Parliamentary Assembly
Assemblée parlementaire
https://pace.coe.int

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Résolution 2343 (2020)[1]

Prévenir les discriminations résultant de l'utilisation de l'intelligence artificielle

Assemblée parlementaire

1. L'intelligence artificielle (IA) est en train de transformer nos modes de vie. Permettant une généralisation massive des processus, elle est déjà utilisée par un large éventail d'entités privées et publiques, dans des domaines aussi divers que des procédures de sélection déterminant l'accès à l'emploi et à l'éducation, l'évaluation des droits d'un individu à des prestations sociales ou à un crédit, ou le ciblage des messages publicitaires et d'information.

2. Bon nombre d'utilisations de l'IA peuvent avoir un impact direct sur l'égalité d'accès aux droits fondamentaux, notamment le droit à la vie privée et la protection des données à caractère personnel; l'accès à la justice et le droit à un procès équitable, en particulier en ce qui concerne la présomption d'innocence et la charge de la preuve; l'accès à l'emploi, à l'éducation, au logement et à la santé; et l'accès aux services publics et à la protection sociale. Il s'avère que l'utilisation de l'IA est susceptible de provoquer ou d'exacerber la discrimination dans ces domaines, entraînant des refus d'accès aux droits qui touchent de manière disproportionnée certains groupes — souvent les femmes, les personnes appartenant aux minorités, ainsi que les personnes déjà les plus vulnérables et les plus marginalisées. Son utilisation dans les flux d'informations a également été liée à la propagation de la haine en ligne qui a des retombées sur toutes les autres interactions sociales.

3. Le processus d'apprentissage par la machine utilisé pour produire des systèmes d'IA repose sur le recours à de vastes ensembles de données *(big data)* qui sont pour une grande part à caractère personnel. Des garanties effectives de la protection des données à caractère personnel revêtent donc une importance essentielle dans ce contexte. Parallèlement, les données sont biaisées par nature dans la mesure où elles reflètent les discriminations déjà présentes dans la société, ainsi que les préjugés des personnes chargées de leur collecte et de leur analyse. Le choix des données à utiliser et à ignorer dans un système fondé sur l'IA, ainsi que l'absence de données sur des questions clés, l'utilisation de variables indirectes et les difficultés inhérentes à la quantification de concepts abstraits peuvent également aboutir à des résultats discriminatoires. Les ensembles de données biaisées sont au cœur de nombreux cas de discrimination du fait de l'utilisation de l'IA et demeurent un problème majeur à résoudre dans ce domaine.

4. La conception et la finalité d'un système fondé sur l'IA sont également cruciales. Les algorithmes optimisés en vue de l'efficacité, de la rentabilité ou d'autres objectifs, sans tenir dûment compte de la nécessité de garantir l'égalité et la non-discrimination, peuvent entraîner une discrimination directe ou indirecte, y compris une discrimination par association, fondée sur toute une série de motifs, dont le sexe, le genre, l'âge, l'origine nationale ou ethnique, la couleur, la langue, les convictions religieuses, l'orientation sexuelle, l'identité de genre, les caractéristiques sexuelles, l'origine sociale, l'état civil, le handicap ou l'état de santé. Il est donc particulièrement important, lorsque l'utilisation de systèmes fondés sur l'IA risque d'avoir une incidence sur l'accès à des droits fondamentaux, que le plein respect de l'égalité et de la non-discrimination soit intégré d'emblée dans la conception de ces systèmes, et qu'ils soient rigoureusement testés avant d'être déployés, mais aussi de façon régulière après leur déploiement, afin de s'assurer que ces droits sont garantis.

---

1. *Texte adopté par la Commission permanente*, agissant au nom de l'Assemblée, le 22 octobre 2020 (voir Doc. 15151, rapport de la commission sur l'égalité et la non-discrimination, rapporteur: M. Christophe Lacroix).
Voir également la Recommandation 2183 (2020).

5. La complexité des systèmes d'IA et le fait qu'ils soient fréquemment développés par des entreprises privées et traités comme des éléments relevant de leur propriété intellectuelle peuvent conduire à de sérieux problèmes de transparence et de responsabilité concernant les décisions prises en utilisant ces systèmes. Ces caractéristiques peuvent rendre la discrimination extrêmement difficile à prouver et entraver l'accès à la justice, en particulier lorsque la charge de la preuve incombe à la victime et/ou lorsque la machine est supposée par défaut avoir pris la bonne décision, ce qui viole la présomption d'innocence.

6. Le manque de diversité au sein de nombreuses entreprises et professions technologiques augmente le risque que des systèmes d'IA soient développés sans tenir compte de leurs impacts potentiellement discriminatoires sur certains individus et groupes de la société. L'accès des femmes et des minorités aux professions scientifiques, technologiques, d'ingénierie et de mathématiques (STIM) doit être amélioré, et une véritable culture de respect de la diversité développée de toute urgence au sein de ces milieux professionnels. Le recours à des approches interdisciplinaires et interculturelles à tous les stades de la conception des systèmes d'IA contribuerait également à leur amélioration du point de vue du respect des principes d'égalité et de non-discrimination.

7. Enfin, des principes éthiques solides, clairs et universellement acceptés et applicables devraient sous-tendre le développement et le déploiement de tous les systèmes fondés sur l'IA. L'Assemblée parlementaire considère que ces principes peuvent être regroupés dans les grandes catégories suivantes: la transparence, y compris l'accessibilité et l'explicabilité; la justice et l'équité, y compris la non-discrimination; la capacité d'imputer la responsabilité des décisions à une personne, y compris l'engagement de la responsabilité de l'intéressé et l'existence de voies de recours; la sûreté et la sécurité; et le respect de la vie privée et la protection des données à caractère personnel.

8. L'Assemblée se félicite de ce que les acteurs publics et privés ont commencé à examiner et à élaborer des normes éthiques ainsi que des normes en matière de droits humains applicables à l'utilisation de l'IA. Elle se félicite en particulier de la Recommandation Rec/CM(2020)1 du Comité des Ministres sur les impacts des systèmes algorithmiques sur les droits de l'homme, complétée par ses lignes directrices, ainsi que de la recommandation de la Commissaire aux droits de l'homme du Conseil de l'Europe intitulée «Décoder l'intelligence artificielle: 10 mesures pour protéger les droits de l'homme». L'Assemblée approuve les propositions générales énoncées dans ces textes qui s'appliquent également dans le domaine de l'égalité et de la non-discrimination.

9. L'Assemblée souligne que le législateur ne doit pas invoquer la complexité de l'IA pour s'abstenir d'introduire des réglementations destinées à protéger et à promouvoir l'égalité et la non-discrimination dans ce domaine: les enjeux en matière de droits humains sont clairs et requièrent une action. Outre les principes éthiques, il faudrait élaborer des procédures, des outils et des méthodes de réglementation et d'audit des systèmes fondés sur l'IA afin de garantir leur conformité aux normes internationales en matière de droits humains, et notamment aux droits à l'égalité et à la non-discrimination. Compte tenu de la forte dimension transnationale et internationale des technologies fondées sur l'IA, des normes internationales apparaissent également nécessaires dans ce domaine.

10. Au vu de ces considérations, l'Assemblée invite les États membres:

10.1. à revoir leur législation antidiscrimination et à la modifier si nécessaire, de manière à ce qu'elle couvre tous les cas où une discrimination directe ou indirecte, y compris par association, pourrait être causée par l'utilisation de l'IA, et à ce que les plaignants aient pleinement accès à la justice; à cet égard, à veiller tout particulièrement à garantir la présomption d'innocence et à éviter d'imposer aux victimes de discrimination une charge de la preuve disproportionnée;

10.2. à élaborer une législation, des normes et des procédures nationales claires visant à garantir que tout système fondé sur l'IA respecte les droits à l'égalité et à la non-discrimination partout où son utilisation risquerait d'affecter la jouissance de ces droits;

10.3. à veiller à ce que les organismes de promotion de l'égalité soient pleinement habilités à traiter les problèmes concernant les principes d'égalité et de non-discrimination résultant d'une utilisation de l'IA et à soutenir les personnes qui portent plainte dans ce domaine, et à s'assurer qu'ils disposent de toutes les ressources nécessaires pour s'acquitter de ces tâches.

11. Afin que l'utilisation par les autorités publiques de technologies fondées sur l'IA soit soumise à un contrôle parlementaire et à un examen public adéquats, l'Assemblée invite les parlements nationaux:

11.1. à veiller à ce que l'utilisation de ces technologies fasse l'objet de débats parlementaires réguliers, et à ce qu'un cadre propice à ces débats soit mis en place;

11.2. à exiger du gouvernement d'informer à l'avance le parlement de l'utilisation de ces technologies;

11.3. à rendre obligatoire l'inscription systématique dans un registre public de chaque utilisation de ces technologies par les autorités.

12. Afin d'aborder les questions sous-jacentes de la diversité et de l'inclusion dans le domaine de l'IA, l'Assemblée invite en outre les États membres:

12.1. à promouvoir l'intégration des femmes, des jeunes filles et des personnes appartenant aux minorités dans les filières d'enseignement des STIM dès le plus jeune âge et jusqu'aux plus hauts niveaux, ainsi qu'à collaborer avec le secteur industriel pour faire en sorte que la diversité et l'intégration soient favorisées tout au long des parcours professionnels;

12.2. à soutenir la recherche sur les biais concernant les données et les moyens de contrer efficacement leur impact dans les systèmes fondés sur l'IA;

12.3. à promouvoir la culture numérique et l'accès aux outils numériques par tous les membres de la société.

13. L'Assemblée invite toutes les entités, tant publiques que privées, travaillant sur des systèmes fondés sur l'IA, et avec ces derniers, à veiller à ce que le respect de l'égalité et de la non-discrimination soit intégré dès le début dans la conception des systèmes en cause et testé de manière adéquate avant leur déploiement, dès lors que ces systèmes pourraient avoir un impact sur l'exercice ou l'accès à des droits fondamentaux. À cette fin, elle invite ces entités à envisager de renforcer les capacités afin de mettre en place un cadre d'évaluation de l'impact sur les droits humains pour le développement et le déploiement de systèmes fondés sur l'IA, tant par les entités privées que par les entités publiques. Elle encourage en outre le recours à des équipes interdisciplinaires et diversifiées à tous les stades du développement et du déploiement des systèmes fondés sur l'IA.

14. Enfin, l'Assemblée invite les parlements nationaux à soutenir les travaux menés au niveau international, notamment par le biais du Comité ad hoc du Conseil de l'Europe sur l'intelligence artificielle (CAHAI), à veiller à ce que les normes internationales relatives aux droits humains soient effectivement appliquées dans le domaine de l'IA, et à garantir le respect des principes d'égalité et de non-discrimination dans ce domaine.