Nations Unies A/HRC/48/31

Assemblée générale

Distr. générale
13 septembre 2021
Français
Original : anglais

Conseil des droits de l'homme
Quarante-huitième session
13 septembre-1er octobre 2021
Points 2 et 3 de l'ordre du jour
Rapport annuel du Haut-Commissaire des Nations Unies aux droits de l'homme et rapports du Haut-Commissariat et du Secrétaire général

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

Le droit à la vie privée à l'ère du numérique*

Rapport de la Haute-Commissaire des Nations Unies aux droits de l'homme

Résumé

Comme l'a demandé le Conseil des droits de l'homme dans sa résolution 42/15, la Haute-Commissaire examine dans le présent rapport les incidences que l'utilisation généralisée par les États et les entreprises de l'intelligence artificielle, y compris le profilage, la prise de décisions automatisée et l'apprentissage automatique, pourrait avoir sur l'exercice du droit à la vie privée et des droits associés. Après avoir analysé le cadre juridique international, la Haute-Commissaire met en avant les aspects de l'intelligence artificielle qui facilitent l'immixtion dans la vie privée et donne des exemples de ses incidences sur le droit au respect de la vie privée et des droits associés dans quatre secteurs clefs. Elle aborde ensuite les approches permettant de relever ces défis, et formule plusieurs recommandations aux États et aux entreprises concernant l'élaboration et la mise en œuvre de mesures de protection visant à prévenir et à réduire au minimum les effets néfastes et à faciliter la pleine jouissance des avantages que l'intelligence artificielle peut offrir.

* Le présent document est soumis après la date prévue pour que l'information la plus récente puisse y figurer.

I. Introduction

1. Le présent rapport est soumis en application de la résolution 42/15 du Conseil des droits de l'homme dans laquelle le Conseil prie la Haute-Commissaire des Nations Unies aux droits de l'homme d'organiser un séminaire d'experts pour examiner les incidences que l'intelligence artificielle, y compris le profilage, la prise de décisions automatisée et l'apprentissage automatique, si elle n'est pas accompagnée des garanties appropriées, pourrait avoir sur l'exercice du droit à la vie privée ainsi que de lui soumettre un rapport thématique sur cette question à sa quarante-cinquième session[1].

2. Aucune autre évolution technologique au cours de ces dernières années n'a eu autant d'emprise sur l'imaginaire collectif que l'intelligence artificielle (IA), en particulier les technologies d'apprentissage automatique[2]. En effet, ces technologies peuvent représenter un formidable outil au service du bien, en aidant les sociétés à surmonter certains des grands défis de l'époque actuelle. Toutefois, elles peuvent également avoir des effets négatifs, voire catastrophiques, si elles sont déployées sans tenir suffisamment compte de leurs incidences sur les droits de l'homme.

3. Bien que le présent rapport ne porte pas sur la pandémie de maladie à coronavirus (COVID-19), la crise sanitaire mondiale actuelle illustre parfaitement et de manière très visible la rapidité et l'ampleur de la pénétration de l'intelligence artificielle dans diverses sphères de la vie à travers le monde ainsi que ses incidences. Des systèmes de traçage des contacts exploitant plusieurs types de données (géolocalisation, carte de crédit, système de transport, santé et démographie) et des informations provenant de réseaux personnels ont été utilisés pour suivre la propagation de la maladie. Des systèmes d'intelligence artificielle ont été utilisés pour signaler les personnes potentiellement contaminées ou infectieuses, les obligeant à s'isoler ou à se mettre en quarantaine. Ils ont été employés pour attribuer des notes prédictives qui ont donné des résultats discriminatoires à l'égard des élèves des écoles publiques et des quartiers défavorisés. Ces faits montrent les incidences diverses et nombreuses que les systèmes d'intelligence artificielle ont sur notre quotidien. L'exercice du droit au respect de la vie privée est entravé dans tous ces cas, l'intelligence artificielle exploitant des informations personnelles et prenant souvent des décisions qui ont des effets tangibles sur la vie des personnes. Néanmoins, la question de la vie privée est intimement liée aux diverses répercussions sur la jouissance d'autres droits, tels que les droits à la santé, à l'éducation, à la liberté de mouvement, à la liberté de réunion pacifique, à la liberté d'association et à la liberté d'expression.

4. En 2019, dans « La plus haute aspiration : Un appel à l'action en faveur des droits humains », le Secrétaire général de l'ONU a reconnu que l'ère numérique ouvrait de nouvelles perspectives en matière de bien-être, de connaissance et de découverte. Il a souligné que les technologies numériques étaient autant de nouveaux outils qui nous permettaient de plaider en faveur de nos droits, de les défendre et de les exercer. Cependant, les nouvelles technologies sont trop souvent utilisées pour porter atteinte à ces droits, en particulier ceux des plus vulnérables et des laissés-pour-compte, qui font l'objet, par exemple, de mesures de surveillance, de répression et de censure et d'actes de harcèlement en ligne, notamment les défenseuses et défendeurs des droits de l'homme. Depuis la numérisation des systèmes de protection sociale, certains parmi ceux qui en ont le plus besoin peuvent même voir leur couverture sociale restreinte, en dépit des améliorations qu'elle apporte. Le Secrétaire général a souligné que les avancées technologiques ne devaient pas servir à remettre en cause les droits fondamentaux, à creuser les inégalités ou à aggraver les discriminations existantes. Il a ajouté que la justice, le respect du principe de responsabilité, l'explicabilité et la transparence devaient être les maîtres mots de sa gouvernance. Dans la

[1] L'établissement du rapport a été reporté. Voir A/HRC/45/26 et A/HRC/47/61.

[2] Il n'existe pas de définition généralement acceptée du terme « intelligence artificielle ». Dans le présent rapport, il est utilisé pour désigner une constellation de techniques et de procédés permettant d'utiliser des ordinateurs pour accompagner ou remplacer des opérateurs humains dans des tâches de résolution de problèmes ou de prise de décisions (A/73/348, par. 3), y compris mais pas exclusivement l'apprentissage automatique et l'apprentissage profond.

sphère de la sécurité, il a réitéré son appel en faveur d'une interdiction des systèmes d'armes létales autonomes.

5. Le présent rapport s'appuie sur les deux précédents rapports de la Haute- Commissaire sur la question du droit à la vie privée à l'ère du numérique[3]. Il tient également compte des enseignements tirés du séminaire virtuel d'experts organisé conformément à la résolution 42/15 du Conseil, qui s'est tenu les 27 et 28 mai 2020, ainsi que des réponses à l'appel à contribution lancé par la Haute-Commissaire afin de l'enrichir[4].

II. Cadre juridique

6. L'article 12 de la Déclaration universelle des droits de l'homme, l'article 17 du Pacte international relatif aux droits civils et politiques et plusieurs autres instruments internationaux et régionaux relatifs aux droits de l'homme reconnaissent le droit à la protection de la vie privée comme un droit humain fondamental[5]. Il joue un rôle central dans l'équilibre des pouvoirs entre l'État et l'individu et constitue un droit fondamental pour une société démocratique[6]. Son importance pour la jouissance et l'exercice d'autres droits de l'homme en ligne et hors ligne[7] dans un monde de plus en plus centré sur les données ne cesse de croître.

7. Le droit à la protection de la vie privée est une expression de la dignité humaine et il est indissociable de la protection de l'autonomie et de l'identité personnelles[8]. Les aspects de la vie privée qui revêtent une importance particulière dans le contexte de l'utilisation de l'IA comprennent la confidentialité des informations personnelles, à savoir qui existent ou qui peuvent être extrapolées au sujet d'une personne et de sa vie, ainsi que les décisions fondées sur cette information[9], et la liberté de prendre des décisions concernant sa propre identité.

8. Nul ne peut être l'objet d'immixtion arbitraire ou illégale dans sa vie privée[10]. Le terme « illégal » signifie que toute atteinte au droit à la protection de la vie privée par un État doit être prévue par la loi et conforme à celle-ci. La loi elle-même ne doit pas être contraire aux dispositions, aux buts et aux objectifs du Pacte international relatif aux droits civils et politiques et doit énoncer en détail les circonstances précises dans lesquelles une telle immixtion est autorisée[11]. L'introduction de la notion d'arbitraire a pour objet de garantir que même une immixtion prévue par la loi soit conforme aux dispositions, aux buts et aux objectifs du Pacte et soit, dans tous les cas, raisonnable eu égard aux circonstances particulières[12]. Par conséquent, toute immixtion dans le droit à la protection de la vie privée doit servir un objectif légitime, être nécessaire pour atteindre cet objectif légitime et être

[3] A/HRC/27/37 et A/HRC/39/29.

[4] Voir www.ohchr.org/EN/Issues/DigitalAge/Pages/cfi-digital-age.aspx pour l'appel à contribution et les contributions reçues.

[5] Voir l'article 16 de la Convention relative aux droits de l'enfant, l'article 14 de la Convention internationale sur la protection des droits de tous les travailleurs migrants et des membres de leur famille, l'article 22 de la Convention relative aux droits des personnes handicapées, l'article 10 de la Charte africaine des droits et du bien-être de l'enfant, l'article 11 de la Convention américaine relative aux droits de l'homme et l'article 8 de la Convention de sauvegarde des droits de l'homme et des libertés fondamentales (la Convention européenne des droits de l'homme).

[6] A/HRC/39/29, par. 11.

[7] Comité des droits de l'enfant, Observation générale n° 25 (2021), par. 67 et 68 ; et A/HRC/39/29, par. 11.

[8] Comité des droits de l'enfant, Observation générale n° 25 (2021), par. 67 et Cour européenne des droits de l'homme, *Goodwin c. Royaume-Uni*, requête n° 28957/95 arrêt du 11 juillet 2002, par. 90.

[9] A/HRC/39/29, par. 5.

[10] Pour une analyse détaillée des termes « arbitraire » et « illégal », voir A/HRC/27/37, par. 21 à 27.

[11] Comité des droits de l'homme, Observation générale n° 16 (1988), par. 3 et 8.

[12] Ibid., par. 4.

proportionnée[13]. Toute restriction doit également être l'option la moins intrusive disponible et ne doit pas porter atteinte à l'essence même du droit à la vie privée[14].

9. Le droit à la protection de la vie privée s'applique à tous. Toutes différences de protection de ce droit fondées sur la race, la couleur, le sexe, la langue, la religion, les opinions politiques ou autres, l'origine nationale ou sociale, la fortune, la naissance ou toute autre situation sont incompatibles avec le principe de non-discrimination énoncé au paragraphe 1 de l'article 2 et à l'article 3 du Pacte international relatif aux droits civils et politiques. Toute discrimination fondée sur ces motifs viole la garantie d'égalité devant la loi au sens de l'article 26 du Pacte.

10. Le paragraphe 1 de l'article 2 du Pacte international relatif aux droits civils et politiques impose aux États de s'engager à respecter et à garantir à tous les individus se trouvant sur leur territoire et relevant de leur compétence les droits reconnus dans le Pacte, sans distinction aucune. En d'autres termes, les États doivent non seulement s'abstenir de violer les droits reconnus dans le Pacte[15], mais ils ont également l'obligation de prendre des mesures positives pour protéger la jouissance de ces droits. Ils ont à ce titre le devoir d'adopter des mesures législatives et autres adéquates pour protéger les individus contre toute immixtion dans leur vie privée, qu'elle émane des autorités de l'État ou de personnes physiques ou morales[16]. Cette obligation figure également dans le premier pilier des Principes directeurs relatifs aux entreprises et aux droits de l'homme, qui précise que les États sont tenus de protéger les citoyens contre les atteintes aux droits de l'homme auxquelles participent des entreprises.

11. Les entreprises ont la responsabilité de respecter tous les droits de l'homme internationalement reconnus. À cet effet, elles devraient éviter de porter atteinte aux droits fondamentaux d'autrui et remédier aux incidences négatives sur les droits de l'homme dans lesquelles elles ont une part. Le pilier II des Principes directeurs relatifs aux entreprises et aux droits de l'homme fournit à toutes les entreprises un plan directeur faisant autorité sur la manière d'assumer cette responsabilité[17], laquelle concerne l'ensemble des activités et des relations commerciales des entreprises.

III. Incidences de l'intelligence artificielle sur le droit à la protection de la vie privée et les autres droits de l'homme

A. Caractéristiques pertinentes des systèmes d'intelligence artificielle

12. Le fonctionnement des systèmes d'AI peut faciliter et multiplier de diverses manières les immixtions dans la vie privée et autres interférences avec les droits. Il s'agit notamment d'applications entièrement nouvelles ainsi que de caractéristiques de ces systèmes qui étendent, intensifient ou encouragent l'ingérence dans le droit à la vie privée, plus particulièrement par la collecte et l'utilisation accrues de données personnelles.

13. Les systèmes d'AI s'appuient généralement sur de vastes ensembles de données, comprenant souvent des données personnelles et encourageant une collecte, un stockage et un traitement à grande échelle. De nombreuses entreprises optimisent leurs services afin de

[13] *Toonen c. Australie* (CCPR/C/50/D/488/1992), par. 8.3, *Van Hulst c. Pays-Bas* (CCPR/C/82/D/903/1999), par. 7.3 et 7.6, *Madhewoo c. Maurice* (CCPR/C/131/D/3163/2018), par. 7.5, et CCPR/C/USA/CO/4, par. 22. Voir également Comité des droits de l'enfant, Observation générale nº 25 (2021), par. 69.

[14] Comité des droits de l'homme, observation générale nº 31 (2004), par. 6. A/HRC/27/37, par. 22, et A/HRC/39/29, par. 10.

[15] Comité des droits de l'homme, observation générale nº 31 (2004), par. 6.

[16] A/HRC/39/29, par. 23. Voir également Comité des droits de l'homme, observations générales nº 16 (1988), par. 1 et 9, et nº 31 (2004), par. 8 ; et Comité des droits de l'enfant, Observation générale nº 25 (2021), par. 36 à 39.

[17] Dans sa résolution 17/4, le Conseil des droits de l'homme a approuvé à l'unanimité les Principes directeurs relatifs aux entreprises et aux droits de l'homme.

collecter le plus grand nombre de données possible[18]. Par exemple, les entreprises en ligne et notamment les réseaux sociaux s'appuient sur la collecte et la monétisation de quantités massives de données sur les internautes[19]. Ce que l'on appelle l'Internet des objets est une source de données dont la croissance est exponentielle et qui est exploitée par les entreprises comme par les États. La collecte de données se fait dans des espaces intimes, privés et publics[20]. Les entreprises de courtage de données acquièrent, fusionnent, analysent et partagent des données personnelles avec d'innombrables destinataires. Ces transactions de données échappent largement au droit de regard du public et ne sont que peu encadrées par les législations en vigueur[21]. Les ensembles de données qui en résultent sont vastes et les informations recueillies sont d'une ampleur sans précédent.

14. Outre le fait que la vie privée des personnes est ainsi dévoilée aux entreprises et aux États, ces ensembles de données rendent les individus vulnérables à bien d'autres égards. Les atteintes à la sécurité des données ont exposé à plusieurs reprises les informations personnelles de millions d'individus[22]. Les grands ensembles de données permettent d'innombrables formes d'analyse et de partage de données avec des tiers, équivalant souvent à d'autres intrusions dans la vie privée avec de nouvelles conséquences négatives sur les droits de l'homme. Les dispositions permettant aux agences gouvernementales d'avoir un accès direct à ces ensembles de données détenus par les entreprises renforcent notamment la probabilité d'une ingérence arbitraire ou illégale dans l'exercice du droit à la protection de la vie privée des personnes concernées[23]. La fusion de données provenant de diverses sources peut faciliter la désanonymisation, ce qui particulièrement préoccupant[24]. Dans le même temps, la création d'ensembles de données peut avoir des répercussions sur l'identité des individus. Par exemple, un ensemble de données qui enregistre le genre de manière binaire ne classe pas correctement les personnes qui ne s'identifient ni comme homme ni comme femme. Le stockage à long terme des données personnelles comporte également des risques particuliers, car celles-ci peuvent être exposées à des formes futures d'exploitation non envisagées au moment de leur collecte[25]. Au fil du temps, les données peuvent devenir inexactes, non pertinentes ou entraîner la répétition d'erreurs d'identification historiques, avec des résultats biaisés ou erronés dans le traitement futur des données[26].

15. Il convient de noter que les systèmes d'IA ne reposent pas exclusivement sur le traitement de données à caractère personnel. Cependant, même lorsque les données personnelles ne sont pas concernées, leur utilisation peut toujours avoir des répercussions négatives sur les droits de l'homme[27], y compris sur le droit à la protection de la vie privée, comme expliqué ci-après.

16. Les outils de l'IA sont largement utilisés pour comprendre les modèles de comportement humain. En ayant accès aux bons ensembles de données, il est possible de tirer des conclusions sur le nombre de personnes d'un quartier donné qui sont susceptibles de fréquenter un certain lieu de culte, sur les émissions de télévision qu'elles préfèrent et même sur l'heure à laquelle elles ont tendance à se lever et à se coucher. Ces outils peuvent

[18] Wolflie Christl, *Corporate surveillance in everyday life* (Vienne, Cracked Lab − Institute for Critical Digital Culture, 2017).

[19] Contribution de Ranking Digital Rights.

[20] Contributions du Centre for Communication Governance at National Law University Delhi, de Derechos Digitales, Digital Rights Watch, Global Partners Digital, International Center for Not-for-Profit Law et Universidade Federal de Uberlândia.

[21] Aaron Rieke et autres, *Data brokers in an open society* (Londres, Open Society Foundation, 2016).

[22] Voir, par exemple, www.ftc.gov/news-events/press-releases/2019/07/equifax-pay-575-million-part-settlement-ftc-cfpb-states-related.

[23] Contribution d'Initiative mondiale des réseaux.

[24] Contributions du Centre for Communication Governance de la National Law University de Delhi, Derechos Digitales et Privacy International.

[25] Contribution d'OVD-Info.

[26] Comité pour l'élimination de la discrimination raciale, recommandation générale nº 36 (2020), par. 33.

[27] Conseil de l'Europe, « Lignes directrices sur le traitement des impacts des systèmes algorithmiques sur les droits de l'homme », (annexe à la recommandation CM/Rec(2020)1 du Comité des Ministres aux États membres sur les impacts des systèmes algorithmiques sur les droits de l'homme), sect. A, par. 6.

permettre de faire des déductions très poussées sur les individus, y compris sur leur état mental et physique[28], et peuvent permettre d'identifier des groupes, tels que des personnes ayant des tendances politiques ou personnelles particulières. L'intelligence artificielle est également utilisée pour évaluer la probabilité de comportements ou d'événements futurs. Les déductions et les prédictions, malgré leur nature probabiliste, peuvent servir de base à la prise de décisions portant atteinte aux droits des personnes, parfois de manière entièrement automatisée.

17. De nombreuses immixtions et prédictions portent gravement atteinte à l'exercice du droit à la vie privée, y compris le droit à l'autonomie de la personne et le droit de se forger une identité. Elles soulèvent également de multiples questions concernant d'autres droits, tels que le droit à la liberté de pensée et d'opinion, le droit à la liberté d'expression, le droit à un procès équitable et les droits connexes.

18. Les décisions fondées sur l'IA ne sont pas exemptes d'erreurs. En fait, l'extensibilité des solutions dans ce domaine peut accroître considérablement les effets négatifs de taux d'erreur apparemment faibles[29]. Les résultats erronés des systèmes d'intelligence artificielle ont plusieurs origines. Tout d'abord, les résultats des algorithmes comportent des éléments probabilistes, ce qui signifie qu'il existe une incertitude liée à ces résultats[30]. En outre, la pertinence et l'exactitude des données utilisées sont souvent sujettes à caution. Par ailleurs, des attentes irréalistes peuvent conduire au déploiement d'outils d'IA qui n'ont pas été conçus pour atteindre les objectifs souhaités. Par exemple dans le domaine médical, une analyse de plusieurs centaines d'outils de diagnostic de la COVID-19 et de prévision des risques liés à cette maladie, dont le développement avait suscité de grands espoirs, a montré qu'aucun d'entre eux n'était adapté pour un usage clinique[31].

19. Les résultats qui s'appuient sur des données erronées peuvent contribuer de multiples façons à porter atteinte aux droits de l'homme, par exemple en signalant à tort qu'un individu est un terroriste potentiel ou qu'il a commis une fraude à l'aide sociale. Les ensembles de données biaisés qui conduisent à des décisions discriminatoires basées sur des systèmes d'intelligence artificielle sont particulièrement préoccupants[32].

20. Les processus de décision de nombreux systèmes d'IA sont opaques. La complexité de l'environnement des données, des algorithmes et des modèles qui sous-tendent le développement et le fonctionnement de ces systèmes, ainsi que le secret voulu par les acteurs gouvernementaux et privés sont des facteurs qui empêchent le public de comprendre de manière significative leurs conséquences pour les droits de l'homme et la société. Les systèmes d'apprentissage automatique sont un autre élément important d'opacité ; ils peuvent être capables d'identifier des modèles et de développer des prescriptions difficiles ou impossibles à expliquer[33]. C'est ce qu'on appelle couramment le problème de la « boîte noire »[34]. Du fait de cette opacité, il est compliqué de procéder à un examen approfondi d'un système d'IA ce qui peut constituer un obstacle à la responsabilisation lorsque ceux-ci ont

[28] Contributions de Derechos Digitales et Privacy International.

[29] Contribution de l'Allemagne.

[30] Agence des droits fondamentaux de l'Union européenne, étude intitulée « #BigData: discrimination in data-supported decision making », (Vienne, 2018) p. 4.

[31] Voir www.technologyreview.com/2021/07/30/1030329/machine-learning-ai-failed-covid-hospital-diagnosis-pandemic/.

[32] Comité pour l'élimination de la discrimination raciale, recommandation générale n° 36 (2020), par. 31 à 36 ; et Groupe de haut niveau sur la coopération numérique, « The age of digital interdependence » (L'ère de l'interdépendance numérique) (juin 2019), p. 17 et 18.

[33] Contribution de l'Allemagne.

[34] Voir www.scientificamerican.com/article/demystifying-the-black-box-that-is-ai/.

des effets néfastes[35]. Néanmoins, il convient de noter qu'il n'est pas indispensable que ces systèmes soient totalement impénétrables[36].

B. Inquiétudes concernant les systèmes d'intelligence artificielle dans des secteurs clefs

21. La présente section illustre comment ces préoccupations sont vécues dans la pratique en examinant quatre domaines clefs où l'application des outils de l'intelligence artificielle a suscité des inquiétudes.

L'intelligence artificielle dans le maintien de l'ordre, la sécurité nationale, la justice pénale et la gestion des frontières

22. Les États ont de plus en plus recours à l'intelligence artificielle dans les domaines du maintien de l'ordre, de la sécurité nationale, de la justice pénale et de la gestion des frontières[37]. Si bon nombre de ces applications peuvent effectivement être source d'inquiétudes, la présente section n'aborde que quelques exemples choisis pour illustrer certains des nouveaux problèmes constatés en matière de droits de l'homme.

23. Ces systèmes sont souvent utilisés comme outils de prévision. Ils s'appuient sur des algorithmes pour analyser de grandes quantités de données, y compris des données historiques, afin d'évaluer les risques et de prédire les tendances futures. En fonction de la finalité, les données de formation et les données analysées peuvent inclure, par exemple, les casiers judiciaires, les procès-verbaux d'arrestation, les statistiques de la criminalité, les dossiers d'intervention de la police dans des quartiers spécifiques, les messages sur les médias sociaux, les données de communication et les registres de voyage[38]. Ces technologies peuvent servir pour créer des profils de personnes, identifier des lieux susceptibles d'être le théâtre d'une activité criminelle ou terroriste accrue, et même signaler des individus comme suspects probables et futurs récidivistes[39].

24. Les conséquences de ces activités sur la protection de la vie privée et les droits de l'homme au sens large sont multiples. Premièrement, les ensembles de données utilisés comprennent des informations sur un grand nombre d'individus, mettant en jeu leur droit à la protection de leur vie privée. Deuxièmement, elles peuvent déclencher des interventions de l'État, telles que des fouilles, des interrogatoires, des arrestations et des poursuites, même si les évaluations de l'intelligence artificielle proprement dites ne devraient pas être considérées comme des motifs raisonnables de suspicion en raison du caractère probabiliste des prévisions. Les droits concernés sont le droit au respect de la vie privée, le droit à un procès équitable, le droit de ne pas être arrêté ou détenu arbitrairement et le droit à la vie. Troisièmement, l'opacité inhérente aux décisions fondées sur l'intelligence artificielle soulève des questions essentielles en ce qui concerne la responsabilité de l'État lorsque l'IA alimente des mesures coercitives, a fortiori dans des domaines qui souffrent généralement d'un manque de transparence, comme les activités des forces antiterroristes[40]. Quatrièmement, les outils d'analyse prédictive comportent un risque inhérent de perpétuation, voire de renforcement de la discrimination, conséquence des préjugés raciaux

[35] Voir www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/international-human-rights-law-as-a-framework-for-algorithmic-accountability/1D6D0A456B36BA7512A6AFF17F16E9B6# ; et www.cambridge.org/core/journals/international-review-of-the-red-cross/article/abs/ai-for-humanitarian-action-human-rights-and-ethics/C91D044210CADF7A0E023862CF4EE758.

[36] Voir, par exemple, Inioluwa Deborah Raji et autres, « Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing », 3 janvier 2020.

[37] Pour une analyse approfondie des conséquences sur les droits de l'homme de l'utilisation de l'intelligence artificielle et des autres technologies numériques dans la gestion des frontières, voir A/75/590.

[38] Voir la contribution de Privacy International. Voir également A/HRC/44/57, par. 35.

[39] Voir www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf.

[40] A/74/335 et A/HRC/43/46, par. 37 et 38.

et ethniques historiques intégrés dans les ensembles de données utilisés, notamment le fait que certaines minorités sont plus particulièrement visées par la police[41].

25. Les progrès réalisés dans le domaine des technologies de reconnaissance biométrique ont conduit à son utilisation croissante par les services de police et de sécurité nationale. La reconnaissance biométrique repose sur la comparaison de la représentation numérique de certaines caractéristiques d'un individu, telles que le visage, les empreintes digitales, le dessin de l'iris, la voix ou la démarche, avec d'autres caractéristiques de ce type dans une base de données[42]. Elle permet ainsi de déterminer le degré de probabilité que l'intéressé soit bien celui dont l'identité est à confirmer ou à établir. Ces processus sont de plus en plus souvent réalisés en temps réel et à distance et les autorités du monde entier ont notamment de plus en plus recours à la reconnaissance faciale à distance en temps réel[43].

26. La reconnaissance biométrique à distance en temps réel suscite de graves préoccupations au regard du droit international des droits de l'homme, que la Haute-Commissaire a déjà soulignées[44]. Certaines d'entre elles portent sur les problèmes associés aux outils prédictifs, notamment la possibilité d'une identification erronée des individus et les conséquences disproportionnées sur les membres de certains groupes[45]. En outre, la technologie de reconnaissance faciale peut être utilisée pour profiler des individus sur la base de leur ethnie, de leur race, de leur origine nationale, de leur sexe et d'autres caractéristiques[46].

27. La reconnaissance biométrique à distance comporte un risque important d'ingérence dans le droit à la vie privée. L'image d'un individu est l'un des attributs principaux de sa personnalité, du fait qu'elle traduit son originalité et lui permet de se différencier de ses semblables[47]. En outre, la reconnaissance biométrique à distance accroît considérablement les moyens dont disposent les autorités de l'État pour identifier et suivre systématiquement les personnes dans les espaces publics, entravant leur capacité à mener leur vie sans être observées et portant directement atteinte à l'exercice des droits à la liberté d'expression, de réunion pacifique et d'association, ainsi qu'à la liberté de circulation[48]. Dans ce contexte, la Haute-Commissaire se félicite donc des efforts déployés récemment afin de limiter ou d'interdire l'utilisation des technologies de reconnaissance biométrique en temps réel[49].

28. Des outils de l'intelligence artificielle ont également été mis au point pour prétendument déduire l'état émotionnel et mental des personnes à partir de leurs expressions faciales en même temps que d'autres « biométries prédictives » pour déterminer si elles constituent une menace pour la sécurité[50]. Les systèmes de reconnaissance des émotions faciales partent du principe qu'il est possible de déduire automatiquement et systématiquement l'état émotionnel des êtres humains à partir de leurs expressions faciales,

[41] Contribution de Tech Hive Advisory Limited. Voir également Comité pour l'élimination de la discrimination raciale, recommandation générale n° 36 (2020), par. 33 et le document de séance du Haut-Commissaire des Nations Unies aux droits de l'homme sur la promotion et la protection des droits de l'homme et des libertés fondamentales des Africains et des personnes d'ascendance africaine contre le recours excessif à la force et les autres violations des droits de l'homme par les responsables de l'application des lois (A/HRC/47/CRP.1), disponible sur www.ohchr.org/Documents/Issues/Racism/A_HRC_47_CRP_1.pdf, par. 15 et 19.

[42] A/HRC/31/64, par. 14.

[43] Contributions de Derechos Digitales et International Center for Not-for-Profit Law.

[44] A/HRC/44/24.

[45] Contribution de Privacy International.

[46] A/HRC/44/57, par. 39 et 40.

[47] A/HRC/44/24, par. 33. Voir également Cour européenne des droits de l'homme, *Reklos et Davourlis c. Grèce* (requête n° 1234/05), arrêt du 15 avril 2009, par. 40.

[48] Voir Conseil européen de la protection des données et Contrôleur européen de la protection des données, avis conjoint 5/2021, par. 30 et les contributions du International Center for Not-for-Profit Law et de Privacy International. Voir également A/HRC/44/24, par. 34, et A/HRC/41/35.

[49] Contribution de l'Union européenne. Voir également https://nymag.com/intelligencer/2020/01/why-we-should-ban-facial-recognition-technology.html ; et Commission européenne, Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (loi sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union, COM(2021) 206 final, 21 avril 2021, art. 5 1) d).

[50] Voir www.ohchr.org/Documents/Issues/Terrorism/biometricsreport.pdf.

ce qui ne repose sur aucune base scientifique solide[51]. Les chercheurs ont trouvé que l'association entre les émotions et les expressions faciales était peu concluante[52] et ont souligné que les expressions faciales variaient selon les cultures et les contextes[53], ce qui rendait la reconnaissance des émotions susceptible d'être biaisée et mal interprétée. Compte tenu de ces préoccupations, l'utilisation de systèmes de reconnaissance des émotions par les autorités publiques, par exemple pour cibler des personnes en vue d'une interpellation ou d'une arrestation par la police ou pour évaluer la véracité de déclarations lors d'interrogatoires, risque de porter atteinte aux droits de l'homme, notamment aux droits à la protection de la vie privée, à la liberté et à un procès équitable.

Systèmes d'intelligence artificielle et services publics

29. Les systèmes d'intelligence artificielle sont de plus en plus utilisés pour aider à fournir des services publics, souvent dans le but déclaré de développer des systèmes plus efficaces pour une prestation de services rapide et précise. Ce phénomène s'observe également de plus en plus dans les contextes humanitaires, où la fourniture de biens et de services humanitaires peut être liée à de tels systèmes. Bien qu'il s'agisse d'objectifs légitimes, voire louables, le déploiement de ces outils dans la prestation de services publics et humanitaires peut avoir un impact négatif sur les droits de l'homme si des garanties appropriées ne sont pas mises en place.

30. L'intelligence artificielle est utilisée dans divers services publics, qu'il s'agisse de prendre des décisions concernant les droits à l'aide sociale ou de signaler aux services de protection de l'enfance les familles à visiter[54]. Ces décisions sont prises à l'aide de vastes ensembles de données, qui comprennent non seulement des données détenues par l'État, mais peuvent également inclure des informations obtenues auprès d'entités privées, telles que des réseaux sociaux ou des courtiers en données, souvent recueillies en dehors des cadres juridiques de protection[55]. En outre, étant donné que les connaissances informatiques et le pouvoir sur les systèmes d'IA sont généralement détenus par des entreprises privées, ces accords signifient souvent que ces dernières ont accès à des ensembles de données contenant des informations sur des segments importants de la population. Cela soulève des problèmes de protection de la vie privée, ainsi que des inquiétudes quant à l'incidence que les préjugés historiques intégrés dans les données auront sur la prise de décisions par les autorités publiques.

31. Une préoccupation majeure concernant l'utilisation de l'IA pour fournir des services publics est qu'elle peut être discriminatoire, notamment à l'égard des groupes marginalisés[56]. Le Rapporteur spécial sur les droits de l'homme et l'extrême pauvreté a mis en garde contre « un monde dystopique où la protection sociale serait totalement dématérialisée », dans lequel le couplage sans entrave des données est utilisé pour exposer, surveiller et punir les bénéficiaires de l'aide sociale et où les conditions imposées aux bénéficiaires portent atteinte à l'autonomie et aux choix individuels[57]. Ces préoccupations ont été illustrées récemment aux Pays-Bas, où une décision de justice largement médiatisée a interdit un système numérique de détection des fraudes à l'aide sociale, estimant qu'il portait atteinte au droit à la vie privée. Le système en question conférait de larges pouvoirs aux autorités centrales et

[51] Voir www.nature.com/articles/d41586-020-00507-5.

[52] Voir https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780190613501.001.0001/acprof-9780190613501-chapter-7.

[53] Voir https://journals.sagepub.com/doi/10.1177/1529100619832930 ; et https://pubmed.ncbi.nlm.nih.gov/22509011/.

[54] Voir A/74/493.

[55] Contribution de Privacy International.

[56] Contribution de Human Rights Watch. Pour une analyse approfondie des conséquences disparates de l'automatisation dans les systèmes de protection sociale, voir Virginia Eubanks, *Automating Inequality* (New York, St. Martin's Press, 2018).

[57] Voir A/74/493. Voir également la lettre IRL 1/2020 du Rapporteur spécial, dans laquelle il fait état de préoccupations similaires concernant une carte de services numériques, et la réponse à cette lettre. Dans le présent rapport il est fait référence à plusieurs reprises aux contributions apportées par les titulaires de mandats au titre des procédures spéciales du Conseil des droits de l'homme. Toutes ces contributions peuvent être consultées à l'adresse suivante : https://spcommreports.ohchr.org/Tmsearch/TMDocuments.

locales pour partager et analyser des données qui étaient auparavant conservées séparément, lesquelles portaient notamment sur l'emploi, le logement, l'éducation, les prestations et l'assurance maladie, ainsi que d'autres données identifiables. En outre, l'outil ciblait les quartiers à faible revenu et les quartiers minoritaires, ce qui a entraîné une discrimination de facto fondée sur le milieu socioéconomique[58].

Utilisation de l'intelligence artificielle dans le contexte de l'emploi

32. Parmi une grande diversité d'employeurs, quels que soient le type et la taille de leur entreprise, il existe une demande croissante de surveillance et de gestion des travailleurs à l'aide de technologies axées sur les données, y compris les systèmes d'IA. Ce que l'on appelle l'analytique RH revendique pouvoir fournir des informations plus rationnelles et plus objectives sur les employés, ce qui peut inclure la prise de décisions automatisée pour le recrutement, les promotions ou les plans de licenciement.

33. Si ces technologies sont principalement axées sur le suivi des comportements et des performances liés à l'emploi, plusieurs applications des systèmes d'intelligence artificielle s'étendent également aux comportements et aux données non liés à l'emploi[59]. La pandémie de COVID-19 a accéléré cette tendance de deux manières. Tout d'abord, certaines entreprises qui proposent aux travailleurs des programmes de prévention en matière de santé collectent de plus en plus de données dans ce domaine. Deuxièmement, étant donné que le nombre de processus exécutés numériquement augmente du fait que le personnel télétravaille, la surveillance du lieu de travail par les systèmes d'IA s'introduit dans les foyers des personnes. Ces deux tendances augmentent le risque de fusionner les données issues de la surveillance du lieu de travail avec des données non liées au travail. Ces pratiques de surveillance basées sur l'IA présentent des risques importants pour la vie privée tout au long du cycle de vie des données. En outre, elles peuvent être utilisées à d'autres fins que celles initialement communiquées aux employés, ce qui peut entraîner un détournement d'usage[60]. Dans le même temps, la base quantitative en sciences sociales d'un grand nombre des systèmes d'IA utilisés dans la gestion des ressources humaines n'est pas solide et peut être biaisée. Par exemple, si une entreprise utilise un algorithme d'embauche construit à partir d'ensembles de données historiques qui favorisent les hommes blancs d'âge moyen, l'algorithme résultant défavorisera les femmes, les personnes de couleur et les personnes plus jeunes ou plus âgées qui auraient été tout aussi qualifiées pour occuper le poste vacant[61]. Dans le même temps, les structures de responsabilité et la transparence pour protéger les travailleurs font souvent défaut et ceux-ci sont de plus en plus confrontés à peu ou pas d'explications sur les pratiques de surveillance basées sur l'IA[62]. Si, dans certaines situations, les entreprises ont un intérêt réel à prévenir les comportements répréhensibles sur le lieu de travail, les mesures visant à défendre cet intérêt ne justifient souvent pas les pratiques très invasives de quantification des modes d'interaction sociale et des objectifs de performance liés au travail. Sur le lieu de travail et compte tenu de la relation de pouvoir entre l'employeur et l'employé, on peut également envisager des scénarios dans lesquels les travailleurs sont contraints de renoncer à leur droit à la vie privée en échange de leur travail[63].

L'intelligence artificielle au service de la gestion de l'information en ligne

34. Les plateformes des réseaux sociaux utilisent des systèmes d'intelligence artificielle pour aider à la prise de décisions en matière de gestion des contenus[64]. Les entreprises ont recours à ces systèmes pour hiérarchiser les contenus et décider de ce qu'il faut mettre en avant ou pas, notamment en personnalisant ces décisions pour différents utilisateurs individuels en fonction de leurs profils. L'automatisation est également utilisée lors de la

[58] Voir www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=25522.

[59] Voir https://journals.sagepub.com/doi/10.1177/20539517211013051.

[60] Christl, *Corporate surveillance in everyday life*.

[61] Contribution de la Pologne. Voir également www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G.

[62] Voir https://journals.sagepub.com/doi/full/10.1177/2053951720938093.

[63] Voir www.californialawreview.org/print/3-limitless-worker-surveillance/.

[64] Voir www.theverge.com/2020/11/13/21562596/facebook-ai-moderation et https://journals.sagepub.com/doi/full/10.1177/2053951720943234.

mise en œuvre de restrictions de contenu, y compris en réponse à différentes exigences légales au sein d'une même juridiction ou d'une juridiction à l'autre[65]. L'adoption de l'obligation pour les intermédiaires de mettre en place des filtres concernant les dangers perçus du monde en ligne risque d'accroître le recours généralisé à l'IA au mépris des conséquences graves de ces systèmes sur les droits à la vie privée et à la liberté d'expression aux niveaux local et mondial.

35. Les vastes ensembles de données sur lesquels s'appuient les systèmes de curation, d'amplification et de modération sont créés et continuellement enrichis grâce à une surveillance et à un profilage en ligne étendus des utilisateurs des plateformes et de leurs réseaux personnels[66]. Ce processus perpétuel de collecte d'informations et de déductions à partir de celles-ci, combiné à une concentration extrême du marché, a conduit à une situation où une poignée d'entreprises au niveau mondial détient et contrôle les profils de milliards d'individus et la sphère publique en réseau au sens large.

36. L'édition de contenu assistée par l'IA par des entreprises disposant d'une énorme puissance commerciale soulève des inquiétudes quant à l'impact sur la capacité de l'individu à former et à développer des opinions, comme l'ont souligné deux titulaires successifs du mandat de Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression[67]. En outre, les systèmes de recommandation des plateformes ont tendance à privilégier la maximisation de l'engagement des utilisateurs en s'appuyant sur des informations sur les préférences, les caractéristiques démographiques et les comportements des personnes, avec pour conséquence fréquente la promotion de contenus sensationnalistes, renforçant potentiellement les tendances à la polarisation[68]. De plus, le ciblage des informations peut être malvenu voire conduire à des intrusions dangereuses dans la vie privée. Les systèmes de recommandation ont, par exemple, permis à des victimes de violence de découvrir que l'auteur de l'agression leur avait été proposé comme ami potentiel par des plateformes de réseaux sociaux, et vice-versa, mettant ainsi la victime en danger. En outre, il a été démontré que le parti pris des groupes majoritaires ou dominants, reflété dans les données provenant des résultats des recherches, influe sur les informations concernant les groupes minoritaires ou vulnérables ou partagées par eux. Par exemple, des recherches ont démontré un niveau inquiétant de préjugés sexistes[69] et raciaux dans les résultats de recherche de Google[70].

IV. Relever les défis

37. Un nombre de plus en plus important d'experts et de parties prenantes ainsi que la communauté internationale sont convaincus de la nécessité d'une approche des nouvelles technologies en général, et de l'intelligence artificielle en particulier, fondée sur les droits de l'homme[71]. Une telle approche offre des outils permettant d'aider les sociétés à identifier les moyens de prévenir et de limiter les dommages tout en maximisant les avantages du progrès technologique.

[65] Voir OTH 71/2018 et OTH 73/2020. Pour une analyse approfondie du filtrage automatique des contenus, voir également https://journals.sagepub.com/doi/full/10.1177/2053951720920686.

[66] A/73/348, par. 17.

[67] Ibid., par. 25, et A/HRC/47/25, par. 36.

[68] Voir www.brookings.edu/techstream/how-youtube-helps-form-homogeneous-online-communities/.

[69] Contributions de l'Autriche et de l'Allemagne.

[70] Safiya Umoja Noble, *Algorithms of Oppression* (New York, New York University Press, 2018).

[71] Résolution 75/176 de l'Assemblée générale, par. 6 ; résolutions 47/16, par. 8 d) et 47/23, seizième alinéa du préambule, A/73/348, par. 47 à 60, A/75/590, par. 57, et A/HRC/43/29 du Conseil des droits de l'homme et contributions de l'Autriche, du Commissaire à la protection de la vie privée du Canada, de Digital Rights Watch, de Global Network Initiative et de Privacy International.

A. Principes fondamentaux

38. Une approche fondée sur les droits de l'homme suppose le respect d'un certain nombre de principes fondamentaux, parmi lesquels l'égalité et la non-discrimination, la participation et la responsabilité, qui sont également au cœur des objectifs de développement durable et des Principes directeurs relatifs aux entreprises et aux droits de l'homme. De plus, les exigences de légalité, de légitimité, de nécessité et de proportionnalité doivent être appliquées de manière cohérente aux technologies de l'IA[72]. Par ailleurs, les nouvelles technologies devraient être déployées de sorte à faciliter la réalisation des droits économiques, sociaux et culturels, en veillant à ce que leurs éléments clefs de disponibilité, d'accessibilité, y compris financière, et de qualité, soient garantis[73]. Les personnes qui subissent des violations des droits de l'homme et des exactions liées à l'utilisation de l'IA doivent avoir accès à des recours judiciaires et non judiciaires utiles[74].

39. Comme souligné plus haut, les restrictions au droit à la vie privée doivent être prévues par la loi, nécessaires pour atteindre un but légitime et proportionnées à celui-ci. En pratique, cela signifie que les États sont tenus de déterminer avec soin si une mesure est susceptible d'atteindre un objectif donné ainsi que l'importance de celui-ci et les incidences de la mesure. Ils doivent également évaluer si des approches moins invasives pourraient permettre d'obtenir les mêmes résultats avec la même efficacité et si c'est le cas, ces mesures doivent être privilégiées. La Haute-Commissaire a déjà souligné la nécessité de ces limites et de ces garanties dans le contexte de la surveillance exercée par les agences de renseignement et les forces de l'ordre[75]. Il convient de noter que les tests de nécessité et de proportionnalité peuvent également conclure à l'abandon de certaines mesures. Par exemple, les exigences de conservation globale et indiscriminée des données de communication imposées aux entreprises de télécommunications et autres ne satisferaient pas au critère de proportionnalité[76]. De même, il serait disproportionné d'imposer des exigences d'identification biométrique aux bénéficiaires de prestations sociales sans prévoir de mesure alternative. En outre, il est crucial que les mesures ne soient pas évaluées isolément, mais que les effets cumulatifs de mesures distinctes mais interactives soient dûment pris en compte. Par exemple, avant de décider de déployer de nouveaux outils de surveillance basés sur l'IA, un État doit faire le point des moyens existants et de leurs effets sur la jouissance du droit à la vie privée et des autres droits.

B. Législation et réglementation

40. La protection effective du droit à la vie privée et des droits connexes dépend des cadres juridiques, réglementaires et institutionnels mis en place par les États[77].

41. L'importance de protections juridiques efficaces en vertu des lois sur la confidentialité des données s'est accrue avec l'émergence des systèmes d'IA axés sur les données. Ces protections devraient répondre aux normes minimales identifiées dans le précédent rapport du Haut-Commissaire sur le droit à la vie privée à l'ère du numérique[78].

[72] A/HRC/43/29, par. 41.

[73] Voir l'analyse détaillée du rôle des nouvelles technologies pour la réalisation des droits économiques, sociaux et culturels dans le document A/HRC/43/29.

[74] Pacte international relatif aux droits civils et politiques, art. 2 3), et Principes directeurs relatifs aux entreprises et aux droits de l'homme, principe 15 c) et pilier III.

[75] A/HRC/39/29, par. 34 à 41.

[76] Ibid., par. 18 ; et Cour de justice de l'Union européenne, *Digital Rights Ireland et autres*, C-293/12 et C-594/12, par. 69. Voir également Cour de justice de l'Union européenne, *Maximilan Schrems c. Commissaire à la protection des données*, C-362/14, par. 94, estimant qu'« une réglementation permettant aux autorités publiques d'accéder de manière généralisée au contenu de communications électroniques doit être considérée comme menaçant l'essence même du droit fondamental au respect de la vie privée ».

[77] A/HRC/39/29, par. 26.

[78] Ibid., par. 28 à 33.

42. Les cadres relatifs à la confidentialité des données doivent tenir compte des nouvelles menaces liées à l'utilisation de l'IA[79]. Par exemple, les lois pourraient imposer des restrictions sur le type de données qui peuvent légalement être extrapolées et/ou utilisées et partagées. Les législateurs devraient également envisager de renforcer les droits individuels, notamment en accordant aux personnes le droit à une explication significative et le droit de s'opposer aux décisions entièrement automatisées qui enfreignent leurs droits[80]. À mesure que les technologies d'IA évolueront, les cadres de protection des données personnelles devront être assortis de garanties renforcées.

43. Les organismes indépendants de surveillance de la confidentialité des données constituent un élément indispensable pour lutter contre la complexité et l'opacité croissantes de l'environnement mondial des données, y compris contre ses vastes asymétries d'information. Ces organismes devraient disposer de pouvoirs d'exécution efficaces et être dotés de ressources suffisantes. Les organisations de la société civile devraient être habilitées à soutenir l'application des lois sur la confidentialité des données, notamment par la mise en place de mécanismes de plainte solides.

44. Au-delà de la législation sur la confidentialité des données, un éventail plus large de lois doit être examiné et éventuellement adopté pour relever les défis que pose l'IA tout en respectant pleinement les droits[81].

45. Compte tenu de la diversité des applications de l'IA, de ses systèmes et de ses utilisations, la réglementation doit être suffisamment spécifique pour aborder les questions propres à chaque secteur et pour adapter les réponses aux risques encourus[82]. Plus le risque pour les droits de l'homme est élevé, plus les exigences juridiques relatives à l'utilisation des technologies de l'IA doivent être strictes. Par conséquent, les secteurs où les enjeux sont particulièrement élevés pour les individus, tels que le maintien de l'ordre, la sécurité nationale[83], la justice pénale, la protection sociale, l'emploi, les soins de santé, l'éducation et le secteur financier, devraient être prioritaires. Une approche de la législation et de la réglementation proportionnée aux risques rendra nécessaire l'interdiction de certaines technologies, applications ou utilisations de l'IA, si elles ont des impacts potentiels ou réels qui ne sont pas justifiés au regard du droit international des droits de l'homme, y compris celles qui échouent aux tests de nécessité et de proportionnalité. En outre, dès lors que certaines utilisations de l'IA sont naturellement incompatibles avec l'interdiction de la discrimination, elles ne devraient pas être autorisées. Par exemple, le contrôle social des individus par les gouvernements[84] ou les systèmes d'IA qui classent les individus en groupes sur la base de motifs discriminatoires prohibés[85] devraient être interdits conformément à ces principes. Pour les systèmes dont l'utilisation présente des risques pour les droits de l'homme lorsqu'ils sont déployés dans certains contextes, les États devront réglementer leur utilisation

[79] Le protocole du Conseil de l'Europe modifiant la Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel, adopté en 2018, constitue notamment une réponse à l'émergence de nouvelles pratiques de traitement des données.

[80] Voir le règlement général sur la protection des données de l'Union européenne, qui énonce de tels droits, et la loi californienne sur le droit à la vie privée, qui autorise l'organisme de réglementation à adopter des règles à cet effet.

[81] Voir Conseil de l'Europe, Recommandation CM/Rec(2020)1 du Comité des Ministres aux États membres sur les impacts des systèmes algorithmiques sur les droits de l'homme.

[82] La proposition de loi sur l'IA de l'Union européenne adopte une telle approche fondée sur les risques. Les contributions de Freedom Online Coalition, de Global Network Initiative et de Global Partners Digital soutiennent les règlementations axées sur les risques.

[83] Dans les documents A/HRC/27/37 et A/HRC/39/29, la Haute-Commissaire a précisé les exigences relatives aux mesures de surveillance prises dans le cadre d'enquêtes criminelles et aux fins de la protection de la sécurité nationale qui devraient guider la législation dans ce domaine.

[84] Contribution de l'Union européenne ; Catelijne Muller, « L'impact du développement de l'intelligence artificielle sur les droits de l'homme, la démocratie et l'État de droit », rapport au Conseil de l'Europe, Comité ad hoc sur l'intelligence artificielle (CAHAI(2020)06-fin), 24 juin 2020, par. 75 et Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO), « Avant-projet de recommandation sur l'éthique de l'intelligence artificielle » (SHS/IGM-AIETHICS/2021/JUN/3, Rev.2), 25 juin 2021, par. 26.

[85] Voir Conseil européen de la protection des données et Contrôleur européen de la protection des données, avis conjoint 5/2021, par. 33.

et leur vente afin de prévenir et d'atténuer les conséquences préjudiciables pour les droits de l'homme[86], tant à l'intérieur qu'à l'extérieur du territoire de l'État. Toute supervision ou prise de décision humaine obligatoire devrait être proscrite lorsque des effets négatifs sur les droits de l'homme sont susceptibles de se produire[87]. Étant donné qu'il peut s'écouler du temps avant que les risques puissent être évalués et éliminés, les États devraient également imposer des moratoires sur l'utilisation de technologies potentiellement à haut risque, telles que la reconnaissance faciale à distance en temps réel, jusqu'à ce qu'il soit garanti que leur utilisation ne peut pas violer les droits de l'homme.

46. Les États devraient également adopter des régimes rigoureux de contrôle des exportations pour le commerce transfrontalier des technologies de surveillance afin d'empêcher la vente de ces technologies lorsqu'il existe un risque qu'elles soient utilisées pour violer les droits de l'homme, notamment en ciblant les défenseurs de ces droits ou les journalistes[88].

47. Compte tenu de la diversité des risques associés aux systèmes d'IA, la nécessité d'une surveillance indépendante et impartiale appropriée de leur développement, déploiement et utilisation s'impose. Ce contrôle peut être exercé par divers organes de contrôle administratif, judiciaire, quasi-judiciaire et/ou parlementaire[89]. Par exemple, outre les autorités chargées de la protection des données, les agences de protection des consommateurs, les organismes de régulation sectorielle, les organismes de lutte contre la discrimination et les institutions nationales pour la promotion et la protection des droits de l'homme devraient aussi faire partie du système de surveillance. De plus, des organismes de régulation intersectorielle chargés de superviser l'utilisation de l'IA peuvent contribuer à définir des normes fondamentales et à assurer la cohérence des politiques et de leur application.

C. Diligence raisonnable en matière de droits de l'homme

48. Les États et les entreprises doivent veiller à ce qu'une politique de diligence raisonnable globale en matière de droits de l'homme soit appliquée lors de l'acquisition, du développement, du déploiement et de l'exploitation des systèmes d'IA, ainsi qu'avant le partage ou l'utilisation des mégadonnées qui ont été collectées sur les personnes[90]. Outre le fait de financer et de diriger de tels processus, les États peuvent également exiger des entreprises qu'elles exercent une diligence raisonnable globale en matière de droits de l'homme, ou les y inciter.

49. Les processus de diligence raisonnable en matière de droits de l'homme ont pour objectif d'identifier, d'évaluer, de prévenir et d'atténuer les incidences négatives sur les droits de l'homme que toute entité peut provoquer ou auxquelles elle peut contribuer ou être directement liée[91]. Lorsque les processus de diligence raisonnable montrent qu'une application dans le domaine de l'IA est incompatible avec les droits de l'homme, en raison de l'absence de moyens suffisants pour en atténuer les effets nocifs, celle-ci devrait être

[86] Contribution de Derechos Digitales.

[87] Voir www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/international-human-rights-law-as-a-framework-for-algorithmic-accountability/1D6D0A456B36BA7512A6AFF17F16E9B6.

[88] A/HRC//41/35, par. 49, et A/HRC/44/24, par. 40. Dans ces rapports, le Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression et la Haute-Commissaire ont également demandé un moratoire sur l'octroi de licences d'exportation pour les technologies de surveillance.

[89] Voir https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64.

[90] Le projet B-Tech du Haut-Commissariat des Nations Unies aux droits de l'homme (HCDH) élabore des orientations sur la mise en œuvre des Principes directeurs relatifs aux entreprises et aux droits de l'homme dans l'industrie technologique, y compris des mesures pour remédier aux incidences de l'utilisation des technologies de l'intelligence artificielle sur les droits de l'homme. Voir www.ohchr.org/EN/Issues/Business/Pages/B-TechProject.aspx.

[91] Pour un aperçu de l'exercice de la diligence raisonnable en matière de droits de l'homme dans le contexte de l'IA, voir https://international-review.icrc.org/articles/ai-humanitarian-action-human-rights-ethics-913, p. 174 à 178.

abandonnée. L'évaluation des répercussions sur les droits de l'homme est un élément essentiel des processus de diligence raisonnable dans ce domaine[92]. La diligence raisonnable doit être exercée tout au long du cycle de vie d'un système d'IA[93]. Une attention particulière doit être accordée aux effets disproportionnés sur les femmes et les filles, les lesbiennes, les gays, les bisexuels, les transgenres et les altersexuels, les personnes handicapées, les personnes appartenant à des minorités, les personnes âgées, les personnes en situation de pauvreté et les autres personnes en situation de vulnérabilité.

50. Des consultations constructives devraient être menées avec les titulaires de droits potentiellement concernés et la société civile, tandis que des experts dotés de compétences interdisciplinaires devraient être impliqués dans les évaluations d'impact, y compris dans l'élaboration et l'évaluation des mesures d'atténuation. Les États et les entreprises devraient surveiller en permanence les incidences des systèmes d'IA qu'ils utilisent afin de vérifier s'ils ont des effets préjudiciables sur les droits de l'homme. Les résultats des études d'impact sur les droits de l'homme, des mesures prises pour prévenir les risques pour ces droits et des consultations publiques devraient eux-mêmes être rendus publics[94].

D. Lien entre l'État et les entreprises

51. Les situations où il existe un lien étroit entre un État et une entreprise technologique nécessitent une attention particulière[95]. L'État est un acteur économique important qui peut façonner la manière dont l'IA est développée et utilisée, au-delà du rôle des États s'agissant des mesures juridiques et politiques. Lorsque les États travaillent avec des développeurs d'IA et des prestataires de services du secteur privé, ils devraient prendre des mesures supplémentaires pour s'assurer qu'elle n'est pas utilisée à des fins incompatibles avec les droits de l'homme. Ces mesures devraient être appliquées à la gestion des entreprises publiques, au financement de la recherche et du développement, au soutien financier et autre apporté par les États aux sociétés d'intelligence artificielle, aux efforts de privatisation et aux pratiques en matière de marchés publics.

52. Lorsque les États agissent en tant qu'acteurs économiques, ils restent les premiers responsables en vertu du droit international des droits de l'homme et doivent s'acquitter de leurs obligations de manière proactive. Dans le même temps, les entreprises sont tenues au respect des droits de l'homme lorsqu'elles collaborent avec les États et devraient chercher des moyens d'honorer leurs obligations en la matière lorsqu'elles sont confrontées à des exigences qui sont en contradiction avec le droit des droits de l'homme[96]. Par exemple, lorsqu'elles doivent répondre à des demandes d'accès à des données personnelles qui ne respectent pas les normes dans ce domaine, elles doivent user de leur poids pour résister ou atténuer le préjudice qui pourrait être causé[97].

53. Les États peuvent renforcer la protection des droits de l'homme en exigeant systématiquement un comportement responsable des entreprises. Par exemple, lorsque des organismes de crédit à l'exportation proposent un appui à des sociétés d'intelligence artificielle, ils doivent s'assurer qu'elles ont des antécédents solides en matière de respect des droits et qu'elles peuvent le démontrer en mettant en œuvre des processus de diligence raisonnable rigoureux.

54. Lorsque les États font appel à ces sociétés pour fournir des biens ou des services publics, ils doivent s'assurer qu'ils peuvent superviser le développement et le déploiement des systèmes d'IA, notamment en demandant des informations sur le niveau de précision de l'application et de risque associé et en l'évaluant. S'il est impossible d'atténuer les risques

[92] Pour un résumé concis des méthodes d'évaluation de l'impact sur les droits de l'homme, voir https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64.

[93] A/HRC/43/29, par. 62 g), et A/HRC/44/24, par. 38, 53 j) i) et 54 c).

[94] A/73/348, par. 68.

[95] Voir www.ohchr.org/Documents/Issues/Business/B-Tech/b-tech-foundational-paper-state-duty-to-protect.pdf.

[96] Principes directeurs relatifs aux entreprises et aux droits de l'homme, principe 23 b).

[97] A/HRC/32/38, par. 58. Voir également www.ohchr.org/Documents/Issues/Terrorism/biometricsreport.pdf, p. 39 et 40.

de manière satisfaisante, les États devraient renoncer à l'utiliser pour fournir des biens ou des services publics.

E. Transparence

55. Les développeurs, les spécialistes du marketing, les opérateurs et les utilisateurs de systèmes d'IA devraient intensifier considérablement leurs efforts en matière de transparence concernant l'utilisation de l'IA. Dans un premier temps, les États, les entreprises et autres utilisateurs devraient mettre à disposition des informations sur le type de systèmes qu'ils utilisent et à quelles fins, ainsi que sur l'identité du développeur et de l'opérateur des systèmes[98]. Les personnes concernées devraient être systématiquement informées en cas de décisions exclusivement automatisées ou prises à l'aide d'un algorithme[99]. Les personnes devraient également être informées lorsque les données personnelles qu'elles fournissent sont intégrées à un ensemble de données utilisé par un système d'IA[100]. Par ailleurs, en ce qui concerne les applications d'importance cruciale sur le plan des droits de l'homme, les États devraient envisager d'établir des registres contenant des informations clefs sur ces outils et leur utilisation[101]. Il convient de veiller au respect des obligations de transparence et des droits d'accès aux données ainsi que d'effacement et de rectification de ces dernières, prévus par les cadres juridiques de protection des données personnelles. Une attention particulière devrait être accordée à la possibilité pour les individus de mieux comprendre et contrôler les profils établis à leur sujet[102].

56. Il faudrait promouvoir encore plus la transparence en déployant des efforts soutenus pour surmonter le problème de la « boîte noire » décrit plus haut. Le développement et le déploiement systématiques de méthodologies visant à rendre les systèmes d'IA plus faciles à expliquer − ce que l'on appelle souvent la transparence algorithmique − sont de la plus haute importance pour garantir une protection adéquate des droits[103]. Ceci est particulièrement essentiel lorsque l'intelligence artificielle tranche des questions cruciales d'ordre judiciaire ou relatives à des services sociaux essentiels à la réalisation des droits économiques, sociaux et culturels. Les chercheurs ont déjà mis au point une série d'approches qui vont dans ce sens[104], et il est indispensable d'accroître les investissements dans ce domaine. Les États devraient également prendre des mesures pour s'assurer que les protections de la propriété intellectuelle n'empêchent pas un examen approfondi des systèmes d'IA qui ont des répercussions sur les droits de l'homme[105]. Les règles de passation de marchés devraient être mises à jour pour tenir compte du besoin de transparence, y compris l'auditabilité des systèmes d'IA[106]. En particulier, les États devraient envisager d'éviter l'utilisation de systèmes susceptibles d'avoir des effets matériels néfastes sur les droits de l'homme, lorsque ces systèmes ne peuvent faire l'objet d'audits approfondis[107].

[98] A/HRC//43/29, par. 52, et A/73/348, par. 49.

[99] Conseil de l'Europe, « Lignes directrices sur le traitement des impacts des systèmes algorithmiques sur les droits de l'homme », (annexe à la recommandation CM/Rec(2020)1 du Comité des Ministres aux États membres sur les impacts des systèmes algorithmiques sur les droits de l'homme), sect. B, par. 4.2.

[100] A/73/348, par. 49.

[101] A/HRC/43/29, par. 52. La proposition de l'Union européenne pour une loi sur l'IA contient des dispositions sur la création d'un registre des systèmes d'IA à haut risque.

[102] Voir https://link.springer.com/article/10.1007/s12394-008-0003-1, p. 67.

[103] Pour un aperçu des éléments qui garantissent la transparence des algorithmes, voir www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/international-human-rights-law-as-a-framework-for-algorithmic-accountability/1D6D0A456B36BA7512A6AFF17F16E9B6, p. 320 à 323.

[104] Voir https://arxiv.org/abs/2001.00973 et https://arxiv.org/pdf/1711.01134.pdf.

[105] Conseil de l'Europe, « Lignes directrices sur le traitement des impacts des systèmes algorithmiques sur les droits de l'homme », (annexe à la recommandation CM/Rec(2020)1 du Comité des Ministres aux États membres sur les impacts des systèmes algorithmiques sur les droits de l'homme), sect. B, par. 4.1.

[106] Voir les contributions de l'Allemagne, Derechos Digitales, Freedom Online Coalition et Global Partners Digital.

[107] A/73/348, par. 55, et A/HRC/43/29, par. 54.

V. Conclusions et recommandations

A. Conclusions

57. **Le présent rapport met en évidence les effets indéniables et en constante augmentation des technologies de l'IA sur l'exercice du droit à la protection de la vie privée et à d'autres droits de l'homme, pour le meilleur comme pour le pire. Il met en évidence des évolutions inquiétantes, notamment un écosystème tentaculaire de collecte et d'échange de données personnelles, en grande partie opaque, qui sous-tend certaines parties des systèmes d'IA largement utilisées. Ces systèmes affectent les approches des pouvoirs publics en matière de maintien de l'ordre et d'administration de la justice, déterminent qui peut bénéficier de services publics, décident de qui a une chance d'être recruté pour un emploi et ont une influence sur l'information que les gens peuvent voir et partager en ligne. En outre, le risque de discrimination lié à des décisions basées sur l'intelligence artificielle n'est que trop réel. Le rapport présente différentes manières de résoudre les problèmes fondamentaux liés à l'IA, soulignant que seule une approche globale fondée sur les droits de l'homme peut garantir des solutions durables pour le bien de tous.**

58. **Néanmoins, étant donné la diversité des nouvelles questions qui se posent dans le contexte de l'IA, le présent rapport n'est qu'un instantané d'une situation en constante évolution à cet égard. Les domaines qui méritent une analyse plus approfondie sont la santé, l'éducation, le logement et les services financiers. Alors que les technologies biométriques s'imposent de plus en plus comme une solution de choix pour les États, les organisations internationales et les entreprises technologiques, il est urgent en l'espèce de fournir davantage de conseils en matière de droits de l'homme. En outre, les travaux futurs qui seront entrepris sous l'angle de ces droits devraient être axés sur la recherche de moyens permettant de combler les énormes lacunes de responsabilité qui existent dans l'environnement mondial des données. Enfin, il est urgent d'identifier et de mettre en œuvre des solutions pour surmonter les discriminations liées à l'IA.**

B. Recommandations

59. **La Haute-Commissaire recommande aux États :**

a) De reconnaître pleinement en tant qu'objectif central la nécessité de protéger et de renforcer tous les droits de l'homme dans le cadre de la mise au point, de l'utilisation et de la gouvernance de l'intelligence artificielle, et de garantir le respect et la réalisation de tous les droits de l'homme, aussi bien en ligne qu'hors ligne ;

b) De veiller à ce que l'utilisation de l'IA respecte tous les droits de l'homme et à ce que toute immixtion dans le droit à la vie privée et les autres droits de l'homme au moyen de cette technologie soit encadrée par la loi, poursuive un objectif légitime, soit conforme aux principes de nécessité et de proportionnalité et ne porte pas atteinte à l'essence même des droits en question ;

c) D'interdire expressément les applications de l'IA qui ne peuvent être exploitées dans le respect du droit international des droits de l'homme et d'imposer des moratoires sur la vente et l'utilisation des systèmes d'IA qui présentent un risque élevé pour la jouissance des droits de l'homme, à moins et jusqu'à ce que des garanties appropriées pour les protéger soient mises en place ;

d) D'imposer un moratoire sur l'utilisation de la reconnaissance faciale dans les espaces publics, au moins jusqu'à ce que les autorités compétentes puissent démontrer le respect des normes de protection de la vie privée et des données, ainsi que l'absence de problèmes de fiabilité et d'effets discriminatoires notables, et jusqu'à ce que toutes les recommandations énoncées à l'alinéa j) i-v) du paragraphe 53 du document A/HRC/44/24 soient mises en œuvre ;

e) D'adopter et d'appliquer effectivement, par l'intermédiaire d'autorités indépendantes et impartiales, une législation sur la confidentialité des données pour les

secteurs public et privé, en tant que condition préalable essentielle à la protection du droit à la vie privée dans le contexte de l'IA ;

f) D'adopter des cadres législatifs et réglementaires qui permettent de prévenir et d'atténuer comme il convient les effets négatifs multiformes sur les droits de l'homme de l'utilisation de l'IA par les secteurs public et privé ;

g) De veiller à ce que les victimes de violations des droits de l'homme et d'atteintes à ceux-ci liées à l'utilisation de systèmes d'IA aient accès à des voies de recours utiles ;

h) D'exiger une explicabilité appropriée de toutes les décisions prises à l'aide de l'IA et susceptibles d'affecter de manière significative les droits de l'homme, en particulier dans le secteur public ;

i) De renforcer les efforts faits pour lutter contre la discrimination liée à l'utilisation des systèmes d'IA par les États et les entreprises, notamment en menant, en exigeant et en soutenant des évaluations et un suivi systématiques des résultats fournis par les systèmes d'IA et des conséquences de leur déploiement ;

j) De veiller à ce que les partenariats public-privé soient transparents et soumis à une surveillance indépendante des droits de l'homme dans la fourniture et l'utilisation des technologies d'IA, et qu'ils n'entraînent pas, de la part des gouvernements, une abdication de leur responsabilité en matière de droits de l'homme.

60. **La Haute-Commissaire recommande aux États et aux entreprises :**

a) D'exercer systématiquement une diligence raisonnable en matière de droits de l'homme tout au long du cycle de vie des systèmes d'IA qu'elles conçoivent, développent, déploient, vendent, acquièrent ou exploitent. De procéder à des évaluations régulières et complètes des incidences sur les droits de l'homme en tant qu'élément clef de l'exercice de leur diligence raisonnable en matière de droits de l'homme ;

b) D'accroître considérablement la transparence de leur utilisation de l'IA, notamment en informant comme il se doit le public et les personnes concernées et en permettant un audit indépendant et externe des systèmes automatisés. Plus les impacts potentiels ou réels sur les droits de l'homme liés à l'utilisation de l'IA sont probables et graves, plus la transparence est nécessaire ;

c) De veiller à ce que toutes les parties prenantes participent aux décisions concernant le développement, le déploiement et l'utilisation de l'IA, en particulier les personnes et groupes concernés ;

d) D'améliorer l'explicabilité des décisions prises à l'aide de l'IA, notamment en finançant et en menant des recherches à cette fin.

61. **La Haute-Commissaire recommande aux entreprises :**

a) De tout mettre en œuvre pour s'acquitter de leur responsabilité de respecter tous les droits de l'homme, notamment en rendant pleinement opérationnels les principes directeurs relatifs aux entreprises et aux droits de l'homme ;

b) D'intensifier les efforts faits pour lutter contre la discrimination liée au développement, à la vente ou au fonctionnement des systèmes d'IA, notamment en procédant à des évaluations et à un suivi systématiques des résultats produits par ces systèmes et des conséquences de leur déploiement ;

c) De prendre des mesures décisives afin de garantir la diversité de la main-d'œuvre chargée du développement de l'IA ;

d) De prévoir des mesures de réparation ou de collaborer à leur mise en œuvre dans le cadre de processus légitimes lorsqu'elles ont eu des incidences négatives ou qu'elles y ont contribué, notamment grâce à des mécanismes de réclamation efficaces au niveau opérationnel.