



Assemblée générale

Distr. générale
13 mai 2026
Français
Original : anglais

Conseil des droits de l'homme

Soixante-deuxième session

15 juin-10 juillet 2026

Points 3 et 5 de l'ordre du jour

**Promotion et protection de tous les droits de l'homme,
civils, politiques, économiques, sociaux et culturels,
y compris le droit au développement**

Organismes et mécanismes de protection des droits de l'homme

Effets des systèmes d'intelligence artificielle sur la bonne gouvernance

Étude du Comité consultatif du Conseil des droits de l'homme*

Résumé

Dans la présente étude, soumise conformément à la résolution [57/5](#) du Conseil des droits de l'homme, le Comité consultatif du Conseil examine les effets des systèmes d'intelligence artificielle (IA) sur la bonne gouvernance.

Le Comité consultatif détermine comment l'intelligence artificielle peut améliorer la prestation des services publics, accroître la transparence, renforcer les mesures de lutte contre la corruption et favoriser une participation inclusive. Il examine les risques en matière de droits de l'homme liés à l'utilisation de l'IA dans la gouvernance publique, notamment la surveillance à grande échelle, la discrimination algorithmique, l'opacité de la prise de décisions automatisée, la manipulation de l'environnement informationnel, le déplacement de la main-d'œuvre et les dommages environnementaux. Il se penche en outre sur la question de savoir en quoi la concentration du développement de l'IA entre les mains d'un petit nombre d'acteurs risque de renforcer les rapports de dépendance numérique, et souligne qu'il est impératif d'assurer la participation inclusive et équitable de tous les États à l'élaboration des normes et règles applicables.

Le Comité consultatif formule des recommandations tendant à garantir que l'utilisation de l'IA dans la gouvernance publique reste conforme aux obligations internationales en matière de droits de l'homme et aux principes de bonne gouvernance.

* Il a été convenu que le présent document serait publié après la date normale de publication en raison de circonstances indépendantes de la volonté du soumetteur.



I. Introduction

A. Mandat

1. Le Conseil des droits de l'homme, dans sa résolution 57/5, a prié son Comité consultatif de réaliser une étude sur les effets des systèmes d'intelligence artificielle (IA) sur la bonne gouvernance, dans laquelle il ferait ressortir, en particulier, les domaines dans lesquels les systèmes d'IA peuvent contribuer à la promotion et à la protection des droits de l'homme grâce à une bonne gouvernance et ceux dans lesquels ces systèmes constituent une menace pour la bonne gouvernance et les droits de l'homme, mettrait en évidence les bonnes pratiques appliquées dans le monde pour développer, mettre en service, utiliser et administrer les systèmes d'IA de manière à promouvoir et protéger les droits de l'homme grâce à une bonne gouvernance, et recenserait les garanties nécessaires.

2. À sa trente-deuxième session, le Comité consultatif a créé un groupe de rédaction, composé des membres dont le nom suit : Miznah Omair Alomair, Nicole Ameline, Joseph Gérard Angoh, Edward Amoako Asante, Buhm-Suk Baek (Rapporteur), Rabah Boudache, Ignacio de Casas, Riva Ganguly Das, Alessandra Devulsky, Veronika Fikfak, Jewel Major, Afsaneh Nadipour, Ren Yisheng, Patrycja Sasnal, Vassilis Tzevelekos (Président) et Frans Viljoen.

B. Portée de l'étude

3. La présente étude examine l'intégration de l'IA dans la gouvernance publique et ses implications pour les droits de l'homme. L'accent n'est donc pas mis sur la gouvernance de l'IA en tant que telle, mais plutôt sur la manière dont les systèmes d'IA peuvent favoriser ou entraver le respect des droits de l'homme lorsqu'ils sont mis en œuvre par les pouvoirs publics. L'étude examine les conditions qui doivent être réunies pour que l'IA aide la gouvernance à tenir ses promesses en matière d'efficacité et de transparence. Elle met en évidence le fait que l'utilisation de l'IA, si elle ne s'appuie pas sur les obligations en matière de droits de l'homme, risque d'ouvrir la voie à de nouvelles violations systématiques des droits de l'homme qui échappent aux mécanismes de recours classiques.

4. L'analyse retrace l'évolution du cadre établi par le Conseil des droits de l'homme reliant bonne gouvernance et droits de l'homme, depuis la résolution 7/11 (2008) du Conseil jusqu'à sa résolution 57/5 (2024), et met en évidence le fait qu'il est de plus en plus admis que ces deux domaines ne sont pas simplement complémentaires, mais se constituent mutuellement. Elle examine la manière dont l'IA met à l'épreuve la capacité d'adaptation du cadre établi. Si l'IA et les systèmes algorithmiques peuvent aider les États à être mieux à même de s'acquitter des obligations mises à leur charge par le droit international des droits de l'homme, notamment en améliorant la prestation des soins de santé, en élargissant l'accès à l'éducation et en renforçant les mécanismes de protection sociale, ils leur donnent, dans le même temps, les moyens d'instaurer des dispositifs de surveillance sans précédent et risquent de perpétuer des pratiques discriminatoires. De plus, le remplacement de procédures administratives transparentes par des processus informatiques opaques risque de nuire à l'efficacité du contrôle exercé par les autorités administratives et judiciaires, et de porter atteinte au droit à un recours utile.

5. La présente étude s'appuie sur les débats menés sous l'égide de l'ONU et sur les contributions des parties prenantes, notamment les 32 réponses au questionnaire du Comité consultatif émanant d'États, d'organisations régionales, d'institutions nationales des droits de l'homme, d'organisations de la société civile et d'établissements universitaires¹. L'analyse établit un cadre global fondé sur les droits de l'homme reposant sur les principes que sont le droit à un recours utile et l'obligation de rendre des comptes, et met systématiquement en évidence la nécessité impérieuse d'intégrer le devoir de diligence en matière de droits de l'homme à l'acquisition, à la conception et à l'utilisation de l'IA au sein de l'administration

¹ À consulter à l'adresse suivante : <https://www.ohchr.org/en/hr-bodies/hrc/advisory-committee/impact-artificial-intelligence-systems-good-governance> (en anglais).

publique, en garantissant une transparence algorithmique qui offre véritablement aux personnes concernées la possibilité de comprendre et de contester les décisions automatisées, et en mettant en place des garanties institutionnelles solides permettant de détecter les biais discriminatoires, les atteintes à la vie privée et les irrégularités de procédure, et d'y remédier. Elle examine les incidences disproportionnées de l'IA sur les populations vulnérables et la question de savoir en quoi l'absence de mesures proactives visant à garantir une conception inclusive et une participation réelle risque de creuser les inégalités existantes. Globalement, l'étude s'interroge sur la question de savoir comment exploiter les possibilités de transformation qu'offre l'IA au service d'une bonne gouvernance centrée sur les droits de l'homme, la dignité humaine et le principe de responsabilité.

C. Bonne gouvernance et droits de l'homme

6. Selon le Haut-Commissariat des Nations Unies aux droits de l'homme (HCDH), la bonne gouvernance est le processus par lequel les institutions publiques conduisent des affaires publiques, gèrent des ressources publiques et garantissent la réalisation des droits de l'homme sans abus ni corruption². Cette conception s'est imposée lorsque la Commission des droits de l'homme, dans sa résolution 2000/64, a officiellement reconnu le lien entre gouvernance, développement et dignité humaine, en affirmant qu'une gouvernance transparente et responsable, attentive aux besoins de la population, était essentielle à la promotion des droits de l'homme. Par conséquent, comme l'a affirmé le Conseil des droits de l'homme, la bonne gouvernance se caractérise par la participation, la responsabilité, la transparence, la réactivité et l'efficacité, et se fonde sur l'état de droit et les principes relatifs aux droits de l'homme³.

7. Depuis 2008, le Conseil des droits de l'homme a progressivement mis en place un cadre normatif reliant bonne gouvernance et droits de l'homme, en reconnaissant qu'un gouvernement transparent, responsable et participatif constitue le fondement même de la promotion des droits de l'homme, en mettant l'accent sur la lutte contre la corruption et en appelant à la mise en place d'une fonction publique professionnelle et fondée sur le mérite, qui respecte les normes d'intégrité les plus strictes⁴. Plus récemment, le Conseil s'est intéressé aux défis de l'ère numérique, reconnaissant qu'avoir accès à Internet et aux technologies numériques était essentiel pour pouvoir accéder à l'information, aux services publics, à l'éducation et à la justice, et saluant les innovations technologiques visant à prévenir, repérer et combattre la corruption, innovations qui devaient néanmoins s'accompagner de garanties suffisantes en matière de protection des données et de respect de la vie privée⁵.

8. La résolution 57/5 du Conseil des droits de l'homme a marqué un tournant en faisant officiellement de l'IA une question relevant de la bonne gouvernance et des droits de l'homme. Le Conseil a constaté que l'IA présentait à la fois des possibilités et des risques, puisqu'elle permettait notamment de lutter contre la corruption et d'améliorer la prestation des services publics, mais posait aussi des difficultés – parmi lesquelles la désinformation – susceptibles de nuire aux processus démocratiques. Le Conseil a exhorté les États, la société civile, les médias et les autres parties prenantes à élaborer des stratégies visant à atténuer les risques tout en promouvant la pleine jouissance des droits de l'homme et des libertés fondamentales. Il a surtout demandé aux États de promouvoir de manière inclusive et équitable des systèmes d'IA sûrs, sécurisés et dignes de confiance au bénéfice de tous, en vertu du principe selon lequel le développement de l'IA doit être guidé par les principes relatifs à la bonne gouvernance et aux droits de l'homme, notamment la transparence, l'équité et la responsabilité. Parallèlement, le Secrétaire général a créé, en 2023, l'Organe consultatif de haut niveau sur l'intelligence artificielle⁶. Le Pacte numérique mondial met en évidence

² Voir <https://www.ohchr.org/fr/good-governance/about-good-governance>.

³ Voir, par exemple, les résolutions 7/11, 19/20, 25/8, 31/14, 37/6, 45/9, 51/5 et 57/5 du Conseil des droits de l'homme.

⁴ Voir les résolutions 7/11, 19/20 et 25/8 du Conseil des droits de l'homme.

⁵ Voir les résolutions 45/9 et 51/5 du Conseil des droits de l'homme.

⁶ Gouverner l'IA au bénéfice de l'humanité : rapport final (New York, Nations Unies, 2024).

la nécessité de faire en sorte d'urgence que l'IA renforce, plutôt qu'elle ne compromette, les droits de l'homme, l'égalité et l'état de droit.

II. La bonne gouvernance assistée par l'intelligence artificielle : possibilités et risques pour les droits de l'homme

A. Les contributions positives de l'intelligence artificielle aux droits de l'homme et à la bonne gouvernance

1. Amélioration de la prestation des services publics et réalisation des droits économiques, sociaux et culturels

9. Il est communément admis que, parmi les atouts qu'elle peut offrir en matière de gouvernance, l'IA peut notamment transformer la prestation des services publics et, partant, promouvoir la réalisation de tous les droits, y compris les droits économiques, sociaux et culturels. Les systèmes d'IA permettent de traiter d'importants volumes de données et d'automatiser des tâches complexes, ce qui peut permettre aux services administratifs de gagner en efficacité, et d'être plus accessibles et mieux adaptés aux besoins de chacun. Plusieurs États et observateurs internationaux ont noté, avec un optimisme prudent, que, lorsqu'elle est correctement conçue et mise en œuvre, et dûment réglementée, l'IA peut renforcer considérablement les capacités des pouvoirs publics dans des domaines tels que la santé, l'éducation et la protection sociale⁷.

10. Les innovations fondées sur l'IA dans les domaines du diagnostic médical et de la gestion de la santé publique peuvent contribuer à faire progresser le droit à la santé⁸, ce qui illustre la capacité de l'IA d'améliorer l'accessibilité et la qualité des services de santé, obligation fondamentale découlant du droit au meilleur état de santé possible⁹. L'IA peut améliorer la précision des diagnostics, optimiser le fonctionnement des hôpitaux, rationaliser le triage des patients, et rendre ainsi les systèmes de santé plus efficaces et plus équitables¹⁰.

11. Dans le domaine de l'éducation, l'IA pourrait contribuer à réduire les inégalités et à améliorer la gestion des établissements¹¹. Les plateformes d'enseignement basées sur l'IA peuvent contribuer à combler les inégalités en adaptant l'enseignement en temps réel aux styles et aux besoins d'apprentissage de chacun, permettant ainsi une meilleure intégration des personnes en difficulté dans les systèmes standardisés, quels que soient l'âge, la langue ou le handicap de celles-ci¹². Les applications permettant de convertir des supports pédagogiques en d'autres formats sont des outils d'aide aux personnes en situation de handicap. L'IA peut également aider les services administratifs des établissements d'enseignement à repérer les élèves susceptibles d'abandonner leurs études et à mettre en place des interventions ciblées, tout en permettant aux enseignants d'être plus attentifs aux besoins individuels de chaque élève. Lorsqu'ils s'inscrivent dans le cadre d'une approche centrée sur l'humain qui donne la priorité à l'accessibilité dans des conditions d'équité, ces outils ont le potentiel de réduire les disparités en matière de résultats scolaires et de promouvoir le principe de bonne gouvernance que constitue une prestation de services inclusive, axée sur l'être humain. Cependant, l'adoption sans discernement de l'IA risque de creuser les fractures numériques et de réduire la capacité d'action des individus, compromettant ainsi, à terme, le droit à une éducation inclusive et équitable¹³.

⁷ Contributions du Brésil, de Singapour, de l'Ouzbékistan et de Solidarité des peuples pour la démocratie participative.

⁸ Contributions du Chili, de Cuba et de l'Équateur.

⁹ Contributions de la Commission des droits de l'homme de Mexico.

¹⁰ Contributions du Chili et de l'Équateur.

¹¹ Voir [A/79/520](#).

¹² Contributions du Venezuela (République bolivarienne du) et de la Commission nationale des droits de l'homme du Myanmar.

¹³ Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO), *Rapport mondial de suivi sur l'éducation, 2023 : les technologies dans l'éducation : qui est aux commandes ?* (Paris, 2023).

12. L'IA peut améliorer la gestion des programmes de protection sociale et contribuer ainsi à garantir le droit à la sécurité sociale et à un niveau de vie suffisant¹⁴. Les algorithmes d'apprentissage automatique permettent d'analyser rapidement les données des ménages afin de repérer ceux qui ont le plus besoin d'aide ou de détecter des erreurs dans l'octroi des prestations. L'automatisation de la traduction et de l'interprétation peut faciliter l'accès à la justice et aux services publics. Lorsqu'elle est transparente, cette automatisation contribue à garantir que des ressources limitées parviennent rapidement aux bénéficiaires auxquels elles sont destinées, ce qui facilite l'exercice effectif des droits socioéconomiques.

13. Le traitement automatisé des demandes peut améliorer l'efficacité et la réactivité des institutions, deux éléments essentiels à une bonne gouvernance. L'intégration de l'IA dans les procédures administratives, telles que la déclaration fiscale et le calcul des prestations, peut permettre de gagner en efficacité et permet aux agents publics de se concentrer sur des tâches plus complexes exigeant une appréciation humaine¹⁵.

14. Plus globalement, les capacités d'analyse prédictive de l'IA peuvent faciliter le passage d'une gouvernance réactive à une gouvernance proactive. Par exemple, l'intégration de l'IA dans la gestion des catastrophes peut aider à prévoir des crises telles que des inondations ou des épidémies, ce qui permet aux autorités de prendre des mesures préventives¹⁶.

15. La technologie de l'IA pourrait être utilisée à l'appui de certaines fonctions dans les domaines du maintien de l'ordre, de la sécurité nationale et des contrôles frontaliers. Les systèmes fondés sur l'IA peuvent faciliter le déroulement des enquêtes pénales, par exemple grâce à la reconnaissance des plaques d'immatriculation, et contribuer aux efforts de détection des activités illicites. Il convient toutefois de souligner que le recours à l'IA dans ces domaines soulève des préoccupations éminemment importantes en matière de droits de l'homme. Ces technologies exigent que l'on instaure des cadres juridiques solides afin de garantir la transparence, l'observance du principe de responsabilité et le plein respect des normes internationales relatives aux droits de l'homme¹⁷.

16. Dans la pratique, ces applications de l'IA contribuent à une bonne gouvernance en facilitant l'élaboration de politiques fondées sur des données factuelles. Grâce aux technologies d'IA, les pouvoirs publics sont mieux à même de prendre des décisions rapides et éclairées, et d'allouer les ressources en fonction des besoins anticipés, ce qui caractérise le principe même de réactivité, qui est au cœur d'une gouvernance axée sur l'être humain. Si les modèles prédictifs doivent être conçus, mis en œuvre et encadrés dans le respect du droit à la vie privée et du principe de non-discrimination et compte tenu d'autres considérations relatives aux droits de l'homme, ils recèlent un potentiel considérable pour ce qui est de transformer l'administration publique de manière à favoriser la réalisation des droits économiques, sociaux et culturels.

2. Promotion de la transparence, du principe de responsabilité et de la lutte contre la corruption

17. L'IA peut être un instrument utile pour accroître la transparence dans l'administration publique et, partant, donner confiance au public et renforcer la responsabilité des institutions publiques. En limitant la marge d'appréciation individuelle dans la prise de décisions courantes et en générant des traces numériques vérifiables, des systèmes d'IA bien conçus peuvent réduire les risques de corruption et faciliter la détection des actes répréhensibles¹⁸. Les innovations en matière d'administration numérique peuvent contribuer à la prévention de la corruption et à la promotion du principe de responsabilité, à condition que les données

¹⁴ Contribution de l'Espagne.

¹⁵ Banque mondiale, *Artificial Intelligence in the Public Sector: Maximizing Opportunities, Managing Risks* (Washington, 2020).

¹⁶ Contributions de Cuba et de la Plateforme de Genève sur les droits humains.

¹⁷ Voir [A/HRC/60/45](#).

¹⁸ Business at OECD, « Harnessing AI for integrity: opportunities, challenges, and the business case against corruption » (Organisation de coopération et de développement économiques (OCDE), Paris, 2024).

personnelles soient protégées et que la vie privée des intéressés soit dûment respectée¹⁹. Plus généralement, les nouvelles technologies numériques, y compris l'intelligence artificielle, peuvent faciliter la coopération internationale au service de la prévention de la corruption et de la lutte contre ce phénomène, pour autant que leur conception et leur utilisation restent conformes aux normes applicables du droit international des droits de l'homme²⁰.

18. L'IA permet également de mieux garantir le respect du principe de responsabilité en améliorant la conservation et l'analyse des données. Les systèmes d'IA du secteur public enregistrent généralement chaque action effectuée ainsi que les données utilisées pour étayer les décisions adoptées, ce qui crée une piste d'audit intégré. Cela peut décourager les comportements répréhensibles en garantissant la traçabilité des décisions et permet d'avoir une base de référence en cas de recours ou de plainte. Dans certains pays, des initiatives visent désormais à concevoir des systèmes d'IA transparents, capables d'expliquer leur raisonnement algorithmique, ce qui permet de faire en sorte que la gouvernance algorithmique vienne compléter l'état de droit plutôt que le compromettre²¹.

19. L'IA peut renforcer encore la participation citoyenne et le contrôle indépendant, et permet, ce faisant, de mieux garantir le respect du principe de responsabilité. Les politiques de données ouvertes, lorsqu'elles sont associées à des analyses fondées sur l'intelligence artificielle, permettent à la société civile, aux médias et à d'autres parties prenantes de contrôler de vastes quantités d'informations publiques afin de repérer d'éventuels cas de malversations ou de violations de droits. Le Programme des Nations Unies pour le développement a souligné qu'il était utile d'utiliser l'IA à cette fin, notamment pour surveiller les marchés publics et les flux financiers illicites, dans le cadre d'une gouvernance plus large²².

20. Si l'IA n'est pas une solution complète aux défis de la gouvernance, elle offre néanmoins de nouveaux moyens de promouvoir le respect des principes de transparence et de responsabilité. Grâce à ces applications, les pouvoirs publics peuvent prendre des mesures efficaces pour mieux garantir l'intégrité des institutions, lutter contre la corruption et faire respecter l'état de droit.

3. Un outil au service de l'inclusion, de la participation et de l'égalité

21. Entre autres aspects positifs, on peut également citer le fait que l'IA peut favoriser l'inclusion et l'égalité dans la gouvernance. Instaurer une bonne gouvernance suppose de se doter d'institutions inclusives, qui garantissent la participation et l'accès de l'ensemble des composantes de la société, y compris des groupes marginalisés. L'IA peut aider à surmonter les obstacles auxquels ces groupes sont confrontés. Par exemple, les technologies d'assistance basées sur l'intelligence artificielle permettent de veiller à ce que les personnes handicapées puissent accéder aux informations et aux services publics au même titre que les autres, conformément à l'obligation, énoncée dans la Convention relative aux droits des personnes handicapées, de garantir l'accessibilité des informations et des services. Ces outils permettent en outre de mieux garantir le droit d'accès à l'information des groupes de population qui ont des difficultés à communiquer avec les pouvoirs publics en raison de barrières liées à la langue et à la lecture ou à l'écriture.

22. L'IA peut favoriser la participation aux processus démocratiques. Ces applications peuvent élargir le champ de la démocratie participative en créant de nouveaux canaux par lesquels les individus, y compris les groupes de tout temps sous-représentés dans l'élaboration des politiques, peuvent contribuer à la gouvernance. Le Conseil des droits de l'homme a encouragé ce type d'innovations, soulignant que les technologies numériques pouvaient favoriser une participation inclusive à la vie publique et demandant aux États de combler la fracture numérique afin d'assurer la participation des populations marginalisées dans des conditions d'égalité²³. De plus, dans les États plurilingues, les outils de traduction

¹⁹ Voir la résolution 51/5 du Conseil des droits de l'homme.

²⁰ Conférence des États Parties à la Convention des Nations Unies contre la corruption, résolution 11/1.

²¹ Voir <https://e-estonia.com/estonia-and-automated-decision-making-challenges-for-public-administration>.

²² New Technologies for Sustainable Development: Perspectives on Integrity, Trust and Anti-Corruption (New York, 2021).

²³ Résolution 47/23 du Conseil des droits de l'homme.

et de traitement du langage basés sur l'IA peuvent faciliter la participation des personnes appartenant à des minorités linguistiques à la vie politique, notamment l'accès de ces personnes aux processus électoraux. Pour assurer une réelle participation de ces personnes grâce à ces outils, il faut toutefois garantir un accès équitable à ceux-ci.

23. L'IA peut également aider à repérer et à corriger les inégalités dans la prestation des services. En analysant les données administratives et autres jeux de données pertinents, les pouvoirs publics peuvent repérer les domaines dans lesquels les services font systématiquement défaut à certains groupes ou dans certaines zones, notamment les quartiers où l'accès aux soins de santé est systématiquement plus limité ou ceux où les délais d'intervention des services d'urgence sont plus longs, ce qui contribue à la promotion des principes d'égalité et de non-discrimination²⁴. Lorsqu'on prend de telles mesures de manière transparente et en tenant compte des contributions des collectivités, ces applications de l'IA sont en adéquation avec le droit au développement et l'engagement commun de ne pas faire de laissés-pour-compte²⁵. Les collectivités concernées devraient participer activement à la conception et à la gouvernance de l'IA, et non être de simples bénéficiaires de ces politiques. Ces avantages ne sont toutefois en aucun cas garantis d'avance et exigent que les systèmes d'IA soient eux-mêmes conçus et entraînés de manière à éviter la reproduction des biais existants.

B. Conséquences négatives de l'intelligence artificielle pour les droits de l'homme et la bonne gouvernance

24. En dépit du potentiel de l'IA, l'intégration de cet outil dans la gouvernance publique pose des difficultés notables pour ce qui est de la protection des droits de l'homme. En l'absence de cadres juridiques solides, de garanties en matière de droits de l'homme et d'un contrôle humain, les systèmes d'IA risquent de faciliter la surveillance de masse, de renforcer les biais systémiques, de porter atteinte à la régularité des procédures légales et de fausser le débat public. Si rien n'est fait pour y remédier, ces risques peuvent ébranler les fondements de la bonne gouvernance et de l'état de droit. Qualifier l'IA de technologie neutre reviendrait à simplifier à l'extrême la situation, car la conception et l'utilisation de cette technologie peuvent refléter et accentuer les déséquilibres de pouvoir existants.

1. L'érosion de la vie privée due à la surveillance de masse et à l'exploitation des données à grande échelle

25. Sur le plan des droits de l'homme, la préoccupation la plus immédiate que soulève l'intégration de l'IA à la gouvernance concerne l'incidence de cet outil sur le respect de la vie privée²⁶. Les systèmes d'IA reposent généralement sur la collecte et le traitement de vastes volumes de données personnelles, qu'il s'agisse d'images faciales, de données de localisation, de contenus publiés sur les réseaux sociaux ou encore d'identifiants biométriques ; lorsque ces systèmes sont utilisés par les pouvoirs publics, ils permettent à ceux-ci de développer considérablement leurs dispositifs de surveillance. Les techniques d'apprentissage automatique sont particulièrement utiles pour dégager des tendances au sein de vastes jeux de données en apparence sans rapport les uns avec les autres et obtenir ainsi avec ces données ce que l'on appelle un « effet de mosaïque » : grâce à ce phénomène, il est possible d'établir des profils complets d'individus en regroupant des informations provenant de sources disparates²⁷. Ces systèmes se caractérisent essentiellement par leur capacité à extraire, par un processus de déduction, des informations personnelles sensibles que les individus concernés n'ont pas expressément divulguées. Dans plusieurs pays, les instances judiciaires ont estimé que certains systèmes de profilage de masse portaient atteinte aux droits fondamentaux en permettant une surveillance secrète et excessive de communautés vulnérables, sans dispositifs de transparence adéquats et sans justifications suffisantes,

²⁴ Voir <https://stories.undp.org/faire-evoluer-l-ia-vers-plus-d-equite>.

²⁵ Organe consultatif de haut niveau sur l'intelligence artificielle en 2023, *Gouverner l'IA au bénéfice de l'humanité*.

²⁶ Voir A/HRC/60/45.

²⁷ Voir A/HRC/48/31 ; David E. Pozen, « The mosaic theory, national security, and the Freedom of Information Act », *Yale Law Journal*, vol. 115, n° 3 (décembre 2005).

soulignant ainsi la nécessité d'instaurer des garanties strictes afin que le profilage automatisé ne porte pas atteinte au droit à la vie privée ni ne perpétue la discrimination.

26. Ces évolutions soulèvent de sérieuses préoccupations concernant le droit à la vie privée, ainsi que le rôle de ce dernier en tant que droit ouvrant la voie à l'exercice des libertés d'expression, d'association et de religion ou de conviction. La surveillance assistée par l'IA risque de porter atteinte à la sphère privée, qui doit être protégée de toute immixtion arbitraire, dans la mesure où la reconnaissance des formes et l'exploration de données à grande échelle peuvent donner lieu à une surveillance généralisée qui outrepassse les limites traditionnellement dictées par les principes de nécessité et de proportionnalité²⁸. La vidéosurveillance en réseau, associée à la technologie de reconnaissance faciale, par exemple, permet de suivre en temps réel les déplacements des personnes dans les espaces publics sans intervention humaine constante ; elle est particulièrement intrusive et nécessite donc une justification très rigoureuse au regard de la protection de la vie privée et des données.

27. Les risques liés à l'IA pour ce qui est de la protection de la vie privée ne se limitent pas aux caméras. Les autorités ont de plus en plus souvent recours à l'apprentissage automatique pour analyser des jeux de données combinés, tels que les données fiscales, celles relatives à la protection sociale et les informations publiées en ligne, afin de repérer les personnes considérées comme présentant un risque. Or, il a été conclu à l'issue d'enquêtes administratives que le recoupement automatisé à grande échelle de données personnelles pouvait donner lieu à des décisions erronées ayant des répercussions considérables, ce qui montre que l'automatisation peut porter atteinte à la vie privée et aux garanties de procédure²⁹.

28. Plus généralement, la surveillance assistée par l'IA marque un tournant qualitatif : on passe d'une interception ciblée fondée sur des soupçons individuels à un filtrage algorithmique à grande échelle, susceptible de soumettre des populations entières à une surveillance constante. Le HCDH a averti que la surveillance sans distinction restreignait l'exercice des libertés d'expression et d'association³⁰. L'effet dissuasif de la surveillance généralisée, qui empêche les individus d'exercer leurs droits légitimes en raison d'un risque perçu de sanctions ou de conséquences négatives, a été amplement attesté, notamment pour ce qui est de la liberté d'expression³¹. Lorsque des individus savent qu'ils peuvent être surveillés, ils sont moins disposés à s'exprimer librement ou à se rassembler. L'analyse prédictive peut accentuer ces préoccupations en concentrant la surveillance dans certains quartiers ou certaines collectivités sur la foi d'indicateurs indirects relatifs à l'origine ethnique ou à la situation socioéconomique, ce qui renforce des pratiques discriminatoires tout en portant atteinte à la vie privée.

29. Les groupes marginalisés et vulnérables sont particulièrement exposés à ces risques. Aux frontières, les réfugiés et les migrants font de plus en plus l'objet d'une collecte de données biométriques et d'estimations algorithmiques des risques³². Ces technologies peuvent conduire à des discriminations et porter atteinte à la vie privée et aux garanties de procédure. Les défenseurs des droits de l'homme, les journalistes et les militants politiques sont également exposés lorsque des dispositifs de surveillance à grande échelle sont mis en place. Dans ce contexte, la protection de la vie privée devient une préoccupation collective. Lorsque la surveillance est omniprésente, l'espace démocratique ouvert à la contestation se rétrécit pour tous. Le Conseil des droits de l'homme a demandé que toute mesure de surveillance numérique réponde aux critères de légalité, de nécessité et de proportionnalité³³ ; or, de nombreuses pratiques fondées sur l'intelligence artificielle reposent sur des cadres juridiques dépassés, qui ne tiennent pas compte des capacités actuelles de ces technologies.

²⁸ Voir [A/HRC/41/35](#).

²⁹ Communication de l'Institute for Digital Rights.

³⁰ Voir [A/HRC/27/37](#).

³¹ Anita R. Gohdes, *Repression in the Digital Age: Surveillance, Censorship, and the Dynamics of State Violence* (Oxford, Oxford University Press, 2024).

³² Voir [A/75/590](#).

³³ Résolution 48/4 du Conseil des droits de l'homme.

30. Dans l'ensemble, face à l'IA, il est désormais d'autant plus urgent de protéger le droit à la vie privée³⁴. Les caractéristiques de ces systèmes qui les rendent utiles pour la gouvernance sont précisément ceux-là mêmes qui facilitent une collecte et un traitement toujours plus étendus des données personnelles, et, partant, portent directement atteinte au droit à la vie privée et entravent l'exercice d'autres droits fondamentaux. En l'absence de garanties suffisantes, l'automatisation de la surveillance risque de se transformer en automatisation de la censure et du contrôle social, ce qui est fondamentalement incompatible avec une bonne gouvernance démocratique.

2. Les biais algorithmiques et la discrimination : persistance des inégalités

31. Parmi les principales préoccupations soulevées par l'IA, la suivante concerne les biais algorithmiques – à savoir le fait que les systèmes d'IA reproduisent ou amplifient les préjugés sociétaux – et les implications de ces biais pour le droit à l'égalité et à la non-discrimination. Garantir une bonne gouvernance suppose d'assurer une prestation des services publics et une prise de décisions impartiales, et de traiter tous les individus équitablement, sans distinction fondée sur la race, le sexe, la religion, le handicap ou tout autre motif de discrimination interdit. Or, les systèmes d'IA ne sont pas neutres par nature et reflètent souvent des schémas discriminatoires ancrés dans les données d'apprentissage ou les choix de conception. Le recours à de tels systèmes dans le cadre de la gouvernance peut donc donner lieu à une discrimination systémique qui peut s'avérer plus difficile à déceler et à contester par les voies de recours existantes que la discrimination résultant des décisions humaines classiques³⁵.

32. Des éléments de biais ont été observés dans les systèmes d'IA utilisés dans le secteur public dans de nombreux domaines. Dans le domaine de la justice pénale, les résultats inégaux obtenus grâce aux outils d'évaluation des risques utilisés pour étayer les décisions relatives à la liberté provisoire, à la détermination de la peine et à la libération conditionnelle ont soulevé des préoccupations. Des problèmes du même ordre ont été constatés dans les systèmes de sélection des candidats à l'emploi, qui peuvent reproduire certains préjugés liés au genre ou au handicap, ainsi que dans les algorithmes financiers ou administratifs, dont les indicateurs peuvent avoir pour effet de pénaliser les interruptions d'activité professionnelle ou de cibler certaines catégories de contribuables. Dans le domaine de la gouvernance, l'utilisation de modèles entraînés à partir de données reflétant des disparités socioéconomiques historiques peut amener à cibler de façon disproportionnée, à des fins de contrôle, les entreprises appartenant à des minorités ou à des communautés à faibles revenus³⁶. Dans leurs contributions à la présente étude, les parties prenantes ont systématiquement mis en avant le risque que l'IA aggrave la discrimination et les inégalités en matière d'accessibilité ; de nombreux exemples montrent en effet que, loin d'être réduites, les inégalités préexistantes, qu'elles soient liées à l'origine ethnique, à la classe sociale, au sexe, au genre, au handicap ou au statut migratoire, sont au contraire amplifiées. Cela va à l'encontre de l'obligation qui incombe à l'État de prévenir les pratiques discriminatoires et de promouvoir l'égalité réelle³⁷.

33. L'une des principales difficultés réside dans le manque de transparence. De nombreux systèmes d'IA ne parviennent pas à exposer leur raisonnement de manière intelligible, et la notion d'intention n'a pas de sens pour une machine, si bien qu'il peut être difficile pour les personnes concernées de reconnaître l'existence même d'un biais, et a fortiori d'en attribuer la responsabilité ou d'obtenir réparation. Les cadres traditionnels de lutte contre la discrimination, qui visent souvent à démontrer l'existence d'une discrimination intentionnelle, ne sont pas adaptés aux préjudices complexes qui peuvent résulter de l'exploitation de données. En conséquence, de nombreuses analyses juridiques et relatives aux droits de l'homme préconisent une approche axée sur les conséquences concrètes, en particulier sur les disparités observées³⁸. S'il est constaté qu'un groupe protégé a obtenu un

³⁴ Voir [A/HRC/48/31](#). Voir aussi [A/HRC/61/68](#), par. 25 et 26.

³⁵ Voir [A/HRC/48/31](#) ; [A/HRC/54/70](#), par. 24.

³⁶ Contribution de la Plateforme de Genève sur les droits humains.

³⁷ Contributions de l'Espagne, de l'Irlande, du Mexique, de Singapour, de l'Association pour la défense des droits de l'homme et des revendications démocratiques/culturelles du peuple azerbaïdjanais, d'Iran – ARC, du Korean Progressive Network – Jinbonet et de Transparency International.

³⁸ Voir [A/HRC/56/68](#).

nombre statistiquement significatif d'issues défavorables, ce constat doit donner lieu à une analyse approfondie et fait naître l'obligation de modifier ou de supprimer le système en question, quelle que soit l'intention de départ. Cette approche est en adéquation avec les interprétations largement acceptées de la notion de discrimination indirecte. Dans l'ensemble, les biais algorithmiques peuvent constituer une menace sérieuse pour le principe de non-discrimination, car ils risquent de renforcer les inégalités structurelles sous le couvert de l'objectivité technologique.

3. Le problème de la « boîte noire » : atteinte à la transparence, aux garanties de procédure et au droit à un recours

34. L'utilisation de l'IA dans le cadre de la gouvernance pose également une autre difficulté tout aussi considérable, qui concerne l'opacité des décisions prises par les algorithmes, souvent désignée sous le nom de « problème de la boîte noire ». De nombreux systèmes d'apprentissage automatique et réseaux neuronaux ne fournissent pas de raisonnement compréhensible pour les personnes concernées. Dans le domaine de la gouvernance, cette opacité peut être contraire aux principes de transparence et de responsabilité, et au principe de l'état de droit, et mettre en péril les garanties de procédure (le droit de comprendre et de contester les décisions ayant une incidence sur des droits et obligations, le droit à un procès équitable et le droit à un recours utile). À l'ère numérique, l'exercice des droits suppose une utilisation transparente et responsable des systèmes automatisés, car un manque d'explicabilité peut entraver l'accès à la justice³⁹. La transparence est essentielle au respect du principe de responsabilité, l'explicabilité, au respect des garanties de procédure et l'établissement clair des responsabilités, à l'utilité des recours⁴⁰.

35. Les juridictions nationales et internationales doivent actuellement faire face à ces questions. Les cadres régionaux régissant la protection des données reconnaissent le droit de n'être pas exclusivement soumis à des décisions automatisées ayant des conséquences importantes et, lorsque de telles décisions sont prises, le droit d'obtenir des informations utiles sur le raisonnement qui y a présidé⁴¹. La question de savoir ce qui constitue une explication utile pour des modèles complexes reste controversée. Tandis que d'aucuns préconisent des justifications simplifiées ou des mesures de transparence, d'autres soulignent que, dans le cas des systèmes d'apprentissage profond, les explications a posteriori peuvent s'avérer trop techniques ou purement déductives⁴². Malgré ces débats, la tendance semble indiquer que l'explicabilité est appelée à devenir une garantie fondamentale.

36. L'opacité entraîne également un déficit de responsabilité. Lorsqu'un système d'IA cause un préjudice, la responsabilité peut être partagée entre les développeurs, les fournisseurs et les pouvoirs publics qui acquièrent et utilisent ces systèmes. Cette incertitude porte atteinte au droit à un recours utile, consacré par l'article 2 (par. 3) du Pacte international relatif aux droits civils et politiques, dans la mesure où les personnes concernées peuvent difficilement déterminer la partie responsable à laquelle demander des comptes. Cela soulève des questions de responsabilité, qui ont trait notamment à l'imputation des violations des obligations en matière de droits de l'homme et à la répartition des responsabilités entre les acteurs concernés, tels que les autorités publiques qui utilisent les systèmes d'IA et les entreprises qui les conçoivent et les fournissent. Le problème est d'autant plus épineux lorsque des fonctions publiques sont, dans les faits, confiées à des systèmes propriétaires protégés par le secret commercial ou par des droits de propriété intellectuelle, ce qui empêche tant les agents publics que les particuliers d'avoir une vision claire des fondements des décisions⁴³. Dans le secteur public, cette opacité peut porter atteinte aux principes établis de la responsabilité administrative et être contraire à l'obligation qui incombe aux autorités publiques de motiver les décisions ayant une incidence sur les droits individuels. La Haute-Commissaire des Nations Unies aux droits de l'homme a souligné que les entreprises

³⁹ Voir A/HRC/48/31.

⁴⁰ Résolution 54/18 du Conseil des droits de l'homme ; contribution de Solidarité des peuples pour la démocratie participative.

⁴¹ Règlement (UE) 2016/679 (règlement général sur la protection des données), art. 22.

⁴² OCDE, Principes de l'IA. À consulter à l'adresse suivante : <https://www.oecd.org/fr/themes/principes-de-l-ia.html>.

⁴³ Contribution du Korean Progressive Network – Jinbonet.

devaient faire preuve de transparence quant à la conception et au fonctionnement des technologies de surveillance et autres technologies à haut risque fournies aux gouvernements⁴⁴. Le Groupe de travail sur la question des droits de l'homme et des sociétés transnationales et autres entreprises ainsi que le HCDH ont souligné que l'IA devait être explicable, que l'on devait réaliser des études d'impact algorithmique et qu'il fallait intégrer le devoir de précaution en matière de droits de l'homme tout au long du cycle de vie de l'IA⁴⁵.

37. Compte tenu de ces enjeux, il est impératif de prendre à la fois des mesures systémiques ex ante en matière de prévention et, parallèlement, des mesures correctives ex post. Entre autres mesures indispensables, il convient de réaliser à titre obligatoire des études d'impact sur les droits de l'homme, d'effectuer régulièrement des audits, de tenir des registres publics des systèmes de prise de décisions automatisée et de se doter d'autorités de contrôle indépendantes. Cette approche préventive repose sur l'idée généralement admise que les mécanismes correctifs classiques risquent d'intervenir trop tard dans des environnements où tout va très vite et où les données jouent un rôle central⁴⁶. Parallèlement, les Principes de l'IA établis par l'OCDE confèrent à la transparence, à l'explicabilité et au contrôle humain une place fondamentale et préconisent la mise en place de mécanismes permettant aux personnes lésées de contester les résultats produits (principe 1.3).

4. Des menaces visant la liberté d'expression, l'accès à l'information et la participation démocratique

38. L'incidence de l'intelligence artificielle sur la liberté d'expression et l'intégrité de la sphère publique appelle une attention urgente, car la bonne gouvernance présuppose un public informé ainsi qu'un environnement pluraliste propice à l'échange d'idées et à des processus politiques équitables et transparents. Entre autres problèmes, il convient de s'intéresser avant tout à la prolifération de la désinformation générée par l'IA, qui menace l'accès à des informations fiables par la production, à moindre coût et à très grande échelle, de textes, d'images et de vidéos réalistes, mais faux ou trompeurs⁴⁷. Par exemple, à l'approche des élections dans divers pays, des observateurs ont mis en évidence l'existence de réseaux coordonnés et de vidéos « deepfakes » visant à discréditer des candidats ou à faire baisser le taux de participation électorale⁴⁸. Cette situation se complique encore lorsque la désinformation se propage au moyen d'applications de messagerie cryptée, ce qui rend difficile la diffusion d'informations fiables provenant d'autres sources⁴⁹. L'utilisation malveillante de ces technologies entraîne une érosion de la confiance dans l'ensemble de l'environnement informatique. Dès lors qu'une personne prend conscience qu'il est possible de produire des contenus mensongers qui paraissent convaincants, même des faits avérés risquent de lui sembler potentiellement faux, ce qui nuit à l'élaboration de politiques fondées sur des données probantes et érode la confiance du public dans les communications des pouvoirs publics. De plus, les acteurs politiques et autres parties intéressées peuvent analyser des données personnelles afin de diffuser des contenus personnalisés et persuasifs, destinés à influencer le comportement électoral ou à décourager la participation. De telles pratiques, en particulier lorsqu'elles reposent sur un profilage intrusif, peuvent porter atteinte au droit à la vie privée et entraver la capacité de l'individu de se forger ses propres opinions, élément essentiel de la liberté de pensée et d'expression et condition préalable à une participation véritable aux affaires publiques⁵⁰. En outre, les algorithmes de recommandation des plateformes utilisant l'IA qui amplifient systématiquement des propos biaisés afin de faire réagir un maximum de personnes peuvent, à la longue, façonner les convictions des utilisateurs à leur insu, ce qui soulève des inquiétudes quant à l'autonomie mentale de ces

⁴⁴ Voir [A/HRC/48/31](#).

⁴⁵ Voir [A/HRC/59/32](#) et [A/HRC/59/53](#).

⁴⁶ UNESCO, « Recommandation sur l'éthique de l'intelligence artificielle » (Paris, 2022) ; Matthew U. Scherer, « Regulating Artificial intelligence systems: risks, challenges, competencies, and strategies », *Harvard Journal of Law & Technology*, vol. 29, n° 2 (printemps 2016).

⁴⁷ Contribution de l'American University.

⁴⁸ Freedom House, « Freedom on the Net 2019: the crisis of social media » (Washington, 2019).

⁴⁹ [A/HRC/61/68](#), par. 74.

⁵⁰ Voir [A/73/348](#) et [A/HRC/47/25](#). Voir également Comité des droits de l'homme, observation générale n° 34 (2011), par. 9 et 10.

derniers. Ces dynamiques sapent la gouvernance en fragmentant un socle factuel commun, en favorisant les clivages et en empêchant la tenue de débats politiques raisonnés.

39. Certains préjudices ont une incidence non seulement sur la liberté d'expression, mais aussi sur l'égalité et la sécurité de la personne. Le Conseil des droits de l'homme s'est déclaré préoccupé par le fait que la manipulation de l'information, notamment au moyen des technologies numériques, risque d'attiser la violence et les conflits, et a demandé aux États d'encourager les acteurs privés à lutter contre la désinformation tout en assurant le respect des droits de l'homme⁵¹. L'IA peut également amplifier les contenus haineux ou discriminatoires en générant de nombreuses variantes qui échappent aux filtres de base, ce qui accroît les risques pour la dignité, la sécurité et l'égalité des communautés visées.

40. Il convient de se pencher, en outre, sur la question de la censure et de la modération des contenus. Les États peuvent recourir à l'IA pour surveiller et restreindre les contenus en ligne à grande échelle. Si la suppression de contenus illicites peut être justifiée, les filtres automatisés ont tendance à supprimer des contenus de manière excessive et ne tiennent souvent pas compte du contexte, de la satire ou du caractère d'utilité publique, et risque ainsi d'entraver indûment l'exercice légitime du droit à la liberté d'expression et de créer un effet dissuasif. Toute restriction de la liberté d'expression, y compris les mesures adoptées pour lutter contre la désinformation, doit se conformer strictement aux prescriptions de l'article 19 (par. 3) du Pacte international relatif aux droits civils et politiques, et respecter les principes de légalité, de nécessité et de proportionnalité. Les mesures visant à lutter contre la désinformation ne doivent pas servir de prétexte à une restriction injustifiée de l'exercice légitime du droit à la liberté d'expression⁵².

41. Ces phénomènes, cumulés, sapent la confiance dans l'environnement informationnel. Lorsque les individus ne peuvent plus se fier à la véracité de ce qu'ils lisent ou voient, ni à l'authenticité de leurs interactions en ligne, la cohésion sociale s'effrite, et un pilier essentiel de la bonne gouvernance est ainsi ébranlé. De plus, cette méfiance a tendance à s'étendre à toutes les formes d'information, ce qui pose des difficultés de taille pour les médias d'information, les organismes de recherche et les organismes publics qui s'appuient sur la communication numérique. Pour surmonter ces difficultés, il faut adopter des approches fondées sur la coopération entre les différentes parties prenantes⁵³. Les États, les plateformes privées et la société civile doivent collaborer pour mettre en place des mesures de protection adéquates. Il s'agit notamment de garantir la transparence de la publicité politique et l'étiquetage des contenus générés par l'IA, d'intégrer des dispositifs appropriés permettant de vérifier l'authenticité humaine, d'investir dans l'éducation aux médias et la littératie numérique afin de renforcer la résilience de la société⁵⁴, et de prendre des mesures pour promouvoir l'équité et le pluralisme dans les systèmes de recommandation. Toute restriction de la liberté d'expression imposée à ces fins doit respecter les normes internationales en matière de légalité, de nécessité et de proportionnalité.

42. Les répercussions négatives que l'IA peut avoir sur la liberté d'expression et la participation démocratique sont vastes et touchent l'ensemble du système. Elles portent atteinte à la capacité du public de mener un débat éclairé et de demander des comptes aux autorités, et sapent ainsi les fondements mêmes de la gouvernance démocratique. Ces conséquences ne sont toutefois pas inéluctables et il est possible de les éviter si les différentes parties prenantes agissent de concert et sans relâche pour préserver la capacité d'action des individus, le pluralisme et la confiance, dans le respect des droits de l'homme et de l'état de droit. Entre autres mesures, il est essentiel d'instaurer des garanties adéquates permettant d'attester que les utilisateurs sont bien des êtres humains afin de préserver la confiance dans les espaces publics numériques, et de s'assurer que la participation aux processus démocratiques reflète une contribution humaine authentique⁵⁵. De telles mesures peuvent contribuer à préserver l'authenticité de la participation citoyenne tout en atténuant les risques

⁵¹ Résolution 55/10 du Conseil des droits de l'homme. Voir aussi A/HRC/61/68, par. 70.

⁵² A/HRC/61/68, par. 21.

⁵³ Voir A/HRC/47/52.

⁵⁴ Résolution 57/5 du Conseil des droits de l'homme.

⁵⁵ Bryan Ford, « Identity and personhood in digital democracy: evaluating inclusion, equality, security, and privacy in pseudonym parties and other proofs of personhood » (2020).

liés à la désinformation, aux comportements inauthentiques coordonnés et à l'influence indûment exercée sur les processus décisionnels.

5. L'automatisation et le droit au travail : effets de déplacement et creusement des inégalités

43. L'automatisation permise par l'IA peut poser des difficultés particulières pour ce qui est de l'exercice du droit au travail consacré par l'article 23 de la Déclaration universelle des droits de l'homme et l'article 6 du Pacte international relatif aux droits économiques, sociaux et culturels. Si les progrès technologiques ont de tout temps entraîné le remplacement de certaines formes d'emploi tout en en créant d'autres, la rapidité et l'ampleur du déploiement actuel de l'IA sont sans précédent et dépassent le cadre du travail manuel répétitif pour s'étendre aux tâches administratives, créatives et décisionnelles. Les effets des déplacements d'emploi imputables à l'automatisation fondée sur l'IA sont inégalement répartis. Les secteurs fortement exposés au risque d'automatisation emploient souvent un nombre disproportionné de personnes issues de groupes marginalisés. Les femmes, qui sont souvent surreprésentées aux postes de bureau et d'assistance administrative, sont nettement plus exposées à la perte d'emplois liée à l'automatisation⁵⁶. De plus, les répercussions ne se limitent pas aux économies avancées. Les pays en développement qui misent traditionnellement sur les industries de main-d'œuvre pour assurer leur développement économique sont exposés à des risques particuliers, car l'automatisation risque d'amoindrir les avantages liés au coût de la main-d'œuvre sur lesquels reposent ces stratégies de développement. Selon les analyses réalisées par des organismes des Nations Unies, une part importante des emplois dans les pays en développement pourrait être menacée par l'automatisation, ce qui soulève des préoccupations quant à la pauvreté, à la stabilité économique et au creusement des inégalités mondiales⁵⁷.

44. Au-delà des conséquences économiques, les licenciements massifs liés à l'IA portent atteinte à la dignité humaine et nuisent à la cohésion sociale. Pour la plupart des gens, le travail est non seulement une source de revenus, mais aussi un élément central de leur identité, de leur raison d'être et de leur participation à la vie sociale. Le licenciement brutal de personnes qui souhaitent pourtant travailler peut avoir un coût psychologique et social important, en ce que cela entraîne notamment un sentiment d'aliénation et une perte d'estime de soi. Cette évolution mérite des interventions proactives de la part des organes dirigeants. Pour veiller à ce que nul ne soit laissé pour compte, il faut mener une action concertée visant à garantir le droit au travail, y compris l'accès à des perspectives de travail décent, élément fondamental de la bonne gouvernance à l'ère de l'IA.

6. La consommation d'énergie et les droits environnementaux : menaces pour le développement durable

45. Les systèmes d'IA utilisés par les pouvoirs publics ont une empreinte matérielle importante, dont il n'est souvent pas tenu compte, et cela a des répercussions sur le droit à un environnement propre, sain et durable⁵⁸. Contrairement à l'idée reçue selon laquelle l'IA serait une technologie purement incorporelle, son infrastructure entraîne une consommation importante de ressources. L'augmentation exponentielle des besoins de puissance de calcul pour l'entraînement de modèles de grande ampleur et l'analyse de données en temps réel a considérablement accru la consommation d'énergie. Selon les prévisions, les besoins énergétiques des centres de données devraient augmenter fortement d'ici à 2030, si bien que l'on risque de devoir avoir davantage recours aux combustibles fossiles pour pouvoir répondre aux besoins de base en matière d'électricité⁵⁹. Une telle évolution risque de

⁵⁶ Paweł Gmyrek *et al.*, « Generative AI and jobs: a global analysis of potential effects on job quantity and quality », document de travail n° 96 de l'OIT (Genève, 2023).

⁵⁷ World Social Report 2020: Inequality in a Rapidly Changing World (publication des Nations Unies, 2020).

⁵⁸ Contributions de l'Arabie saoudite, du Chili, du Maroc, de l'Association pour la défense des droits de l'homme et des revendications démocratiques/culturelles du peuple azerbaïdjanais, d'Iran – ARC, de l'Institute for Digital Rights, du Center for Advanced Study in Humanitarian Law et de la Plateforme de Genève sur les droits humains.

⁵⁹ Agence internationale de l'énergie, Energy and AI: World Energy Outlook Special Report (2025).

compromettre le respect des engagements mondiaux en matière de climat et contribue à la pollution atmosphérique à l'échelon local, ce qui fait peser des risques sur la santé publique et les écosystèmes.

46. Au-delà des émissions de carbone, la gestion thermique des infrastructures d'IA exerce une pression considérable sur les ressources en eau douce⁶⁰. L'importante consommation d'eau nécessaire au refroidissement des centres de données entre souvent en concurrence avec les besoins des collectivités locales, en particulier dans les régions exposées au stress hydrique. Dans ce contexte, l'opacité qui entoure les données environnementales constitue un défi majeur pour la bonne gouvernance. Les grands fournisseurs de technologies omettent souvent de communiquer des données précises sur les prélèvements et la consommation d'eau, ce qui empêche les pouvoirs publics et la société civile de demander des comptes aux acteurs concernés et porte atteinte au droit d'accès à l'information, indispensable à une gouvernance participative efficace en matière de gestion des ressources.

47. De plus, le cycle de vie de l'IA contribue à perpétuer les schémas d'injustice environnementale et d'inégalités mondiales. Le matériel nécessaire au fonctionnement des systèmes d'IA alimente la demande de minerais critiques, souvent extraits dans des pays en développement dans des conditions susceptibles de porter atteinte aux droits des travailleurs et aux normes environnementales. L'obsolescence rapide de ce type de matériel accélère la production de déchets électroniques, dont les conséquences sont assumées de manière disproportionnée par les pays en développement. Cette dynamique traduit une forme de fracture numérique : les avantages économiques découlant de l'IA se concentrent dans les pays développés, tandis que les externalités environnementales et la dégradation des ressources sont assumées par les pays en développement, ce qui a pour effet de creuser les inégalités mondiales. Les États qui souhaitent mettre l'IA au service de la bonne gouvernance doivent donc prendre en considération ces conséquences environnementales et les effets sur la répartition des coûts et des bénéfices, qui font partie intégrante de leurs obligations en matière de droits de l'homme.

7. L'IA dans le système judiciaire et les lieux de détention : garantir l'équité et la dignité humaine

48. Dans le domaine de l'administration de la justice, les inquiétudes liées au recours à l'IA sont particulièrement vives. L'IA est de plus en plus utilisée au sein des systèmes judiciaires pour l'exécution de tâches administratives, notamment la gestion des dossiers, la traduction, l'anonymisation automatique des décisions de justice et le dépôt électronique d'actes. Elle est en outre de plus en plus souvent intégrée aux moteurs de recherche juridiques et utilisée pour assister les juges et les avocats dans la rédaction de résumés d'affaires ou l'analyse de précédents⁶¹. Les tribunaux et les autorités judiciaires ont également commencé à utiliser, à titre expérimental, des outils d'évaluation des risques fondés sur l'IA afin d'éclairer la prise de décisions en matière de mise en liberté sous caution, de libération conditionnelle et de détermination de la peine⁶². Si ces applications peuvent permettre de gagner en efficacité devant une charge de travail importante, elles soulèvent toutefois la question fondamentale de savoir dans quelle mesure des fonctions judiciaires essentielles peuvent être confiées à des systèmes algorithmiques. L'indépendance de la magistrature risque d'être compromise lorsque le pouvoir exécutif ou législatif, et non le pouvoir judiciaire lui-même, est chargé de sélectionner les données d'apprentissage, de concevoir les algorithmes ou de superviser les systèmes d'IA utilisés dans les tribunaux, en particulier lorsque ces fonctions sont confiées à des entreprises du secteur privé⁶³. Par nature, l'équité d'un procès repose sur l'appréciation humaine et le respect du principe de responsabilité. Si l'IA peut faciliter le déroulement des procédures judiciaires, les juges jouent un rôle irremplaçable en exerçant leur pouvoir d'appréciation et en tenant compte des particularités

⁶⁰ Joyeeta Gupta, Hilmer Bosch et Luc van Vliet, « AI's excessive water consumption threatens to drown out its environmental contributions », *The Conversation*, 21 mars 2024.

⁶¹ Voir [A/79/296](#).

⁶² Conseil de l'Europe, Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires (Strasbourg, 2019).

⁶³ Voir [A/79/296](#) ; UNESCO, *Manuel de formation mondial : l'IA et l'état de droit pour le pouvoir judiciaire* (Paris, 2024).

de chaque affaire. Les responsables judiciaires ont donc souligné qu'en matière de justice, la décision devait appartenir, en dernier ressort, à l'humain et non à la machine⁶⁴. Dans tous ces cas, un contrôle humain rigoureux, la transparence et la possibilité pour les parties de comprendre et de contester le rôle de l'IA sont essentiels. L'opacité des systèmes d'IA peut porter atteinte au droit à la liberté et à la sécurité de la personne ainsi qu'au droit à un procès équitable, garantis par les articles 9 et 14 du Pacte international relatif aux droits civils et politiques, lorsque les mis en cause ne sont pas en mesure de contester les décisions dont ils font l'objet. Conscients de ces risques, certains pays ont commencé à classer parmi les systèmes à haut risque les systèmes d'IA destinés à être utilisés par les autorités judiciaires et à interdire les estimations des risques fondées sur l'IA lorsque celles-ci reposent exclusivement sur le profilage ou sur les traits de personnalité pour prédire la commission d'infractions⁶⁵. En conséquence, l'intégration de l'IA dans les procédures judiciaires doit être considérée avec la prudence qui s'impose, l'objectif étant de préserver l'indépendance de la magistrature, l'équité des procédures et la confiance du public dans l'administration de la justice.

49. Dans les établissements pénitentiaires, la mise en place de systèmes basés sur l'IA a soulevé de nouvelles préoccupations en matière de droits de l'homme. Certaines installations de construction récente, souvent qualifiées de « prisons intelligentes », ont recours à la reconnaissance faciale, à la surveillance biométrique et à l'analyse prédictive pour renforcer la sécurité et améliorer l'efficacité opérationnelle. Si ces outils peuvent contribuer à réduire la violence et le trafic de marchandises de contrebande, ils risquent toutefois de nuire à la dimension humaine du milieu carcéral en limitant les contacts interpersonnels et l'exercice, par les professionnels de ce secteur, de leur pouvoir discrétionnaire. La surveillance algorithmique constante à laquelle elles sont soumises peut accentuer le sentiment d'isolement des personnes privées de liberté, et un recours excessif à l'automatisation dans la gestion de la détention risque de compromettre les efforts de réinsertion en supprimant la dimension humaine indispensable à la prise en charge et à l'évaluation individualisée des détenus. L'Ensemble de règles minima des Nations Unies pour le traitement des détenus (Règles Nelson Mandela) met en évidence l'importance d'un contact humain réel, en précisant qu'un isolement prolongé sans contact humain constitue une forme de traitement cruel, inhumain ou dégradant (règles 43 et 44). Par conséquent, l'automatisation dans les lieux de détention, si elle n'est pas maîtrisée, peut être incompatible avec l'obligation, au regard du droit international des droits de l'homme, de traiter toutes les personnes privées de liberté avec humanité, en respectant la dignité intrinsèque de l'être humain.

III. Développer une intelligence artificielle respectueuse des droits au service de la bonne gouvernance

A. Biais des données et déséquilibre mondial dans le développement de l'intelligence artificielle : défis dans le domaine des droits de l'homme

50. L'une des principales difficultés réside dans les biais présents dans les données et les algorithmes, problème d'autant plus grave que le développement de l'IA est concentré entre les mains d'un petit nombre d'acteurs, dans un éventail limité de contextes. Les systèmes d'IA sont bien souvent l'expression de valeurs, d'hypothèses et de limites inhérentes à leur conception et aux données à partir desquelles ils sont entraînés. Le secteur de l'IA et la recherche dans ce domaine restent dominés par un petit nombre d'États dotés de technologies perfectionnées et par une poignée de grandes entreprises, de sorte que ces systèmes traduisent souvent un éventail restreint de points de vue. Il est à craindre, de ce fait, que les technologies d'IA ne tiennent pas compte de la diversité des contextes culturels, sociaux et régionaux, ou ne représentent pas comme il se doit ces différents contextes.

⁶⁴ UNESCO, « Recommandation sur l'éthique de l'intelligence artificielle ». Voir également UNESCO, « Lignes directrices pour l'utilisation des systèmes d'IA dans les tribunaux » (Paris, 2025).

⁶⁵ Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle.

51. Ce phénomène a été qualifié de forme de colonialisme numérique, par lequel la position dominante d'une poignée de grands fournisseurs de technologies et leurs cadres normatifs perpétuent des schémas historiques d'exclusion⁶⁶. Les pays où le secteur de l'IA en est encore à ses balbutiements adoptent souvent des systèmes d'IA ou des cadres réglementaires élaborés à l'étranger, principalement conçus pour les économies technologiquement avancées. Les responsables locaux et les populations concernées interviennent parfois très peu dans la mise en œuvre de ces solutions. Les déséquilibres d'influence au cours des processus de mise en œuvre risquent de creuser les inégalités. Une approche uniforme de l'IA peut avoir pour effet de reléguer au second plan les systèmes de connaissances locaux, en particulier ceux des peuples autochtones, lorsque la technologie n'est pas adaptée au contexte culturel et social. En pareil cas, les cadres d'IA homogènes importés de sources extérieures risquent de marginaliser les valeurs et le savoir-faire locaux⁶⁷. Les biais des données et des algorithmes ont des répercussions directes et importantes sur les droits de l'homme. Les algorithmes biaisés ou discriminatoires peuvent produire des résultats inéquitables et injustes, ce qui constitue une violation effective des principes d'égalité et de non-discrimination. Si rien n'est fait, ces systèmes risquent de renforcer ce que l'on a qualifié d'« injustice automatisée », phénomène par lequel des décisions algorithmiques opaques privent des individus d'un traitement équitable, ou de l'accès à des services ou à des possibilités, sur la base de données erronées ou biaisées. Ces préjudices risquent de toucher tout particulièrement les personnes appartenant à des groupes marginalisés et vulnérables, ce qui ne ferait que creuser encore les inégalités.

52. Pour surmonter ces difficultés, il faut intégrer les principes des droits de l'homme, l'équité et la diversité à toutes les étapes du développement de l'IA⁶⁸. De simples ajustements techniques ne suffisent pas : il faut opérer des changements plus profonds qui concernent la diversité des personnes qui développent les systèmes d'IA et les finalités pour lesquelles ces systèmes sont conçus. Il est essentiel de garantir la diversité des participants et d'assurer une conception inclusive tout au long du cycle de vie de l'IA, depuis la collecte des données jusqu'à l'utilisation des systèmes, en passant par la conception des algorithmes, afin de réduire les biais et de promouvoir le respect du principe de responsabilité. En tenant compte d'un éventail plus large de points de vue, les systèmes d'IA et les dispositifs de gouvernance dans lesquels ils s'inscrivent pourront représenter plus fidèlement la diversité des besoins et des valeurs de la société.

53. Des initiatives encourageantes de démocratisation de l'IA voient le jour au plan local. Des spécialistes des données de différentes régions, notamment d'Afrique et d'Asie, collaborent pour créer des jeux de données et des outils adaptés aux langues peu représentées et aux besoins locaux. Ces initiatives visent à rendre l'IA plus inclusive et à mieux l'adapter aux différentes cultures, pour faire contrepoids aux approches uniformes qui ne tiennent pas compte de la diversité des contextes. Les initiatives internationales en matière de normalisation commencent à tenir compte de ces préoccupations⁶⁹. En 2021, l'UNESCO a demandé aux États de veiller à ce que le développement de l'IA respecte les droits de l'homme, valorise la diversité et garantisse l'équité⁷⁰. Appliquer ces principes suppose de remettre en question les choix de conception par défaut et d'associer des experts locaux au processus de développement. Les systèmes d'IA devraient être conçus de manière à respecter la pluralité des valeurs juridiques et culturelles, plutôt qu'à imposer, de façon non intentionnelle, des normes extérieures. De plus, le cadre international des droits de l'homme doit lui-même être appliqué à l'IA dans le respect des particularités culturelles. Certains spécialistes soulignent que le fait de se fonder sur le droit international des droits de l'homme sans tenir dûment compte des contextes locaux peut traduire une orientation implicitement individualiste qui ne rend pas pleinement compte des conceptions collectives ou communautaires des droits qui existent dans certains contextes culturels. En conséquence, les normes internationales relatives à la protection des données ou à la transparence

⁶⁶ Jake Okechukwu Effoduh et Miracle Okumu Mudeyi, « TWAILing AI governance: will third world countries be left behind? », TWAIL Review (7 août 2025).

⁶⁷ Contribution de Solidarité des peuples pour la démocratie participative.

⁶⁸ Voir [A/HRC/59/32](#).

⁶⁹ Voir [A/HRC/53/42](#).

⁷⁰ « Recommandation sur l'éthique de l'intelligence artificielle ».

algorithmique devraient être interprétées et appliquées de manière à respecter les droits fondamentaux de l'homme, tout en tenant compte des différences culturelles et juridiques pertinentes dans les mécanismes de responsabilité au plan local.

B. Assurer une participation et une inclusion véritables

54. Si l'on veut mettre l'IA au service d'une bonne gouvernance, l'une des principales difficultés à surmonter tient au défaut de participation réelle des différentes parties prenantes aux processus décisionnels clés. Les stratégies de haut niveau en matière d'IA et les directives relatives à l'éthique sont souvent élaborées par des experts issus d'une poignée d'États où se concentre le développement de l'IA, sans une contribution suffisante des autres régions. Les premiers concernés par l'utilisation de l'IA dans les services publics ou la mise en place de dispositifs de surveillance basés sur l'IA, notamment le grand public, les membres de communautés marginalisées et les populations des petits États ou des États peu avancés sur le plan technologique, sont souvent exclus des processus d'élaboration des politiques en la matière. Cette exclusion illustre, plus généralement, un déficit démocratique dans la gouvernance des technologies de l'information. Les principes relatifs à l'IA adoptés à l'échelle mondiale pourront difficilement être appliqués s'ils ne tiennent pas compte des priorités de la majeure partie de la population mondiale.

55. Au plan mondial, on observe une grande diversité entre les États pour ce qui est du degré de préparation à l'IA et des capacités de gouvernance en la matière. Certains gouvernements, notamment dans les pays développés, ont mis en place des cadres nationaux et régionaux en matière d'IA et participent activement aux forums internationaux, tandis que beaucoup d'autres commencent seulement à élaborer des politiques ou des lignes directrices dans ce domaine. La coopération régionale permet de donner davantage la parole aux parties sous-représentées. Dans certaines régions, les États ont commencé à collaborer par l'intermédiaire d'organisations régionales afin d'adopter, concernant l'éthique de l'IA et son application à la gouvernance, des positions communes qui reflètent à la fois les valeurs locales partagées et les principes universels des droits de l'homme, ce qui montre que les pays en développement peuvent œuvrer collectivement en vue d'influencer les nouvelles normes mondiales relatives à l'IA. Toutefois, de telles initiatives supposent le soutien de l'ensemble de la communauté internationale.

56. Il est tout aussi essentiel d'assurer une participation véritable au plan national. Lorsque les gouvernements utilisent l'IA dans l'administration publique, que ce soit pour automatiser l'octroi des prestations sociales, mettre en place des systèmes de surveillance urbaine intelligente ou utiliser des outils judiciaires assistés par l'IA, les personnes et les collectivités concernées devraient pouvoir exercer, outre le droit à l'information et à la transparence, une influence réelle sur la conception et l'utilisation de ces systèmes. Le fait que les individus doivent pouvoir participer à la prise des décisions qui concernent leurs droits est un principe général du droit. Ce principe s'étend à l'adoption des technologies d'IA dans la gouvernance. Pourtant, dans de nombreux pays, la participation citoyenne à l'élaboration des politiques technologiques reste limitée ou superficielle.

57. Pour combler cette lacune en matière de gouvernance, les États devraient mettre en place des mécanismes officiels de consultation et de contrôle chaque fois que doivent être adoptés des systèmes d'IA ayant une incidence importante sur le plan social. Il convient de mettre en place des procédures de consultation publique avant de mettre en œuvre tout système fondé sur l'IA susceptible d'avoir une incidence sur les droits de l'homme ou l'intérêt général. Il convient de garantir la représentation de la population et de la société civile au sein des organes de contrôle et de mettre en place des comités d'éthique pour les initiatives gouvernementales en matière d'IA. En associant des représentants des collectivités concernées, des organisations de défense des droits de l'homme et des experts indépendants, ces instances et comités sont mieux à même d'évaluer les implications sociétales de l'utilisation de l'IA et de formuler des recommandations sur les mesures de protection à adopter. Ces groupes apportent de précieuses connaissances sur les répercussions en matière de droits de l'homme et les contextes culturels dont il n'est bien souvent pas tenu compte dans le cadre des approches technocratiques. Pour que l'IA réponde aux objectifs d'une bonne gouvernance, son utilisation doit être encadrée avec la participation véritable des

personnes auxquelles elle est destinée, le but étant de renforcer la confiance du public et d'assurer une mise en œuvre au service de l'intérêt général.

C. Définir les responsabilités des différentes parties prenantes pour une mise en œuvre de l'intelligence artificielle fondée sur les droits

58. Pour mettre en place un écosystème d'IA qui garantisse le respect des droits de l'homme et permette de prévenir les abus, toutes les parties prenantes doivent mener une action concertée fondée sur les principes d'équité, de souveraineté⁷¹ et de justice sociale. C'est aux États qu'il incombe au premier chef de veiller à ce que le développement et le déploiement de l'IA respectent les normes internationales relatives aux droits de l'homme. Cette obligation commence par la mise en place de cadres juridiques et réglementaires solides, comprenant des exigences en matière de transparence, de responsabilité et de contrôle des dispositifs de prise de décisions fondés sur des algorithmes dans le secteur public⁷². Les gouvernements devraient en outre investir dans l'éducation publique et la littératie numérique afin que chacun comprenne comment l'IA est utilisée dans la gouvernance et connaisse ses droits dans ce contexte. Les pays en développement ont besoin d'une autonomie et d'un soutien suffisants pour réglementer l'IA conformément à leurs priorités sociales spécifiques. Étant donné que de nombreux pays en développement ne disposent pas des ressources nécessaires pour mettre en place de manière autonome des dispositifs de gouvernance efficaces pour les systèmes algorithmiques, la coopération internationale devrait se concentrer sur le renforcement des capacités techniques par le transfert de technologie, la formation en matière de réglementation et les échanges de connaissances Sud-Sud, le but étant de permettre à tous les pays de tirer parti de l'IA pour promouvoir la bonne gouvernance. En conséquence, les normes mondiales devraient demeurer fondées sur des principes et suffisamment souples pour permettre aux États de les appliquer en tenant compte de leurs contextes institutionnels et culturels⁷³. Les gouvernements doivent ancrer fermement la gouvernance de l'IA dans les valeurs universelles des droits de l'homme, tout en orientant activement le développement technologique afin qu'il réponde aux besoins nationaux et locaux.

59. Lorsque les États confient à des acteurs privés la prestation de services publics, notamment dans les domaines de l'éducation, des soins de santé et de l'approvisionnement en eau, leurs obligations en matière de droits de l'homme restent pleinement applicables. Lorsque des systèmes d'IA sont utilisés dans la prestation de services publics, les obligations en matière de droits de l'homme s'appliquent conformément aux principes établis, notamment l'imputation d'un comportement à l'État, le cas échéant et, en tout état de cause, l'obligation de faire preuve de diligence raisonnable à l'égard des acteurs privés. Les États doivent veiller à ce que ce déploiement ne compromette pas la disponibilité, l'accessibilité, l'acceptabilité ni la qualité des services, et ne donne pas lieu à des discriminations ni à d'autres répercussions négatives sur les droits de l'homme⁷⁴. Cela relève des obligations positives des États en matière de droits de l'homme, qui supposent d'établir et de faire respecter des normes réglementaires s'appliquant de la même manière à la prestation de services privés fondés sur l'IA, d'assurer un suivi et une évaluation continus, de garantir la conduite d'enquêtes efficaces en cas de violations des droits de l'homme et d'assurer des recours utiles aux personnes dont les droits ont été lésés. Les obligations de l'État dans ce contexte restent pleinement applicables quels que soient les moyens techniques utilisés pour la prestation des services⁷⁵.

⁷¹ Dans ce contexte, on entend par souveraineté la capacité des États, en particulier des pays en développement, d'exercer un contrôle effectif et autonome sur les décisions gouvernementales relatives à l'IA, la gouvernance des données et les infrastructures numériques, sans dépendance extérieure excessive. Concernant les difficultés que pose la concentration du développement de l'IA et les risques liés à la dépendance numérique, voir les paragraphes 50 à 54 ci-dessus.

⁷² Voir [A/HRC/48/31](#).

⁷³ UNESCO, « Recommandation sur l'éthique de l'intelligence artificielle ».

⁷⁴ Voir [A/HRC/59/53](#).

⁷⁵ Comité des droits économiques, sociaux et culturels, observation générale n° 24 (2017), par. 11 ; Commission africaine des droits de l'homme et des peuples, observation générale n° 7, par. 43.

60. Une grande partie des technologies d'IA que les gouvernements acquièrent ou utilisent est développée par le secteur privé. Conformément aux Principes directeurs relatifs aux entreprises et aux droits de l'homme, toutes les entreprises, y compris celles qui conçoivent et fournissent des systèmes d'intelligence artificielle, ont pour responsabilité de respecter les droits de l'homme. Dans la pratique, les entreprises devraient faire preuve de diligence raisonnable en matière de droits de l'homme tout au long de leurs processus de développement de l'IA⁷⁶. Cela suppose de consulter les parties prenantes, d'évaluer les risques possibles tels que les biais algorithmiques, la discrimination et d'autres formes d'abus, et d'adapter les choix de conception afin d'atténuer les préjudices. Les entreprises doivent également s'attaquer aux biais systémiques et à la prédominance des points de vue émanant d'un nombre limité de régions dans leurs outils d'IA. Pour y parvenir, elles doivent diversifier leurs effectifs ainsi que les jeux de données utilisés pour l'entraînement des modèles, afin d'assurer une meilleure représentation de la diversité mondiale. Lorsqu'on y associe véritablement les chercheurs locaux et que l'on tient compte, comme il se doit, des contextes locaux, ces efforts peuvent contribuer à garantir que les produits fondés sur l'IA sont mieux adaptés à une plus grande diversité de contextes sociaux et culturels.

61. La société civile, notamment les organisations non gouvernementales, les établissements universitaires, les médias, les associations locales et les jeunes, joue un rôle indispensable pour ce qui est de garantir que l'utilisation de l'IA dans la gouvernance est fondée sur le respect des droits de l'homme. Il convient de donner à ces acteurs les moyens d'agir et de leur offrir des voies officielles de participation aux processus décisionnels. Par leurs compétences, la société civile et les acteurs du milieu universitaire peuvent contribuer à l'élaboration des garanties nécessaires, telles que l'obligation de procéder à des audits algorithmiques indépendants ou la mise en place de mécanismes permettant aux populations concernées de faire part de leurs observations sur les systèmes d'IA utilisés dans le secteur public. Les organisations de la société civile jouent également un rôle essentiel pour ce qui est de sensibiliser le public à l'IA et d'améliorer ses connaissances en la matière. En rendant compréhensibles les implications de l'IA en matière de droits de l'homme, ces organisations permettent aux populations de comprendre ces enjeux et de demander des comptes aux responsables.

62. Les organisations internationales, notamment l'ONU, ont pour responsabilité de mettre en place des cadres inclusifs visant à faire en sorte que l'IA soit mise au service d'une bonne gouvernance, respectueuse des droits de l'homme. Les initiatives prises au plan mondial doivent expressément tenir compte des points de vue des pays en développement. Des voix provenant des pays en développement et des communautés marginalisées mettent en garde contre le risque que les initiatives internationales ne reproduisent les hiérarchies de pouvoir existantes sous couvert de gouvernance mondiale. En conséquence, l'élaboration de normes internationales, qu'il s'agisse de lignes directrices non contraignantes ou d'éventuels instruments juridiquement contraignants, devrait s'inscrire dans le cadre de processus inclusifs et transparents au sein desquels tous les États ont voix au chapitre dans des conditions d'égalité. Il pourrait s'avérer nécessaire d'étudier de nouveaux mécanismes de gouvernance mondiale afin de mieux concilier l'IA et les droits de l'homme.

IV. Recommandations

A. À l'intention des États

63. Les États devraient adopter des cadres législatifs et réglementaires complets régissant l'utilisation de l'IA dans l'administration publique ; ces cadres devraient rendre obligatoire la réalisation d'études d'impact en matière de droits de l'homme avant la mise en œuvre et régulièrement par la suite, exiger la transparence et l'explicabilité des algorithmes, prévoir la mise en place d'organes de contrôle indépendants chargés de vérifier que les systèmes d'IA ne présentent pas de biais, qu'ils sont précis et qu'ils respectent le droit international des droits de l'homme, et garantir

⁷⁶ Voir [A/HRC/59/32](#).

la conduite d'enquêtes efficaces et l'existence de recours utiles lorsque l'utilisation de l'IA entraîne des violations des droits de l'homme.

64. Les États devraient garantir une participation véritable du public aux décisions relatives à l'introduction de systèmes d'IA dans le secteur public en mettant en place des procédures de consultation transparentes qui associent activement la société civile, les experts techniques et les populations concernées, afin de veiller à ce que les technologies de gouvernance tiennent compte de la diversité des besoins et des valeurs de la société.

65. Les États devraient réglementer strictement l'utilisation des systèmes de surveillance basés sur l'IA afin de prévenir toute atteinte aux droits à la vie privée, à la liberté d'expression et à la liberté d'association, en veillant à ce que, dans tous les cas, l'utilisation de ces systèmes respecte les principes de légalité, de nécessité et de proportionnalité et soit soumise à un contrôle juridictionnel efficace et indépendant.

66. Les États devraient préserver les processus démocratiques et l'intégrité des écosystèmes informationnels en favorisant l'éducation aux médias et l'acquisition de compétences numériques, en imposant la transparence de la publicité politique et l'étiquetage des contenus générés par l'intelligence artificielle, en examinant la possibilité de mettre en place des mécanismes appropriés permettant de vérifier l'intervention effective d'êtres humains ou en facilitant l'utilisation de ces mécanismes, et en soutenant le pluralisme, tout en veillant à ce que les mesures visant à lutter contre la désinformation préservent un juste équilibre avec la liberté d'expression et ne donnent pas lieu à une censure injustifiée.

67. Les États devraient veiller à ce que l'intégration de l'IA dans les procédures judiciaires et les établissements pénitentiaires préserve le contrôle humain, l'équité des procédures et la dignité inhérente à l'être humain, en interdisant le recours exclusif à des systèmes automatisés pour la prise de décisions ayant une incidence sur la liberté, tout en continuant de garantir des contacts humains véritables dans les lieux de détention.

68. Les États devraient s'attaquer de manière proactive aux répercussions socioéconomiques de l'automatisation fondée sur l'IA en adaptant les systèmes éducatifs, en proposant des programmes de reconversion professionnelle et en renforçant les mécanismes de protection sociale afin de préserver le droit au travail et de garantir des transitions justes.

69. Les États devraient évaluer et atténuer les effets qu'a l'utilisation de l'IA sur l'environnement, notamment la consommation importante d'énergie et d'eau des centres de données ainsi que la production de déchets électroniques, en veillant à ce que la mise en place d'une gouvernance fondée sur l'IA ne porte pas atteinte au droit à un environnement propre, sain et durable, ni ne transfère la charge environnementale vers les pays en développement.

B. À l'intention de la communauté internationale

70. La communauté internationale devrait veiller à assurer une gouvernance mondiale inclusive et équitable de l'IA en garantissant aux pays en développement et aux représentants des pays du Sud une participation effective dans les processus d'élaboration des normes internationales, en veillant à ce qu'ils aient voix au chapitre, dans des conditions d'égalité, dans ces processus, et en évitant que soient imposés des cadres réglementaires uniformes qui ne tiennent pas compte des contextes locaux ou qui perpétuent des rapports de dépendance numérique.

71. La communauté internationale devrait apporter un soutien en matière de renforcement des capacités, de transfert de technologies et d'aide financière afin de permettre à tous les États, en particulier aux pays à faible revenu, d'élaborer et de mettre en œuvre des dispositifs de gouvernance de l'IA qui soient efficaces et adaptés au contexte, et devrait subordonner l'aide internationale au développement destinée

aux projets numériques à la mise en place de solides garanties en matière de droits de l'homme.

72. La communauté internationale devrait envisager d'instaurer, au sein du système des Nations Unies, des mécanismes spécifiques, tels qu'un rapporteur spécial ou un groupe de travail sur les nouvelles technologies numériques et leurs incidences sur les droits de l'homme, afin de suivre l'évolution de la situation et de formuler des recommandations sur les mesures de protection à adopter. Il conviendrait également d'envisager de mettre en place, sous les auspices de l'ONU, un cadre spécifique qui engagerait les États à respecter des obligations concrètes concernant les répercussions de l'IA sur les droits de l'homme.

C. À l'intention des entreprises

73. Les entreprises qui conçoivent, développent ou utilisent des systèmes d'IA doivent respecter les droits de l'homme, notamment en faisant preuve de diligence raisonnable en la matière tout au long du cycle de vie de l'IA, en évaluant et en atténuant les risques de biais algorithmique, de discrimination et d'atteinte à la vie privée, en consultant les communautés concernées, en mettant en place des mécanismes de réclamation efficaces permettant aux utilisateurs de signaler les préjudices subis et d'obtenir réparation, et en participant au bon fonctionnement de ces mécanismes.

74. Les entreprises devraient promouvoir une IA inclusive et équitable en diversifiant leurs effectifs et les jeux de données d'entraînement afin de tenir compte de la diversité mondiale, en veillant à ce que les outils d'IA s'adaptent aux différents contextes sociaux, culturels et régionaux, et en s'abstenant de toute pratique susceptible de perpétuer des préjugés historiques ou de creuser les inégalités numériques.

D. À l'intention de l'ensemble des parties prenantes

75. Toutes les parties prenantes, notamment les États, les organisations internationales, les entreprises, la société civile, le milieu universitaire et les populations concernées, devraient collaborer dans le cadre de processus transparents et inclusifs faisant intervenir différentes parties, afin d'élaborer des mesures de gouvernance qui garantissent le respect des droits de l'homme. Cette collaboration devrait s'appuyer sur un suivi indépendant des incidences sociétales de l'IA, sur l'existence de voies de recours accessibles et sur des outils communs, tels que des cadres d'évaluation de l'incidence des algorithmes et des normes d'audit, afin de garantir que la gouvernance de l'IA reste centrée sur les droits de l'homme, qu'elle répond aux besoins de tous et qu'elle respecte la dignité humaine.

76. Toutes les parties prenantes devraient soutenir activement les organisations de la société civile, en particulier dans les pays en développement, au moyen d'initiatives de financement et de renforcement des capacités, afin qu'elles puissent permettre au public de mieux comprendre ce qu'est l'IA, surveiller les incidences de l'IA et faire en sorte que les personnes les plus touchées par les technologies émergentes puissent contribuer à définir les modalités d'utilisation de ces technologies.