la santé ou des services financiers. Ces derniers vérifient également l'utilisation de l'IA et des technologies afférentes, par exemple dans le cadre de leurs programmes de certification. Les personnes interrogées issues du secteur privé ont exprimé le souhait de pouvoir s'adresser à des organes capables de leur fournir des conseils d'experts sur les possibilités et la légalité des utilisations potentielles de l'IA.

Il existe, au sein de l'UE, un ensemble bien défini d'organismes indépendants dont le mandat consiste à protéger et promouvoir les droits fondamentaux. Parmi ces organismes, on trouve les autorités chargées de la protection des données, des organismes de promotion de l'égalité, des institutions nationales des droits de l'homme et des médiateurs. Les travaux de recherche montrent que les personnes qui utilisent ou prévoient d'utiliser l'IA ont souvent contacté différents organismes au sujet de leur utilisation de l'IA, tels que des organismes de protection des consommateurs.

S'agissant du traitement de données à caractère personnel, dans la majorité des cas, les utilisateurs de l'IA ont contacté les autorités chargées de la protection des données pour obtenir des conseils, des informations ou des autorisations. Les experts interrogés soulignent le rôle important des autorités chargées de la protection des données dans le contrôle des systèmes d'IA en ce qui concerne l'utilisation des données à caractère personnel. Ils notent cependant aussi que les autorités chargées de la protection des données manquent de ressources, mais également de connaissances spécifiques sur les questions d'IA, pour mener à bien cette mission.

Les experts, y compris ceux qui travaillent pour des organes de contrôle, tels que les organismes de promotion de l'égalité et les autorités chargées de la protection des données, s'accordent à dire que l'expertise des organes de contrôle existants devrait être renforcée pour leur permettre d'assurer un contrôle efficace des questions liées à l'IA. Mais selon les experts, la tâche s'annonce difficile car les ressources de ces organismes sont déjà utilisées au maximum. Ils ont également souligné le rôle important que peuvent jouer certaines organisations de la société civile spécialisées dans les domaines de la technologie, des droits numériques et des algorithmes, en renforçant la responsabilisation dans l'utilisation des systèmes d'IA.



NON-DISCRIMINATION, PROTECTION DES DONNÉES ET ACCÈS À LA JUSTICE: TROIS THÉMATIQUES HORIZONTALES

Les travaux de recherche montrent que l'utilisation de l'IA nuit à divers droits fondamentaux. Outre les aspects spécifiques liés au contexte qui ont une incidence sur différents droits à des degrés divers, les thématiques relatives aux droits fondamentaux qui se sont dégagées de la recherche et qui s'appliquent de manière répétée à la plupart des cas d'IA incluent: la nécessité de garantir une utilisation non discriminatoire de l'IA (droit à la non-discrimination), l'obligation de traiter les données légalement (droit à la protection des données à caractère personnel) et la possibilité de contester les décisions fondées sur l'IA et de demander réparation (droit à un recours effectif et à un procès équitable).

Les deux principaux droits fondamentaux mis en avant pendant les entretiens sont la protection des données et la non-discrimination. En outre, la mise en place de moyens efficaces pour contester l'utilisation de l'IA a été évoquée à plusieurs reprises, en lien avec le droit à un procès équitable et à un recours effectif. Les trois avis suivants de la FRA, qui reflètent ces conclusions, devraient être lus en combinaison avec les autres avis, qui appellent à une reconnaissance plus complète de l'ensemble des droits fondamentaux affectés par l'IA et à une réponse à cette situation.

AVIS 4 DE LA FRA

Les États membres de l'UE devraient envisager d'encourager les entreprises et l'administration publique à évaluer les résultats susceptibles d'être discriminatoires lors de l'utilisation de systèmes d'IA.

La Commission européenne et les États membres devraient envisager de financer des recherches ciblées sur les effets potentiellement discriminatoires de l'utilisation de l'IA et des algorithmes. Ces recherches bénéficieraient de l'adaptation de méthodologies de recherche établies, issues des sciences sociales et qui sont utilisées pour identifier les discriminations potentielles dans différents domaines, allant du recrutement à l'établissement de profils clients.

En s'appuyant sur les résultats de ces recherches, il conviendrait d'élaborer des orientations et des outils afin d'aider ceux qui utilisent l'IA à détecter d'éventuels résultats discriminatoires.

Des garanties spécifiques pour garantir la non-discrimination lors de l'utilisation de l'IA

Les personnes interrogées n'ont que rarement mentionné la réalisation d'évaluations détaillées des discriminations potentielles lors de l'utilisation de l'IA, ce qui suggère un manque d'évaluations approfondies de ces discriminations dans la prise de décision automatisée.

L'obligation de respecter le principe de non-discrimination est consacrée à l'article 2 du TUE, à l'article 10 du traité sur le fonctionnement de l'Union européenne (TFUE) (qui impose à l'Union de combattre toute discrimination fondée sur un certain nombre de motifs) ainsi qu'aux articles 20 et 21 de la Charte (égalité en droit et non-discrimination fondée sur divers motifs). Des dispositions plus spécifiques et plus détaillées de plusieurs directives de l'UE consacrent également ce principe, avec des champs d'application variables.

L'automatisation et l'utilisation de l'IA permettent d'accroître considérablement l'efficacité des services et d'intensifier des tâches que les humains ne seraient pas en mesure d'entreprendre. Il convient cependant de s'assurer que les services et les décisions basés sur l'IA ne sont pas discriminatoires. À partir de ce constat, la Commission européenne a récemment souligné la nécessité d'une législation supplémentaire pour préserver la non-discrimination lors de l'utilisation de l'IA dans le **plan d'action de l'UE contre le racisme 2020-2025**.

La majorité des personnes interrogées sont, en principe, conscientes du risque de discrimination. Cependant, elles ont rarement soulevé cette question spontanément. Seules quelques-unes d'entre elles estiment que leur système pourrait effectivement être discriminatoire.

En outre, les personnes interrogées ont rarement mentionné des évaluations détaillées des discriminations potentielles, ce qui semble indiquer qu'il n'y a pas d'évaluation approfondie du risque de discrimination.

Il est courant de penser que l'omission d'informations sur certaines caractéristiques protégées, telles que le sexe, l'âge ou l'origine ethnique, permet de garantir l'absence de discrimination dans un système d'IA. Or cela n'est pas nécessairement vrai. En effet, les informations indiquant potentiellement des caractéristiques protégées (proxys), que l'on trouve souvent dans les ensembles de données, pourraient donner lieu à des discriminations.

Dans certains cas, les systèmes d'IA peuvent également être utilisés pour tester et détecter des comportements discriminatoires, qui peuvent être encodés dans des ensembles de données. Néanmoins, très peu de personnes interrogées ont fait mention de la possibilité de recueillir de telles informations sur des groupes défavorisés afin de déceler d'éventuelles discriminations. En l'absence d'une analyse approfondie des discriminations potentielles dans le cadre de l'utilisation réelle des systèmes d'IA, il n'existe pratiquement aucune discussion ni aucune analyse de l'effet positif potentiel de l'utilisation d'algorithmes sur la prise de décisions plus justes. Qui plus est, aucune des personnes interrogées travaillant sur l'IA n'a mentionné l'utilisation de l'IA pour détecter d'éventuelles discriminations comme un résultat positif, en ce sens que l'analyse des données pour déceler un éventuel biais permettrait de mieux détecter les discriminations.

Étant donné que la détection d'éventuelles discriminations par l'intermédiaire de l'IA et des algorithmes reste difficile et que les personnes interrogées n'ont que brièvement abordé cette question, différentes mesures doivent être prises pour remédier à ce problème. Parmi ces mesures figurent notamment l'obligation de prendre en considération les questions liées à la discrimination lors de l'évaluation de l'utilisation de l'IA, et l'investissement dans des études complémentaires sur les discriminations potentielles qui s'appuient sur un large éventail de méthodologies.

Ces méthodes pourraient impliquer, par exemple, des tests de discrimination. Elles pourraient en outre s'appuyer sur des méthodologies similaires visant à tester les préjugés à l'œuvre dans la vie quotidienne, par exemple en ce qui concerne les demandes d'emploi, où le nom du candidat est modifié pour identifier (indirectement) son origine ethnique. S'agissant des applications d'IA, ces tests pourraient impliquer la création éventuelle de faux profils pour les outils en ligne, qui ne diffèrent que par des attributs protégés. De cette façon, les résultats pourraient être vérifiés pour déceler d'éventuelles discriminations. La recherche pourrait également bénéficier d'une analyse statistique avancée afin de détecter les différences dans les ensembles de données portant sur des groupes protégés, et donc servir de base pour étudier les discriminations potentielles.

Enfin, certains entretiens menés dans le cadre de la recherche ont permis de constater que les résultats d'algorithmes d'apprentissage automatisé complexes sont souvent très difficiles à comprendre et à expliquer. En conséquence, d'autres recherches visant à mieux comprendre et expliquer ces résultats (l'IA dite «explicable») pourraient également contribuer à mieux détecter les discriminations lors de l'utilisation de l'IA.



AVIS 5 DE LA FRA

Le comité européen de la protection des données et le Contrôleur européen de la protection des données devraient envisager de fournir des orientations et un soutien supplémentaires pour mettre en œuvre efficacement les dispositions du RGPD qui s'appliquent directement à l'utilisation de l'IA pour la sauvegarde des droits fondamentaux, en particulier en ce qui concerne la signification des données à caractère personnel et leur utilisation dans l'IA, y compris dans les ensembles de données d'apprentissage de l'IA.

Il existe un niveau élevé d'incertitude concernant la signification de la prise de décision automatisée et le droit à une vérification par un être humain en lien avec l'utilisation de l'IA et la prise de décision automatisée. Partant, le comité européen de la protection des données et le Contrôleur européen de la protection des données devraient également envisager de clarifier davantage les notions de «prise de décision automatisée» et de «vérification par un être humain», lorsqu'elles sont mentionnées dans la législation européenne.

Par ailleurs, les organismes nationaux de protection des données devraient fournir des orientations pratiques sur la manière dont les dispositions relatives à la protection des données s'appliquent à l'utilisation de l'IA. Ces orientations pourraient inclure des recommandations et des listes de contrôle, basées sur des cas concrets d'utilisation de l'IA, afin de favoriser le respect des dispositions relatives à la protection des données.

Orientations supplémentaires sur la protection des données

Il est nécessaire de clarifier davantage le champ d'application et la signification des dispositions juridiques relatives à la prise de décision automatisée.

La protection des données est fondamentale dans le développement et l'utilisation de l'IA. L'article 8, paragraphe 1, de la Charte et l'article 16, paragraphe 1, du TFUE prévoient que toute personne a droit à la protection des données à caractère personnel la concernant. Le RGPD et la directive (UE) 2016/680 en matière de protection des données dans le domaine répressif développent ce droit et comportent de nombreuses dispositions applicables à l'utilisation de l'IA.

Les personnes interrogées ont indiqué que la plupart des systèmes d'IA qu'elles emploient utilisent des données à caractère personnel, ce qui signifie que la protection des données est affectée de nombreuses manières différentes. Cependant, d'après les personnes interrogées, certaines applications n'utilisent pas de données à caractère personnel, ou n'utilisent que des données anonymisées, de sorte que la loi sur la protection des données ne s'appliquerait pas. Si des données à caractère personnel sont utilisées, tous les principes et dispositions relatifs à la protection des données s'appliquent.

Le rapport met en lumière une question importante liée à la protection des données, qui concerne également d'autres droits fondamentaux au regard de la prise de décision automatisée. Selon une enquête Eurobaromètre, seuls 40 % des Européens savent qu'ils peuvent avoir un droit de regard en cas de décisions automatisées. La connaissance de ce droit est nettement plus élevée chez les personnes travaillant

avec l'IA. La majorité des personnes interrogées ont abordé cette question. Cependant, de nombreuses personnes interrogées, y compris des experts, ont fait valoir qu'il fallait clarifier le champ d'application et la signification des dispositions juridiques sur la prise de décision automatisée.

Dans le domaine des prestations sociales, les personnes interrogées n'ont mentionné qu'un seul exemple de décisions entièrement automatisées et fondées sur des règles. Toutes les autres applications qu'elles ont mentionnées sont vérifiées par des humains. Les personnes interrogées travaillant dans l'administration publique ont souligné l'importance de la vérification des décisions par des êtres humains. Pour autant, elles ont rarement expliqué ce qu'implique réellement une telle vérification par un être humain ou comment d'autres informations sont utilisées lors de la vérification des résultats des systèmes d'IA.

Si les personnes interrogées ne sont pas toutes d'accord en ce qui concerne la question de savoir si la législation existante est suffisante ou non, nombre d'entre elles ont appelé de leurs vœux une interprétation plus concrète des règles de protection des données existantes au regard de la prise de décision automatisée, telle que consacrée à l'article 22 du RGPD.

Accès effectif à la justice dans les cas impliquant des décisions fondées sur l'IA

Pour contester efficacement les décisions fondées sur le recours à l'IA, les personnes devraient être informées de l'utilisation de l'IA, et savoir où et comment déposer une réclamation. Les organisations qui utilisent l'IA doivent être en mesure d'expliquer leur système d'IA ainsi que les décisions fondées sur l'IA.

L'accès à la justice, qui est à la fois une procédure et un objectif, est essentiel pour les personnes qui cherchent à bénéficier d'autres droits procéduraux et matériels. Il englobe un certain nombre de droits fondamentaux, au premier rang desquels figure le droit à un procès équitable et à un recours effectif visé aux articles 6 et 13 de la CEDH et à l'article 47 de la charte des droits fondamentaux de l'UE. En conséquence, la notion d'accès à la justice oblige les États à garantir le droit de chaque individu à saisir un tribunal – ou, dans certaines circonstances, une instance alternative de résolution des litiges – pour obtenir un recours s'il s'avère que les droits de l'individu ont été enfreints.

Conformément à ces normes, toute victime d'une atteinte aux droits humains découlant du développement ou de l'utilisation d'un système d'IA par une entité publique ou privée doit avoir accès à une voie de recours devant une autorité nationale. Conformément à la jurisprudence applicable au titre de l'article 47 de la Charte et de l'article 13 de la CEDH, le recours doit être «effectif en pratique comme en droit».

Les conclusions de la recherche identifient les conditions préalables suivantes pour que le recours soit effectif, dans la pratique, dans les cas impliquant des systèmes d'IA et leur impact sur les droits fondamentaux: chacun doit être informé de l'utilisation de l'IA et savoir où et comment déposer une réclamation. Les organisations qui utilisent l'IA doivent veiller à ce que le public soit informé de leur système d'IA et des décisions qui en découlent.

Les conclusions montrent qu'il peut être difficile d'expliquer, en des termes simples, les systèmes d'IA et la façon dont ils prennent des décisions. Les droits de propriété intellectuelle peuvent entraver la communication d'informations détaillées sur le fonctionnement d'un algorithme. Qui plus est, certains systèmes d'IA sont complexes. Il est donc difficile de fournir des informations significatives sur le fonctionnement d'un système et sur les décisions afférentes.

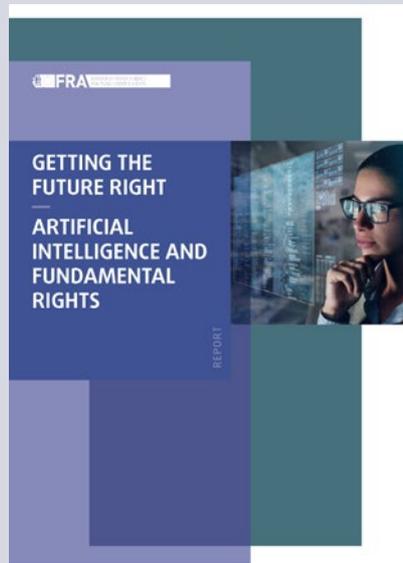
Pour remédier à ce problème et être en mesure d'expliquer les décisions, certaines entreprises interrogées préfèrent éviter d'utiliser des méthodes trop complexes pour certaines prises de décision. Dans le même esprit, elles utilisent aussi des méthodes d'analyse de données plus simples afin de mieux comprendre les principaux facteurs qui influencent certains résultats. Quelques-unes des personnes interrogées issues du secteur privé ont cité les efforts déployés pour améliorer progressivement leur compréhension de la technologie de l'IA.

AVIS 6 DE LA FRA

Le législateur européen et les États membres devraient garantir un accès effectif à la justice aux individus concernés dans les cas impliquant des décisions fondées sur l'IA.

Afin de garantir que les recours disponibles sont accessibles dans la pratique, le législateur européen et les États membres pourraient envisager d'introduire l'obligation légale, pour l'administration publique et les entreprises privées qui utilisent des systèmes d'IA, de fournir aux personnes cherchant à obtenir réparation des informations sur le fonctionnement de leurs systèmes d'IA, y compris des informations sur la manière dont ces systèmes d'IA prennent des décisions automatisées. Cette obligation permettrait de respecter le principe de l'égalité des armes dans les cas où des individus cherchent à obtenir justice. Elle favoriserait également l'efficacité de la surveillance externe et du contrôle des systèmes d'IA en matière de droits humains (voir l'avis 3 de la FRA).

Compte tenu de la difficulté d'expliquer les systèmes complexes d'IA, l'UE, conjointement avec les États membres, devrait envisager d'élaborer des lignes directrices afin d'encourager les efforts de transparence dans ce domaine. Ce faisant, ils devraient s'appuyer sur l'expertise des organismes nationaux de défense des droits humains et des organisations de la société civile actifs dans ce domaine.



Ce résumé présente les principales conclusions du rapport de la FRA intitulé *Bien préparer l'avenir: l'intelligence artificielle et les droits fondamentaux*. Le rapport principal peut être consulté sur le site web de la FRA.



PROMOUVOIR ET PROTÉGER VOS DROITS FONDAMENTAUX DANS TOUTE L'UE

Le rapport de la FRA consacré à l'intelligence artificielle et aux droits fondamentaux présente des exemples concrets de la manière dont les entreprises et les administrations publiques de l'UE utilisent ou tentent d'utiliser l'IA. Il aborde principalement quatre domaines essentiels: les prestations sociales, la police prédictive, les services de soins de santé et la publicité ciblée. Le rapport examine les implications potentielles pour les droits fondamentaux et analyse la manière dont ces droits sont pris en considération dans le cadre de l'utilisation ou du développement des applications d'IA.

Ce résumé présente les principales idées exposées dans le rapport. Ces idées peuvent orienter les efforts de l'UE et des États membres pour élaborer des politiques visant à réglementer l'utilisation des outils d'IA dans le respect des droits de l'homme et des droits fondamentaux.



FRA — AGENCE DES DROITS FONDAMENTAUX DE L'UNION EUROPÉENNE
Schwarzenbergplatz 11 — 1040 Vienne — Autriche
Tél. +43 158030-0 — Fax +43 158030-699

fra.europa.eu

 facebook.com/fundamentalrights
 twitter.com/EURightsAgency
 linkedin.com/company/eu-fundamental-rights-agency



Office des publications
de l'Union européenne