



## Groupe d'experts ad hoc (GEAH) pour l'élaboration d'un avant-projet de recommandation sur l'éthique de l'intelligence artificielle

Distribution limitée

SHS/BIO/AHEG-AI/2020/4 REV.2  
Paris, le 7 septembre 2020  
Original anglais

### DOCUMENT FINAL

#### AVANT-PROJET DE RECOMMANDATION SUR L'ÉTHIQUE DE L'INTELLIGENCE ARTIFICIELLE

Conformément à la décision adoptée par la Conférence générale de l'UNESCO à sa 40<sup>e</sup> session ([résolution 40 C/37](#)), la Directrice générale a constitué, en mars 2020, un Groupe d'experts ad hoc (GEAH) chargé d'élaborer un projet de recommandation sur l'éthique de l'intelligence artificielle.

Compte tenu de la situation difficile liée à la pandémie de COVID-19, le GEAH a conduit ses travaux à distance de fin mars à début mai 2020, aboutissant à la production de la première version d'un projet de recommandation sur l'éthique de l'intelligence artificielle.

Un vaste processus de consultation multipartite sur cette première version a été menée de juin à août 2020, sur la base de trois éléments : (i) une consultation publique en ligne (qui a recueilli plus de 800 réponses) ; (ii) des consultations régionales et sous-régionales coorganisées avec les pays/institutions hôtes et menées par voie électronique dans toutes les régions de l'UNESCO (qui ont impliqué plus de 500 participants) ; (iii) des ateliers ouverts, organisés par des partenaires, permettant à différentes parties prenantes et aux citoyens de discuter (qui ont réuni environ 500 participants). Le processus de consultation a donné lieu à plus de 50 000 commentaires sur le texte.

En tenant compte des commentaires reçus au cours de ce processus de consultation, le GEAH a révisé la première version du projet de texte d'août à début septembre 2020 afin d'élaborer l'avant-projet de recommandation contenu dans le présent document, lequel sera transmis aux États membres en septembre 2020 pour qu'ils communiquent leurs observations par écrit.

Le GEAH a bénéficié du soutien de la Sous-Directrice générale pour les sciences sociales et humaines et de la Section de la bioéthique et de l'éthique des sciences.

Le présent document ne prétend pas être exhaustif et ne représente pas nécessairement les opinions des États membres de l'UNESCO.

## AVANT-PROJET DE RECOMMANDATION SUR L'ÉTHIQUE DE L'INTELLIGENCE ARTIFICIELLE

### PRÉAMBULE

La Conférence générale de l'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO), réunie à Paris du xx au xx, à l'occasion de sa xx session,

**Reconnaissant** les répercussions profondes et dynamiques de l'intelligence artificielle (IA) sur les sociétés, les écosystèmes et la vie humaine, y compris l'esprit humain, en raison notamment des nouvelles façons dont elle agit sur la pensée, les interactions et la prise de décisions des êtres humains et dont elle retentit sur l'éducation, les sciences sociales et humaines et les sciences exactes et naturelles, la culture et la communication et l'information,

**Rappelant** que l'UNESCO se propose, aux termes de son Acte constitutif, de contribuer à la paix et à la sécurité et de resserrer, par l'éducation, les sciences, la culture et la communication et l'information, la collaboration entre nations, afin d'assurer le respect universel de la justice, de la loi, des droits de l'homme et des libertés fondamentales reconnus à tous les peuples,

**Convaincue** que l'instrument normatif présenté ici, fondé sur le droit international et sur une approche normative globale, et centré sur la dignité humaine et les droits de l'homme, ainsi que sur l'égalité des genres, la justice sociale et économique, le bien-être physique et mental, la diversité, l'interdépendance, l'inclusion et la protection de l'environnement et des écosystèmes, peut donner une orientation responsable aux technologies de l'IA,

**Considérant** que les technologies de l'IA peuvent rendre de grands services à l'humanité, mais qu'elles soulèvent également des préoccupations éthiques de fond, à l'égard, par exemple, des biais qu'elles sont susceptibles de comporter et d'accentuer, lesquels pourraient entraîner inégalité et exclusion, menacer la diversité culturelle, sociale et écologique et entraîner des clivages sociaux ou économiques ; la nécessité d'assurer la transparence et l'intelligibilité du fonctionnement des algorithmes et des données à partir desquelles ils ont été entraînés ; et leurs éventuelles conséquences sur la dignité humaine, les droits de l'homme, l'égalité des genres, la vie privée, la liberté d'expression, l'accès à l'information, les processus sociaux, économiques, politiques et culturels, les pratiques scientifiques et d'ingénierie, le bien-être animal, et l'environnement et les écosystèmes,

**Ayant à l'esprit** que les technologies de l'IA pourraient creuser les écarts et les inégalités qui existent dans le monde, au sein des pays et entre eux, et qu'il faut maintenir la justice, la confiance et l'équité pour que personne ne soit laissé de côté, qu'il s'agisse de profiter des avantages des technologies de l'IA ou de se prémunir contre leurs conséquences négatives, tout en reconnaissant les différences de situation qui prévalent entre les pays et le souhait de certaines personnes de ne pas prendre part à toutes les innovations technologiques,

**Consciente** que tous les pays connaissent une accélération de l'utilisation des technologies de l'information et de la communication et des technologies de l'IA, ainsi qu'un besoin croissant en matière d'éducation aux médias et à l'information, et que l'économie numérique représente des défis et des possibilités immenses aux niveaux sociétal, économique et environnemental en termes de partage des bienfaits, en particulier pour les pays à revenu faible et intermédiaire, y compris, sans s'y limiter, les pays les moins avancés (PMA), les pays en développement sans littoral (PDSL) et les petits États insulaires en développement (PEID), ce qui exige de prendre en compte, de protéger et de promouvoir les cultures, valeurs et connaissances endogènes pour développer des économies numériques durables,

**Reconnaissant** que les technologies de l'IA pourraient avoir des effets bénéfiques sur l'environnement et les écosystèmes, mais que la matérialisation de ces avantages exige

d'assurer un accès équitable à ces technologies sans ignorer mais au contraire en tenant compte des dommages et des impacts qu'elles peuvent avoir sur l'environnement et les écosystèmes,

**Notant** que la prise en compte des risques et des préoccupations éthiques ne devrait pas freiner l'innovation, mais plutôt offrir de nouvelles possibilités et encourager de nouvelles pratiques de recherche et d'innovation responsables plaçant les droits de l'homme, les valeurs, les principes et la réflexion morale et éthique au fondement des technologies de l'IA,

**Rappelant** que la Conférence générale de l'UNESCO, à sa 40<sup>e</sup> session en novembre 2019, a adopté la résolution 40 C/37, par laquelle elle a chargé la Directrice générale « d'élaborer un instrument normatif international sur l'éthique de l'intelligence artificielle, sous la forme d'une recommandation », qui doit lui être présenté à sa 41<sup>e</sup> session en 2021,

**Reconnaissant** que l'essor des technologies de l'IA entraîne une multiplication des informations qui nécessite de développer l'éducation aux médias et à l'information en conséquence, de même que l'accès à des sources d'information critiques,

**Constatant** que le cadre normatif applicable aux technologies de l'IA et à ses implications sociales trouve son fondement dans l'éthique, les droits de l'homme, les libertés fondamentales, l'accès aux données, aux informations et aux connaissances, les cadres juridiques internationaux et nationaux, la liberté de recherche et d'innovation, le bien-être de l'humanité, de l'environnement et des écosystèmes, et établit un lien entre les valeurs et principes éthiques et les défis et possibilités liés aux technologies de l'IA, sur la base d'une compréhension commune et d'objectifs partagés,

**Reconnaissant** que les valeurs et principes éthiques peuvent profondément influencer la définition et l'application de mesures politiques et de normes juridiques fondées sur les droits, en fournissant des orientations lorsque le champ d'application des normes n'est pas clairement circonscrit, ou lorsque ces normes ne sont pas encore établies en raison de la rapidité du progrès technologique et de la relative lenteur des réponses politiques,

**Convaincue** que des normes éthiques mondialement reconnues pour les technologies de l'IA et le droit international, en particulier les textes, principes et normes des droits de l'homme, peuvent jouer un rôle essentiel pour harmoniser les normes juridiques relatives à l'IA dans le monde,

**Reconnaissant** la Déclaration universelle des droits de l'homme (1948), notamment l'article 27 qui énonce le droit de participer au progrès scientifique et aux bienfaits qui en résultent ; les instruments internationaux relatifs aux droits de l'homme, dont la Convention internationale sur l'élimination de toutes les formes de discrimination raciale (1965), le Pacte international relatif aux droits civils et politiques (1966), le Pacte international relatif aux droits économiques, sociaux et culturels (1966), la Convention des Nations Unies sur l'élimination de toutes les formes de discrimination à l'égard des femmes (1979), la Convention des Nations Unies relative aux droits de l'enfant (1989), la Convention des Nations Unies relative aux droits des personnes handicapées (2006), et la Convention de l'UNESCO sur la protection et la promotion de la diversité des expressions culturelles (2005),

**Prenant acte** de la Déclaration de l'UNESCO sur les responsabilités des générations présentes envers les générations futures (1997) ; de la Déclaration des Nations Unies sur les droits des peuples autochtones (2007) ; du Rapport du Secrétaire général de l'ONU de 2011 sur la suite donnée à la deuxième Assemblée mondiale sur le vieillissement (A/66/173), qui traite de la situation des droits des personnes âgées ; du Rapport du Représentant spécial du Secrétaire général de l'ONU chargé de la question des droits de l'homme et des sociétés transnationales et autres entreprises de 2011 (A/HRC/17/31), qui présente les « Principes

directeurs relatifs aux entreprises et aux droits de l'homme : mise en œuvre du cadre de référence « protéger, respecter et réparer » des Nations Unies » ; de la résolution de l'Assemblée générale des Nations Unies sur l'examen des textes issus du Sommet mondial sur la société de l'information (A/68/302) ; de la résolution du Conseil des droits de l'homme sur « Le droit à la vie privée à l'ère du numérique » (A/HRC/RES/42/15) adoptée le 26 septembre 2019 ; du rapport du Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression (A/73/348) ; de la Recommandation de l'UNESCO concernant la science et les chercheurs scientifiques (2017) ; des indicateurs de l'UNESCO sur l'universalité de l'Internet (approuvés par le Programme international de l'UNESCO pour le développement de la communication en 2019), incluant les principes ROAM (approuvés par la Conférence générale de l'UNESCO en 2015) ; de la Recommandation de l'UNESCO concernant la préservation et l'accessibilité du patrimoine documentaire, y compris le patrimoine numérique (2015) ; du rapport du Groupe de haut niveau du Secrétaire général de l'ONU sur la coopération numérique, intitulé « L'ère de l'interdépendance numérique » (2019) ; du Plan d'action de coopération numérique du Secrétaire général de l'Organisation des Nations Unies (2020) ; de la Déclaration universelle sur la bioéthique et les droits de l'homme (2005) ; de la Déclaration de principes éthiques en rapport avec le changement climatique (2017) de l'UNESCO ; de l'Initiative Global Pulse des Nations Unies ; et des résultats et rapports des sommets mondiaux de l'Union internationale des télécommunications sur l'intelligence artificielle au service de l'intérêt général,

**Prenant acte également** des cadres relatifs à l'éthique de l'IA établis par d'autres organisations intergouvernementales, tels que les instruments des droits de l'homme et autres textes juridiques pertinents adoptés par le Conseil de l'Europe, y compris les travaux de son Comité ad hoc sur l'intelligence artificielle (CAHAI) ; les travaux de l'Union européenne traitant de l'IA et ceux du Groupe d'experts de haut niveau de la Commission européenne sur l'IA, notamment les Lignes directrices en matière d'éthique pour une IA digne de confiance ; les travaux du premier Groupe d'experts de l'Organisation de coopération et de développement économiques (OCDE) (AIGO) et de son successeur, le réseau d'experts de l'OCDE sur l'IA (ONE AI), la Recommandation du Conseil de l'OCDE sur l'IA et l'Observatoire des politiques de l'IA de l'OCDE (OECD.AI) ; les Principes sur l'IA du G20, inspirés des travaux de l'OCDE et présentés dans la Déclaration ministérielle du G20 sur le commerce et l'économie numérique ; la Vision commune de Charlevoix sur l'avenir de l'IA adoptée par le G7 ; les travaux du Groupe de travail de l'Union africaine sur l'IA ; et les travaux du Groupe de travail de la Ligue des États arabes sur l'IA,

**Soulignant** qu'il est nécessaire de prêter une attention particulière aux pays à revenu faible et intermédiaire, y compris, sans s'y limiter, aux pays les moins avancés (PMA), aux pays en développement sans littoral (PDSL) et aux petits États insulaires en développement (PEID), car ils disposent de capacités mais sont sous-représentés dans le débat sur l'éthique de l'IA, ce qui soulève des préoccupations quant à une prise en compte insuffisante des savoirs locaux, du pluralisme culturel et éthique, des systèmes de valeurs et des exigences d'équité mondiale dans la gestion des incidences positives et négatives des technologies de l'IA,

**Consciente** qu'il existe de nombreuses politiques et autres cadres éthiques et réglementaires relatifs aux technologies de l'IA à l'échelle nationale,

**Consciente également** qu'il existe nombre d'initiatives et de cadres relatifs à l'IA émanant du secteur privé, des organisations professionnelles et des organisations non gouvernementales, notamment l'Initiative mondiale de l'Institut des ingénieurs électriciens et électroniciens (IEEE) sur l'éthique des systèmes autonomes et intelligents et ses travaux sur une conception conforme à l'éthique ; le livre blanc du Forum économique mondial sur une gouvernance multipartite mondiale des technologies ; les « 10 grands principes pour une intelligence artificielle éthique » de l'UNI Global Union ; la Déclaration de Montréal pour un développement responsable de l'IA ; la Déclaration de Toronto – Protéger le droit à l'égalité et à la non-

discrimination dans les systèmes reposant sur l'apprentissage automatique ; les principes pour une intelligence artificielle harmonieuse ; et les principes du Partenariat sur l'IA,

**Convaincue** que les technologies de l'IA peuvent être porteuses d'importants avantages, mais que leur matérialisation pourrait amplifier les tensions en lien avec une dette d'innovation, un accès asymétrique aux connaissances, des limitations du droit à l'information, des écarts en matière de capacité de créativité, de cycles de développement et de capacités humaines et institutionnelles, des restrictions de l'accès à l'innovation technologique, et un manque d'infrastructures et de cadres réglementaires physiques et numériques adéquats concernant les données,

**Soulignant** que la coopération et la solidarité mondiales sont nécessaires pour aborder les défis posés par les technologies de l'IA dans des cultures et systèmes éthiques divers et interdépendants, pour réduire les risques d'utilisation abusive et pour garantir que les stratégies et cadres réglementaires relatifs à l'IA ne sont pas uniquement guidés par des intérêts nationaux et commerciaux et par la concurrence économique,

**Tenant pleinement compte** du fait que le développement rapide des technologies de l'IA rend difficiles une mise en œuvre et une gouvernance éthiques, en raison de la diversité des orientations éthiques et des cultures du monde, du manque de souplesse de la législation concernant la technologie et les sociétés du savoir, et du risque que les technologies de l'IA perturbent les normes et valeurs éthiques locales et régionales,

1. **Adopte** la présente Recommandation sur l'éthique de l'intelligence artificielle ;
2. **Recommande** aux États membres d'appliquer les dispositions de la présente Recommandation en prenant des mesures appropriées, notamment législatives, conformes aux pratiques constitutionnelles et aux structures de gouvernance de chaque État, en vue de donner effet, dans leurs juridictions, aux principes et normes énoncés dans la Recommandation conformément au droit international et aux pratiques constitutionnelles ;
3. **Recommande également** aux États membres de veiller à ce que toutes les parties prenantes assument leurs responsabilités, y compris les entreprises du secteur privé travaillant dans le domaine des technologies de l'IA, et de porter la présente Recommandation à la connaissance des autorités, organismes, instituts de recherche et organismes universitaires, et institutions et organisations des secteurs public, privé et de la société civile engagés dans les technologies de l'IA, de sorte de garantir que le développement et l'utilisation des technologies de l'IA sont guidés par des recherches scientifiques de qualité ainsi que par une analyse et une évaluation éthiques.

## I. CHAMP D'APPLICATION

1. La présente Recommandation traite des questions éthiques soulevées par l'intelligence artificielle. Elle aborde l'éthique de l'IA en tant que réflexion normative systématique, basée sur un cadre global et évolutif de valeurs, de principes et d'actions interdépendants de nature à orienter les sociétés pour qu'elles apportent des réponses responsables aux conséquences connues et inconnues des impacts des technologies de l'IA sur les êtres humains, les sociétés, l'environnement et les écosystèmes, et leur offre une base pour accepter ou rejeter les technologies de l'IA. Plutôt que d'assimiler l'éthique au droit, aux droits de l'homme ou à un complément normatif aux technologies, elle considère l'éthique comme une base dynamique pour l'évaluation et l'orientation normatives des technologies de l'IA, en faisant de la dignité humaine, du bien-être et de la prévention des dommages des repères et en s'appuyant sur l'éthique des sciences et des technologies.

2. La présente Recommandation ne cherche pas à donner de définition unique de l'IA, celle-ci étant appelée à évoluer en fonction des progrès technologiques. Son objectif est plutôt de traiter des caractéristiques des systèmes d'IA qui revêtent une importance majeure sur le plan éthique et font l'objet d'un vaste consensus international. En conséquence, la présente Recommandation envisage les systèmes d'IA comme des systèmes technologiques capables de traiter l'information par un processus s'apparentant à un comportement intelligent, et comportant généralement des fonctions de raisonnement, d'apprentissage, de perception, d'anticipation, de planification ou de contrôle. Trois éléments occupent une place centrale dans cette approche :

- (a) les systèmes d'IA sont des technologies de traitement des informations qui intègrent des modèles et des algorithmes, lesquels génèrent une capacité d'apprentissage et d'exécution de tâches cognitives conduisant à des résultats tels que l'anticipation et la prise de décisions dans des environnements réels et virtuels. Les systèmes d'IA sont conçus pour fonctionner avec une certaine autonomie, au moyen de la modélisation et la représentation des connaissances, de l'exploitation des données et du calcul de corrélations. Ils peuvent intégrer plusieurs méthodes, telles que, sans s'y limiter :
  - (i) l'apprentissage automatique, y compris l'apprentissage profond et l'apprentissage par renforcement ;
  - (ii) le raisonnement automatique, y compris la planification, la programmation, la représentation des connaissances et le raisonnement, et la recherche et l'optimisation ;
  - (iii) des systèmes cyberphysiques, y compris l'Internet des objets, les systèmes robotiques, la robotique sociale et les interfaces homme-machine qui impliquent des fonctions de contrôle, de perception et de traitement de données recueillies par des senseurs, et le fonctionnement d'actionneurs dans l'environnement du système d'IA ;
- (b) les questions éthiques concernant les systèmes d'IA se rapportent à toutes les étapes du cycle de vie de ces systèmes, compris ici comme allant de la recherche, la conception et le développement au déploiement et à l'utilisation, et incluant la maintenance, l'exploitation, la commercialisation, le financement, le suivi et l'évaluation, la validation, la fin de l'utilisation, le démontage et la mise hors service. En outre, les acteurs de l'IA peuvent désigner tout acteur impliqué dans au moins une étape du cycle de vie de l'IA, et peuvent renvoyer à des personnes physiques comme à des personnes morales, telles que les chercheurs, les programmeurs, les ingénieurs, les spécialistes des données, les utilisateurs finaux, les grandes entreprises spécialisées dans la technologie, les petites et moyennes entreprises, les start-up, les universités et les entités publiques, entre autres ;
- (c) les systèmes d'IA soulèvent de nouveaux types de questions éthiques qui comprennent, sans s'y limiter, leur impact sur la prise de décisions, l'emploi et le travail, les interactions sociales, les soins de santé, l'éducation, les médias, la liberté d'expression, l'accès à l'information, la protection de la vie privée, la démocratie, la discrimination et la militarisation. En outre, de nouveaux défis éthiques sont générés par la reproduction potentielle des biais par les algorithmes d'IA, par exemple concernant le genre, l'appartenance ethnique et l'âge, ce qui exacerbe des formes déjà existantes de discriminations, de préjugés liés à l'identité et de stéréotypes. Certaines de ces questions ont trait à la capacité des systèmes d'IA d'effectuer des tâches qu'auparavant seuls des êtres vivants pouvaient réaliser, parfois même uniquement des êtres humains. Ces

caractéristiques confèrent aux systèmes d'IA un rôle déterminant et inédit dans les pratiques et les sociétés humaines, ainsi que dans leur relation avec l'environnement et les écosystèmes, créant un nouveau contexte dans lequel les enfants et les jeunes vont grandir, développer une compréhension du monde et d'eux-mêmes, comprendre de manière critique les médias et l'information, et apprendre à prendre des décisions. Les systèmes d'IA pourraient, à long terme, disputer aux êtres humains le sentiment d'expérience et la capacité d'agir qui leur sont propres, ce qui susciterait de nouvelles inquiétudes quant à la compréhension qu'ont les êtres humains d'eux-mêmes, leurs interactions sociales, culturelles et environnementales, leur autonomie, leur capacité d'agir, leur valeur et leur dignité.

3. La présente Recommandation prête une attention particulière aux implications éthiques plus larges des systèmes d'IA dans les domaines centraux de l'UNESCO : éducation, science, culture et communication et information, lesquelles font l'objet de l'étude préliminaire sur l'éthique de l'intelligence artificielle réalisée par la Commission mondiale d'éthique des connaissances scientifiques et des technologies (COMEST) de l'UNESCO en 2019 :

- (a) l'éducation, parce que vivre dans des sociétés en cours de numérisation exige de nouvelles pratiques éducatives, une réflexion éthique, une pensée critique, des pratiques de conception responsables et de nouvelles compétences, au vu des conséquences sur le marché de l'emploi et l'employabilité ;
- (b) la science, au sens le plus large et incluant tous les domaines d'étude, des sciences exactes et naturelles et des sciences médicales aux sciences sociales et humaines, car les technologies de l'IA apportent de nouvelles capacités de recherche, influencent nos conceptions de la compréhension et de l'explication scientifiques, et créent une nouvelle base pour la prise de décisions ;
- (c) l'identité et la diversité culturelles, car les technologies de l'IA peuvent enrichir les industries culturelles et créatives, mais aussi aboutir à une concentration accrue de l'offre, des données, des marchés et des revenus de la culture entre les mains d'un petit nombre d'acteurs, avec des répercussions potentiellement négatives sur la diversité et le pluralisme des langues, des médias, des expressions culturelles, la participation et l'égalité ;
- (d) la communication et l'information, car les technologies de l'IA jouent un rôle croissant dans le traitement, la structuration et la transmission de l'information, et les questions du journalisme automatisé et de l'utilisation d'algorithmes pour diffuser des actualités et pour modérer et organiser des contenus sur les réseaux sociaux et les moteurs de recherche ne sont que quelques exemples de problématiques liées à l'accès à l'information, à la désinformation, à la mésinformation, à l'incompréhension, à l'émergence de nouvelles formes de récits sociétaux, à la discrimination, à la liberté d'expression, à la protection de la vie privée et à l'éducation aux médias et à l'information, entre autres.

4. La présente Recommandation s'adresse aux États, à la fois en tant qu'acteurs de l'IA et en tant que responsables de l'élaboration de cadres juridiques et réglementaires tout au long du cycle de vie des systèmes d'IA, et de la promotion de la responsabilité des entreprises. Elle fournit également des orientations éthiques à tous les acteurs de l'IA, y compris du secteur privé, en constituant une base pour une évaluation des incidences éthiques des systèmes d'IA tout au long de leur cycle de vie.

## II. BUTS ET OBJECTIFS

5. La présente Recommandation a pour objet de servir de base afin de mettre les systèmes d'IA au service de l'humanité, des individus, des sociétés, de l'environnement et des écosystèmes, ainsi que de prévenir les préjudices.

6. En plus des cadres éthiques concernant l'IA qui ont déjà été élaborés par diverses organisations à travers le monde, la présente Recommandation vise à fournir un instrument normatif accepté au niveau mondial qui ne se concentre pas seulement sur la définition des valeurs et des principes, mais aussi sur leur application pratique, par le biais de recommandations stratégiques concrètes, mettant fortement l'accent sur les questions d'égalité des genres et de protection de l'environnement et des écosystèmes.

7. Parce que la complexité des questions éthiques qui entourent l'IA exige la coopération des multiples parties prenantes aux différents niveaux et dans les différents secteurs des communautés internationales, régionales et nationales, la présente Recommandation vise à permettre aux parties prenantes d'assumer leur part de responsabilité par le biais d'un dialogue mondial et interculturel.

8. Les objectifs de la présente Recommandation sont les suivants :

- (a) offrir un cadre universel de valeurs, de principes et d'actions pour guider les États dans la formulation de leur législation, de leurs politiques ou d'autres instruments concernant l'IA ;
- (b) guider les actions des individus, des groupes, des communautés, des institutions et des entreprises du secteur privé afin de garantir la prise en compte de l'éthique à tous les stades du cycle de vie des systèmes d'IA ;
- (c) promouvoir le respect de la dignité humaine et de l'égalité des genres, préserver les intérêts des générations présentes et futures, et protéger les droits de l'homme, les libertés fondamentales, et l'environnement et les écosystèmes à tous les stades du cycle de vie des systèmes d'IA ;
- (d) favoriser un dialogue multipartite, pluridisciplinaire et pluraliste au sujet des questions éthiques en lien avec les systèmes d'IA ;
- (e) promouvoir un accès équitable aux progrès et aux connaissances dans le domaine de l'IA, ainsi que le partage des bienfaits qui en découlent, en accordant une attention particulière aux besoins et contributions des pays à revenu intermédiaire inférieur (PRITI), notamment aux PMA, aux PDSL et aux PEID.

## III. VALEURS ET PRINCIPES

9. Les valeurs et principes énoncés ci-dessous doivent être respectés par tous les acteurs du cycle de vie des systèmes d'IA, en premier lieu, et être promus par l'élaboration de nouvelles législations, réglementations et directives commerciales et la modification de celles qui existent déjà. Cette démarche doit respecter le droit international ainsi que les règles, principes et normes du droit international des droits de l'homme, et doit s'inscrire dans le cadre d'objectifs de durabilité sociale, politique, environnementale, éducative, scientifique et économique.

10. Les valeurs ont un rôle important en tant qu'idéaux motivant l'orientation des mesures politiques et des normes juridiques. Alors que l'ensemble des valeurs énoncées ci-dessous inspirent ainsi des comportements souhaitables et constituent les fondements des principes, les principes quant à eux explicitent les valeurs de manière plus concrète, de façon à faciliter l'application de ces dernières dans les déclarations et actions politiques.

11. Si toutes les valeurs et tous les principes décrits ci-dessous sont souhaitables en soi, dans tout contexte réel, il y a inévitablement des compromis à faire, ce qui exige de procéder à des choix complexes concernant la hiérarchisation des contextes, sans pour autant compromettre d'autres principes ou valeurs. Les compromis doivent tenir compte des préoccupations liées à la proportionnalité et à la légitimité de l'objectif. Pour faire des choix judicieux parmi de tels scénarios, il faudra généralement s'engager avec un large éventail de parties prenantes compétentes, guidées par les règles, normes et principes du droit international des droits de l'homme, en recourant au dialogue social, aux débats sur l'éthique, à la diligence requise et à l'étude d'impact.

12. Si elles sont atteintes, la fiabilité et l'intégrité du cycle de vie des systèmes d'IA sont au service de l'humanité, des individus, des sociétés, de l'environnement et des écosystèmes, et concrétisent les valeurs et principes énoncés dans la présente Recommandation. Les individus doivent avoir de bonnes raisons de ne pas douter de la répartition des avantages des systèmes d'IA lorsque des mesures appropriées sont prises pour atténuer les risques. Une condition essentielle de la fiabilité consiste à soumettre, tout au long de leur cycle de vie, les systèmes d'IA à un suivi exercé par les autorités, les entreprises du secteur privé, des organisations de la société civile et d'autres parties prenantes indépendantes. La fiabilité résultant de l'application des principes énoncés dans le présent document, les actions stratégiques proposées dans la Recommandation visent toutes à promouvoir la fiabilité à tous les stades du cycle de vie de l'IA.

### **III.1. VALEURS**

#### **Respect, protection et promotion de la dignité humaine, des droits de l'homme et des libertés fondamentales**

13. La dignité de chaque être humain est au fondement du système indivisible des droits de l'homme et des libertés fondamentales, et tient une place essentielle tout au long du cycle de vie des systèmes d'IA. La dignité humaine a trait à la reconnaissance de la valeur intrinsèque de chaque être humain, et n'est donc pas liée au sexe, au genre, à la langue, à la religion, aux opinions politiques ou autres, à l'origine nationale, ethnique, autochtone ou sociale, à l'orientation sexuelle et à l'identité de genre, à la fortune, à la naissance, au handicap, à l'âge ou à toute autre situation.

14. Aucun être humain ne doit subir de préjudice physique, économique, social, politique ou mental au cours d'une phase quelconque du cycle de vie des systèmes d'IA. Les systèmes d'IA doivent, tout au long de leur cycle de vie, améliorer la qualité de vie de chaque être humain, en laissant aux individus ou groupes le soin de définir la notion de « qualité de vie », tant qu'il n'y a ni violation des droits de l'homme, ni atteinte à la dignité humaine au sens de cette définition.

15. Les individus peuvent interagir avec les systèmes d'IA tout au long de leur cycle de vie et recevoir leur aide, par exemple pour les soins aux personnes vulnérables, incluant, mais sans s'y limiter, les enfants, les personnes âgées, les personnes handicapées ou les malades. Dans le cadre de ces interactions, les individus ne doivent jamais être traités comme des objets, leur dignité ne doit pas être mise à mal, et il ne doit y avoir ni violation des droits de l'homme ni atteinte à ces droits.

16. Les droits de l'homme et les libertés fondamentales doivent être respectés, protégés et promus tout au long du cycle de vie des systèmes d'IA. Les autorités, le secteur privé, la société civile, les organisations internationales, les communautés techniques et les universités doivent respecter les instruments et cadres relatifs aux droits de l'homme lorsqu'ils interviennent dans les processus entourant le cycle de vie des systèmes d'IA. Les nouvelles technologies doivent fournir de nouveaux moyens de promouvoir, de défendre et d'exercer les droits de l'homme, et non de les enfreindre.

## **Un environnement et des écosystèmes qui prospèrent**

17. Tout au long de leur cycle de vie, les systèmes d'IA doivent reconnaître et encourager l'importance d'avoir un environnement et des écosystèmes qui prospèrent. En outre, l'environnement et les écosystèmes sont une nécessité existentielle pour que l'humanité et les autres êtres vivants puissent profiter des bienfaits liés aux progrès de l'IA.

18. Tous les acteurs impliqués dans le cycle de vie des systèmes d'IA doivent respecter le droit international et les lois, normes et pratiques nationales pertinentes, telles que le principe de précaution, conçues pour la protection et la restauration de l'environnement et des écosystèmes, et pour le développement durable. Ils devraient réduire l'impact environnemental des systèmes d'IA, ce qui inclut, sans s'y limiter, leur empreinte carbone, afin de réduire autant que possible les facteurs de risque associés au changement climatique et aux changements environnementaux, et d'empêcher l'exploitation, l'utilisation et la transformation non durables des ressources naturelles, qui contribuent à la détérioration de l'environnement et à la dégradation des écosystèmes.

## **Assurer la diversité et l'inclusion**

19. Il convient d'assurer le respect, la protection et la promotion de la diversité et de l'inclusion tout au long du cycle de vie des systèmes d'IA, en se conformant au minimum aux règles, principes et normes du droit international des droits de l'homme, ainsi qu'à la diversité et à l'inclusion démographiques, culturelles, sociales et de genre. Cela peut passer par le fait d'encourager la participation active au cycle de vie des systèmes d'IA de tous les individus ou groupes en fonction de leur sexe, de leur genre, de leur langue, de leur religion, de leurs opinions politiques ou autres, de leur origine nationale, ethnique, autochtone ou sociale, de leur orientation sexuelle et de leur identité de genre, de leur fortune, de leur naissance, de leur handicap, de leur âge ou de toute autre situation. Toute tendance à l'homogénéisation doit être surveillée et traitée.

20. La diversité des choix de style de vie, des croyances, des opinions, des expressions ou des expériences personnelles, ce qui inclut le caractère facultatif de l'utilisation des systèmes d'IA et la coconception de ces architectures, ne doit être restreinte en aucune façon à aucun stade du cycle de vie des systèmes d'IA.

21. En outre, il faudra s'efforcer de pallier, mais jamais d'exploiter, l'absence des infrastructures, formations et compétences technologiques ainsi que des cadres juridiques nécessaires dans certaines communautés, en particulier dans les PRITI, les PMA, les PDSL et les PEID.

## **Vivre en harmonie et en paix**

22. Les acteurs de l'IA doivent favoriser la mise en œuvre d'une vie basée sur l'harmonie et la paix en assurant un avenir interconnecté pour le bien de tous. La vie en harmonie et en paix, en tant que valeur, suppose que les systèmes d'IA peuvent contribuer tout au long de leur cycle de vie à l'interdépendance de toutes les créatures vivantes entre elles et avec le milieu naturel.

23. La notion d'interdépendance des êtres humains repose sur le fait que chaque être humain appartient à un ensemble plus vaste, qui s'affaiblit lorsque d'autres êtres humains se trouvent diminués d'une quelconque façon. Vivre en harmonie et en paix nécessite un lien de solidarité organique, immédiat et non calculé, caractérisé par une recherche permanente de relations non conflictuelles et pacifiques, tendant vers le consensus avec les autres et l'harmonie avec le milieu naturel au sens large du terme.

24. Pour assurer le respect de cette valeur, la paix doit être promue tout au long du cycle de vie des systèmes d'IA, dans la mesure où les processus du cycle de vie des systèmes d'IA ne devraient pas mettre à l'écart les êtres humains, les chosifier ou menacer leur sécurité, diviser les individus et les groupes et les dresser les uns contre les autres, ou compromettre la coexistence harmonieuse entre les êtres humains, les non-humains et le milieu naturel, car cela nuirait à l'humanité tout entière.

### **III.2. PRINCIPES**

#### **Principes de proportionnalité et d'innocuité**

25. Il faut reconnaître que les technologies de l'IA ne garantissent pas nécessairement, en soi, l'épanouissement des êtres humains ni le maintien d'un environnement et d'écosystèmes prospères. Qui plus est, aucun des processus liés au cycle de vie des systèmes d'IA ne devrait aller au-delà de ce qui est nécessaire pour atteindre des buts ou objectifs légitimes, et devraient être adaptés au contexte dans lequel ils interviennent. S'il s'avérait qu'un préjudice quelconque puisse affecter les êtres humains ou l'environnement et les écosystèmes, il conviendrait d'assurer la mise en œuvre de procédures d'évaluation des risques et l'adoption de mesures visant à empêcher qu'un tel préjudice ne survienne.

26. Le choix de la méthode d'IA devrait être justifié des manières suivantes : (a) la méthode d'IA retenue devrait être souhaitable et adéquate pour atteindre un objectif légitime donné ; (b) la méthode d'IA retenue ne devrait pas porter d'atteinte aux valeurs fondamentales énoncées dans la présente Recommandation ; (c) la méthode d'IA retenue devrait être adaptée au contexte et reposer sur des bases scientifiques rigoureuses. Dans les scénarios qui impliquent des décisions de vie et de mort, la décision finale devrait être prise par l'homme.

#### **Sûreté et sécurité**

27. Les préjudices non désirés (risques liés à la sûreté) et les vulnérabilités aux attaques (risques liés à la sécurité) doivent être évités tout au long du cycle de vie des systèmes d'IA afin de garantir la sûreté et la sécurité des personnes, de l'environnement et des écosystèmes. La sûreté et la sécurité de l'IA seront rendues possibles par l'élaboration de cadres d'accès aux données durables et respectueux de la vie privée, qui favorisent un meilleur entraînement des modèles d'IA utilisant des données de qualité.

#### **Équité et non-discrimination**

28. Les acteurs de l'IA doivent promouvoir la justice sociale en respectant l'équité. L'équité implique le partage des bénéfices des technologies de l'IA aux niveaux local, national et international, tout en prenant en considération les besoins spécifiques des différents groupes d'âge, des systèmes culturels, des différents groupes linguistiques, des personnes handicapées, des filles et des femmes, et des populations défavorisées, marginalisées et vulnérables. Au niveau local, il s'agit d'œuvrer pour donner aux communautés un accès aux systèmes d'IA dans les langues de leur choix et dans le respect des différentes cultures. Au niveau national, les gouvernements sont tenus de faire preuve d'équité entre les zones rurales et urbaines, et entre toutes les personnes sans distinction de sexe, de genre, de langue, de religion, d'opinion politique ou autre, d'origine nationale, ethnique, autochtones ou sociale, d'orientation sexuelle et d'identité de genre, de fortune, de naissance, de handicap, d'âge ou toute autre situation, en termes d'accès et de participation au cycle de vie des systèmes d'IA. Au niveau international, les pays les plus avancés sur le plan technologique ont une obligation de solidarité envers les pays les moins avancés afin de garantir le partage des bienfaits des technologies de l'IA de telle sorte que l'accès et la participation de ces derniers au cycle de vie des systèmes d'IA contribuent à un ordre mondial plus équitable en matière d'information, de communication, de culture, d'éducation, de recherche et de stabilité socioéconomique et politique.

29. Les acteurs de l'IA doivent faire tout leur possible pour réduire au maximum et éviter de renforcer ou de perpétuer des biais sociotechniques inappropriés basés sur les préjugés liés à l'identité, tout au long du cycle de vie des systèmes d'IA afin de garantir l'équité de ces systèmes. Il devrait être possible d'avoir un recours contre la détermination et la discrimination algorithmiques injustes.

30. En outre, la discrimination, les fractures numérique et cognitive, et les inégalités mondiales doivent être traitées tout au long du cycle de vie des systèmes d'IA, notamment en termes d'accès aux technologies, aux données, à la connectivité, aux connaissances et aux compétences, et de participation des communautés concernées à la phase de conception, de sorte que chaque personne soit traitée équitablement.

### **Durabilité**

31. Le développement de sociétés durables repose sur la réalisation d'un ensemble complexe d'objectifs concernant divers aspects sociaux, culturels, économiques et environnementaux. L'arrivée des technologies de l'IA peut soit favoriser les objectifs de durabilité, soit entraver leur réalisation, selon la manière dont elles sont appliquées dans des pays ayant des niveaux de développement différents. L'évaluation continue de l'impact social, culturel, économique et environnemental des technologies de l'IA doit donc être effectuée en tenant pleinement compte des incidences de ces technologies pour la durabilité en tant qu'ensemble d'objectifs en constante évolution dans toute une série de dimensions, tels qu'ils sont actuellement définis dans les objectifs de développement durable (SDG) des Nations Unies.

### **Protection de la vie privée**

32. La vie privée, qui constitue un droit essentiel pour la protection de la dignité, de l'autonomie et de la capacité d'action des êtres humains, doit être respectée, protégée et promue tout au long du cycle de vie des systèmes d'IA, tant au niveau personnel que collectif. Il est essentiel que les données pour l'IA soient collectées, utilisées, partagées, archivées et supprimées d'une manière qui soit conforme aux valeurs et principes énoncés dans la présente Recommandation.

33. Des organismes de réglementation doivent mettre en place, au niveau national ou supranational, des cadres de protection des données et des mécanismes de gouvernance adéquats, protégés par les systèmes judiciaires et appliqués tout au long du cycle de vie des systèmes d'IA. Ces mécanismes et cadres de protection concernent la collecte, le contrôle et l'utilisation des données et l'exercice de leurs droits par les personnes concernées, ainsi que le droit des individus de faire effacer leurs données personnelles. Ils assurent ainsi un objectif légitime et une base juridique valable pour le traitement des données personnelles ainsi que pour la personnalisation, la dépersonnalisation et la repersonnalisation des données, la transparence, la mise en place de garanties appropriées pour les données sensibles, et un contrôle indépendant efficace.

34. Les systèmes algorithmiques nécessitent des évaluations approfondies de l'impact sur la vie privée qui incluent également des considérations sociétales et éthiques de leur utilisation et une approche innovante consistant à appliquer le respect de la vie privée dès la conception des systèmes.

### **Surveillance et décision humaines**

35. Il doit toujours être possible d'attribuer la responsabilité éthique et juridique de tout stade du cycle de vie des systèmes d'IA à des personnes physiques ou des entités juridiques existantes. Ainsi, la surveillance humaine renvoie non seulement à la surveillance humaine individuelle, mais aussi à la surveillance publique, selon le cas.

36. Les êtres humains devront peut-être, dans certains cas, s'appuyer sur les systèmes d'IA à des fins d'efficacité, mais la décision de céder les fonctions de contrôle dans des contextes limités reste leur prérogative. En effet, les êtres humains peuvent recourir aux systèmes d'IA pour la prise de décisions et l'exécution de tâches, mais un système d'IA ne peut jamais se substituer à la responsabilité ultime des êtres humains et à leur obligation de rendre compte.

### **Transparence et explicabilité**

37. La transparence des systèmes d'IA est souvent une condition préalable essentielle pour garantir le respect, la protection et la promotion des droits humains fondamentaux et des principes éthiques. La transparence est nécessaire pour que la législation nationale et internationale en matière de responsabilité fonctionne efficacement.

38. S'il convient de tout mettre en œuvre pour améliorer la transparence et l'explicabilité des systèmes d'IA tout au long de leur cycle de vie afin de favoriser une gouvernance démocratique, le degré de transparence et d'explicabilité devrait toujours être adapté au contexte, car des arbitrages sont effectués entre la transparence et l'explicabilité et d'autres principes, tels que la sûreté et la sécurité. Les individus ont le droit de savoir lorsqu'une décision est prise sur la base d'algorithmes d'IA et, dans ces circonstances, d'exiger ou de demander des explications aux entreprises du secteur privé ou aux institutions du secteur public.

39. D'un point de vue sociotechnique, une plus grande transparence contribue à des sociétés plus pacifiques, plus justes et plus inclusives. Elle permet un contrôle public qui peut réduire la corruption et la discrimination, et peut également aider à détecter et à prévenir les impacts négatifs sur les droits de l'homme. La transparence peut contribuer à établir la confiance des êtres humains dans les systèmes d'IA. Concernant spécifiquement les systèmes d'IA, la transparence peut permettre aux individus de comprendre comment chaque étape d'un système d'IA est mis en place, en fonction du contexte et du degré de sensibilité du système concerné. Il est possible également de fournir des informations sur les facteurs qui influencent une prévision ou une décision particulière, et sur l'existence ou non des garanties appropriées (telles que des mesures de sécurité ou d'équité). Dans les cas où l'on prévoit de graves répercussions négatives sur les droits de l'homme, la transparence peut également exiger la communication de codes ou d'ensembles de données spécifiques.

40. L'explicabilité implique de rendre les résultats des systèmes d'IA intelligibles et de fournir des renseignements à leur sujet. L'explicabilité des systèmes d'IA renvoie également à l'intelligibilité des intrants, des extrants, du comportement des différents modules algorithmiques et de leur contribution aux résultats des systèmes. L'explicabilité est donc étroitement liée à la transparence, puisque les résultats et les sous-processus qui y conduisent doivent être intelligibles et traçables, en fonction du contexte d'utilisation.

41. La transparence et l'explicabilité sont étroitement liées à des mesures adéquates en matière de responsabilité et d'obligation de rendre compte, ainsi qu'à la fiabilité des systèmes d'IA.

### **Responsabilité et redevabilité**

42. Les acteurs de l'IA devraient respecter, protéger et promouvoir les droits de l'homme et favoriser la protection de l'environnement et des écosystèmes, endossant une responsabilité éthique et juridique, conformément au droit national et international en vigueur, en particulier les règles, principes et normes du droit international des droits de l'homme, et aux directives éthiques établies tout au long du cycle de vie des systèmes d'IA. La responsabilité éthique des décisions et actions fondées d'une quelconque manière sur un système d'IA devrait toujours incomber en dernier ressort aux acteurs de l'IA.

43. Des mécanismes appropriés de surveillance, d'étude d'impact et de diligence requise devraient être mis en place pour assurer la redevabilité des systèmes d'IA et de leur impact tout au long de leur cycle de vie. Des dispositifs techniques et institutionnels devraient garantir la vérifiabilité et la traçabilité (du fonctionnement) des systèmes d'IA, en particulier pour résoudre tout conflit concernant les droits de l'homme et traiter les menaces pesant sur le bien-être de l'environnement et des écosystèmes.

### **Sensibilisation et éducation**

44. La sensibilisation du public et sa compréhension des technologies de l'IA et de la valeur des données devraient être favorisées par une éducation ouverte et accessible, l'engagement civique, l'acquisition de compétences numériques et la formation à l'éthique de l'IA, l'éducation aux médias et à l'information et la formation menée conjointement par les gouvernements, les organisations intergouvernementales, la société civile, les universités, les médias, les dirigeants locaux et le secteur privé, et en tenant compte de la diversité linguistique, sociale et culturelle existante, afin de garantir une participation publique efficace de sorte que tous les membres de la société puissent prendre des décisions éclairées concernant leur utilisation des systèmes d'IA et soient protégés contre toute influence induite.

45. La sensibilisation à l'impact des systèmes d'IA doit inclure un apprentissage sur, par et pour les droits de l'homme, ce qui signifie que l'approche et la compréhension des systèmes d'IA doivent se baser sur l'impact de ces systèmes sur les droits de l'homme et sur l'accès à ces droits.

### **Gouvernance et collaboration multipartites et adaptatives**

46. Le droit international et la souveraineté doivent être respectés dans l'utilisation des données. La souveraineté des données signifie que les États, conformément au droit international, réglementent les données générées sur leur territoire ou transitant par celui-ci, et prennent des mesures en vue d'une réglementation efficace des données fondée sur le respect du droit à la vie privée et d'autres droits humains.

47. La participation des différentes parties prenantes tout au long du cycle de vie des systèmes d'IA est nécessaire pour une gouvernance inclusive de l'IA, le partage des bénéfices de l'IA, la réalisation de progrès technologiques équitables et leur contribution aux objectifs de développement. Les parties prenantes comprennent, sans s'y limiter, les gouvernements, les organisations intergouvernementales, la communauté technique, la société civile, les chercheurs et les universitaires, les médias, les responsables de l'éducation, les décideurs, les entreprises du secteur privé, les institutions de défense des droits de l'homme et les organismes de promotion de l'égalité, les organismes de contrôle des mesures de lutte contre la discrimination et les groupes de jeunes et d'enfants. Il convient d'adopter des normes ouvertes et d'assurer l'interopérabilité pour faciliter la collaboration. Des mesures doivent être adoptées pour tenir compte de l'évolution des technologies et de l'émergence de nouveaux groupes de parties prenantes, et pour permettre aux groupes, communautés et individus marginalisés d'intervenir efficacement.

## **IV. DOMAINES D'ACTION STRATÉGIQUE**

48. Les actions stratégiques décrites dans les domaines suivants traduisent dans les faits les valeurs et les principes énoncés dans la présente Recommandation. La principale action consiste pour les États membres à mettre en place des cadres ou des mécanismes stratégiques et à veiller à ce que les autres parties prenantes, telles que les entreprises du secteur privé, les établissements universitaires, les instituts de recherche et la société civile, y adhèrent, notamment en aidant toutes les parties prenantes à élaborer des outils d'évaluation de l'impact éthique et de diligence requise. Le processus d'élaboration de ces politiques ou

mécanismes devrait inclure toutes les parties prenantes et tenir compte de la situation et des priorités de chaque État membre. L'UNESCO peut être un partenaire et aider les États membres dans l'élaboration ainsi que le suivi et l'évaluation des mécanismes stratégiques.

49. L'UNESCO reconnaît que les États membres se trouveront à différents stades de préparation pour mettre en œuvre la présente Recommandation, des points de vue scientifique, technologique, économique, éducatif, juridique, réglementaire, infrastructurel, sociétal et culturel, entre autres. Il convient de noter que dans le cas présent, « l'état de préparation » est un état dynamique. Afin de permettre la mise en œuvre effective de la présente Recommandation, l'UNESCO va donc : (1) élaborer une méthode d'évaluation de l'état de préparation pour aider les États membres à identifier leur statut à des moments spécifiques de l'avancement de leur préparation à travers un ensemble de dimensions ; et (2) assurer un soutien aux États membres pour l'élaboration d'une méthodologie acceptée au niveau mondial concernant l'évaluation de l'impact éthique des technologies de l'IA, le partage de bonnes pratiques, de lignes directrices pour l'évaluation et d'autres mécanismes et travaux analytiques.

## **DOMAINE STRATÉGIQUE 1 : ÉTUDES D'IMPACT EN MATIÈRE D'ÉTHIQUE**

50. Les États membres devraient mettre en place des études d'impact pour identifier et analyser les avantages et les risques des systèmes d'IA et les questions qu'ils soulèvent, ainsi que des mesures de prévention, d'atténuation et de suivi des risques. L'étude d'impact en matière d'éthique devrait mettre en évidence les répercussions sur les droits de l'homme, notamment les droits des groupes vulnérables, le droit du travail, l'environnement et les écosystèmes, ainsi que les incidences éthiques et sociales conformément aux principes énoncés dans la présente Recommandation.

51. Les États membres et les entreprises du secteur privé devraient mettre en place des mécanismes de diligence requise et de supervision pour identifier, prévenir et atténuer les risques et rendre compte de la manière dont ils traitent les répercussions des systèmes d'IA sur les droits de l'homme, l'État de droit et les sociétés inclusives. Les États membres devraient également être capables d'évaluer les effets socioéconomiques des systèmes d'IA sur la pauvreté et de s'assurer que le fossé entre les riches et les pauvres, ainsi que la fracture numérique entre les pays et à l'intérieur de ceux-ci, ne s'accroissent pas avec l'adoption massive des technologies d'IA aujourd'hui et à l'avenir. Pour ce faire, des protocoles de transparence exécutoires devraient être mis en œuvre, correspondant au droit d'accès à l'information, notamment aux informations d'intérêt public détenues par des entités privées.

52. Les États membres et les entreprises du secteur privé devraient mettre en œuvre des mesures appropriées pour surveiller toutes les phases du cycle de vie d'un système d'IA, y compris le comportement des algorithmes utilisés pour la prise de décisions, les données, ainsi que les acteurs de l'IA impliqués dans le processus, en particulier dans les services publics et lorsqu'une interaction directe avec l'utilisateur final est nécessaire.

53. Les gouvernements devraient adopter un cadre réglementaire qui définit une procédure permettant en particulier aux autorités publiques de mener à bien des études sur l'impact des systèmes d'IA afin d'anticiper les répercussions, d'atténuer les risques, d'éviter les conséquences préjudiciables, de faciliter la participation des citoyens et de faire face aux défis sociétaux. L'étude devrait également établir des mécanismes de supervision adaptés, notamment les principes de vérifiabilité, de traçabilité et d'explicabilité, permettant d'évaluer les algorithmes, les données et les processus de conception, ainsi qu'inclure un examen externe des systèmes d'IA. Les études d'impact en matière d'éthique réalisées par les autorités publiques doivent être transparentes et ouvertes au public. Elles devraient par ailleurs être multidisciplinaires, multipartites, multiculturelles, pluralistes et inclusives. Les États membres sont encouragés à mettre en place des mécanismes et des outils, par exemple

des bacs à sable réglementaires ou des centres d'évaluation, qui devraient permettre de suivre et d'évaluer les impacts de manière multidisciplinaire et multipartite. Les autorités publiques devraient être tenues de surveiller les systèmes d'IA qu'elles mettent en œuvre et/ou déploient, en établissant des mécanismes et des outils appropriés.

54. Les États membres devraient mettre en place des mécanismes de suivi et d'évaluation pour les initiatives et les politiques liées à l'éthique de l'IA. Ces mécanismes peuvent être : un répertoire couvrant le développement de systèmes d'IA éthique et conforme aux droits de l'homme ; un mécanisme de partage d'expériences permettant aux États membres de demander à d'autres États membres leur avis sur leurs politiques et initiatives ; un guide permettant à tous les acteurs de l'IA d'évaluer le respect des recommandations stratégiques mentionnées dans le présent document ; ainsi que des outils de suivi. Les règles, principes et normes du droit international des droits de l'homme devraient faire partie des aspects éthiques des évaluations des systèmes d'IA.

## **DOMAINE STRATÉGIQUE 2 : GOUVERNANCE ET GESTION ÉTHIQUES**

55. Les États membres devraient s'assurer que tout mécanisme de gouvernance de l'IA est inclusif, transparent, multidisciplinaire, multilatéral (cela inclut la possibilité d'atténuer et de réparer les dommages au-delà des frontières) et multipartite. La gouvernance devrait comporter des aspects d'anticipation, de protection, de suivi des conséquences, de mise en application et de réparation.

56. Les États membres devraient veiller à ce que les préjudices causés à des utilisateurs par le biais de systèmes d'IA fassent l'objet d'enquêtes et de réparations, en mettant en place des mécanismes d'exécution solides et des mesures correctives afin de s'assurer que les droits de l'homme et l'État de droit sont respectés dans le monde numérique comme ils le sont dans le monde physique. Ces mécanismes et mesures devraient inclure les mécanismes de réparation fournis par des entreprises du secteur privé. La vérifiabilité et la traçabilité des systèmes d'IA devraient être encouragées à cette fin. En outre, les États membres devraient renforcer leurs capacités institutionnelles à s'acquitter de ce devoir de diligence et devraient collaborer avec les chercheurs et les autres parties prenantes pour analyser, prévenir et freiner toute utilisation potentiellement malveillante des systèmes d'IA.

57. Les États membres sont encouragés à envisager l'adoption de formes de gouvernance souple, par exemple un mécanisme de certification pour les systèmes d'IA et la reconnaissance mutuelle de leurs certifications, selon la sensibilité du domaine d'application et selon l'impact escompté sur les vies humaines, l'environnement et les écosystèmes, ainsi que les autres considérations éthiques énoncées dans la présente Recommandation. Un tel mécanisme pourrait inclure différents niveaux d'audit des systèmes, des données et de la conformité aux principes éthiques, et devrait être validé par les parties autorisées dans chaque pays. Dans le même temps, il ne doit pas entraver l'innovation ni désavantager les petites et les moyennes entreprises ou les start-up en exigeant de grandes quantités de documents administratifs. Ce mécanisme inclurait en outre un volet consacré à un suivi régulier pour garantir la fiabilité du système et le maintien de son intégrité et de sa conformité aux principes éthiques tout au long de sa durée de vie, en exigeant une nouvelle certification si nécessaire.

58. Les gouvernements et les autorités publiques devraient être tenus de procéder à une auto-évaluation des systèmes d'IA existants et proposés, qui devrait notamment analyser si l'adoption de l'IA est appropriée et, le cas échéant, prévoir une évaluation supplémentaire pour déterminer quelle est la méthode appropriée, ainsi qu'une évaluation visant à déterminer si la méthode adoptée transgresse toute règle, tout principe ou toute norme en matière de droits de l'homme.

59. Les États membres devraient encourager les entités publiques, les entreprises du secteur privé et les organisations de la société civile à associer différentes parties prenantes à leur gouvernance en matière d'IA et à envisager d'ajouter une fonction de responsable de l'éthique de l'IA indépendant ou tout autre mécanisme visant à superviser les efforts déployés concernant l'étude d'impact en matière d'éthique, le contrôle et le suivi continu, ainsi qu'à assurer la conformité éthique des systèmes d'IA. Les États membres, les entreprises du secteur privé et les organisations de la société civile, avec le soutien de l'UNESCO, sont encouragés à créer un réseau de responsables de l'éthique de l'IA indépendants pour appuyer ce mécanisme aux niveaux national, régional et international.

60. Les États membres devraient favoriser le développement et l'accessibilité d'un écosystème numérique pour le développement éthique des systèmes d'IA à l'échelle nationale, tout en encourageant la collaboration internationale. Cet écosystème engloberait notamment des technologies et infrastructures numériques et des mécanismes de partage des connaissances en matière d'IA, le cas échéant. À cet égard, les États membres devraient envisager de revoir leurs politiques et leurs cadres réglementaires, notamment en ce qui concerne l'accès à l'information et les principes d'un gouvernement ouvert, pour rendre compte des exigences propres à l'IA et promouvoir des mécanismes, tels que des dépôts en libre accès pour les données et les codes sources publics ou financés par des fonds publics ainsi que des fiduciaires de données (« data trusts »), afin de favoriser le partage des données de façon sûre, équitable, légale et éthique, entre autres.

61. Les États membres devraient établir des mécanismes, en collaboration avec les organisations internationales, les sociétés transnationales, les établissements universitaires et la société civile, afin de garantir la participation active de tous les États membres, en particulier les PRITI, notamment les PMA, les PDSL et les PEID, aux discussions internationales concernant la gouvernance de l'IA. Cela peut passer par la mise à disposition de fonds, le fait d'assurer une participation régionale égale, ou tout autre mécanisme. En outre, afin de s'assurer que les forums sur l'IA soient ouverts à tous, les États membres devraient faciliter les déplacements des acteurs de l'IA, notamment ceux originaires des PRITI et en particulier des PMA, des PDSL et des PEID, au sein et en dehors de leur territoire, afin qu'ils puissent participer à ces forums.

62. Les modifications apportées aux législations nationales existantes ou l'élaboration de nouvelles législations nationales concernant les systèmes d'IA doivent respecter le droit international des droits de l'homme et promouvoir les droits humains et les libertés fondamentales tout au long de la durée de vie du système d'IA. La promotion des droits de l'homme devrait également prendre la forme d'initiatives de gouvernance, de modèles de bonnes pratiques de collaboration concernant les systèmes d'IA, ainsi que de directives techniques et méthodologiques élaborées par les autorités nationales et internationales pour s'adapter aux progrès des technologies d'IA. Différents secteurs, y compris le secteur privé, dans leurs pratiques concernant les systèmes d'IA, doivent respecter, protéger et promouvoir les droits de l'homme et les libertés fondamentales au moyen d'instruments nouveaux et existants, en association avec la présente Recommandation.

63. Les États membres devraient prévoir des mécanismes permettant d'évaluer l'impact de la surveillance et du contrôle des systèmes d'IA sur l'exercice des droits de l'homme ainsi que leur impact économique et social, ainsi que d'autres mécanismes de gouvernance tels que des autorités indépendantes de protection des données, une surveillance sectorielle, des organismes publics chargés de contrôler l'acquisition de systèmes d'IA pour les cas d'utilisation sensibles en matière de droits de l'homme, comme la justice pénale, l'application des lois, la protection sociale, l'emploi, les soins de santé, entre autres, ainsi que des systèmes judiciaires indépendants.

64. Les États membres devraient s'assurer que les gouvernements et les organisations multilatérales jouent un rôle de premier plan pour garantir la sûreté et la sécurité des systèmes d'IA. En particulier, les États membres, les organisations internationales et les autres organes concernés devraient élaborer des normes internationales qui décriraient des niveaux de sûreté et de transparence mesurables et vérifiables pour pouvoir évaluer les systèmes de manière objective et déterminer les niveaux de conformité. Les États membres devraient en outre apporter un soutien constant à la recherche stratégique sur les risques de sûreté et de sécurité potentiels des technologies d'IA et devraient encourager la recherche sur la transparence et l'explicabilité en y consacrant des fonds supplémentaires pour différents domaines et à différents niveaux, tels que le langage technique et naturel.

65. Les États membres devraient mettre en œuvre des politiques visant à garantir que les activités des acteurs de l'IA sont conformes aux règles, aux normes et aux principes du droit international des droits de l'homme tout au long de la durée de vie des systèmes d'IA, et qu'elles prennent en considération et respectent les diversités culturelles et sociales actuelles, notamment les coutumes locales et les traditions religieuses.

66. Les États membres devraient mettre en place des mécanismes pour exiger des acteurs de l'IA qu'ils révèlent et combattent tout stéréotype présent dans les résultats et les données des systèmes d'IA, volontairement ou par négligence, et qu'ils veillent à ce que les ensembles de données d'entraînement des systèmes d'IA ne favorisent pas les inégalités culturelles, économiques ou sociales, les préjugés, la propagation d'informations non fiables ou la diffusion d'idées antidémocratiques. Une attention particulière devrait être portée aux régions où les données sont rares.

67. Les États membres devraient mettre en œuvre des politiques visant à promouvoir et à accroître la diversité des équipes de développement de l'IA et des ensembles de données d'entraînement, et à garantir l'accès équitable aux technologies d'IA et à leurs avantages, en particulier pour les groupes marginalisés, aussi bien dans les zones rurales que dans les zones urbaines.

68. Les États membres devraient revoir et adapter, le cas échéant, les cadres réglementaires et juridiques pour établir la redevabilité et la responsabilité des contenus et des résultats des systèmes d'IA aux différentes phases de leur cycle de vie. Les États membres devraient définir des cadres de responsabilité ou clarifier l'interprétation des cadres existants pour assurer l'attribution de la responsabilité des résultats et du comportement des systèmes d'IA. En outre, lors de l'élaboration de cadres réglementaires, les États membres devraient notamment tenir compte du fait que la responsabilité et la redevabilité devraient toujours incomber au final à une personne physique ou morale, et que les systèmes d'IA ne devraient pas eux-mêmes bénéficier d'une personnalité juridique. Pour ce faire, ces cadres réglementaires devraient être compatibles avec le principe de la surveillance humaine et établir une approche globale, axée sur les acteurs et les processus technologiques impliqués dans les différentes étapes du cycle de vie des systèmes d'IA.

69. Les États membres devraient renforcer la capacité du pouvoir judiciaire à prendre des décisions relatives aux systèmes d'IA, dans le respect de l'État de droit et conformément aux normes internationales, notamment en ce qui concerne l'utilisation des systèmes d'IA dans leurs délibérations, tout en veillant au respect du principe de la surveillance humaine.

70. Afin d'établir des normes lorsqu'il n'y en a pas, ou d'adapter les cadres réglementaires existants, les États membres devraient impliquer tous les acteurs de l'IA (chercheurs, représentants de la société civile et des forces de l'ordre, assureurs, investisseurs, industriels, ingénieurs, avocats et utilisateurs, entre autres). Les normes peuvent évoluer en bonnes pratiques, en lois et en règlements. Les États membres sont également encouragés à utiliser des mécanismes tels que des modèles de politiques et des bacs à sable réglementaires pour

accélérer la formulation des lois, règlements et politiques et suivre ainsi l'évolution rapide des nouvelles technologies, et pour que les lois et règlements puissent être testés dans un environnement sûr avant d'être officiellement adoptés. Les États membres devraient aider les autorités locales à élaborer des politiques, des règlements et des lois à l'échelle locale, conformément aux cadres juridiques nationaux et internationaux.

71. Les États membres devraient définir des exigences claires en matière de transparence et d'explicabilité des systèmes d'IA afin d'aider à garantir la crédibilité des systèmes tout au long de leur durée de vie. Ces exigences devraient inclure la conception et la mise en œuvre de mécanismes d'évaluation de l'impact qui prennent en considération la nature du domaine d'application (s'agit-il d'un domaine à haut risque tel que le maintien de l'ordre, la sécurité, l'éducation, l'emploi et la santé ?), l'utilisation prévue (quels sont les risques en termes de transgression de la sécurité et des droits de l'homme ?), le public cible (qui demande les informations ?) et la faisabilité (l'algorithme est-il explicable ou non, et quels sont les compromis trouvés entre exactitude et explicabilité ?) de chaque système d'IA spécifique.

### **DOMAINE STRATÉGIQUE 3 : POLITIQUES EN MATIÈRE DE DONNÉES**

72. Les États membres devraient s'efforcer d'élaborer des stratégies de gouvernance des données qui garantissent l'évaluation continue de la qualité des données d'entraînement des systèmes d'IA, notamment l'adéquation des processus de collecte et de sélection des données, et qui prévoient des mesures de sécurité et de protection des données appropriées, ainsi que des mécanismes de retour d'information permettant de tirer des enseignements des erreurs et de partager de bonnes pratiques entre tous les acteurs de l'IA. Il est fondamental de trouver un équilibre entre la collecte de métadonnées et la protection de la vie privée des utilisateurs lors de l'élaboration de ces stratégies.

73. Les États membres devraient mettre en place des garanties appropriées pour reconnaître et protéger le droit fondamental des individus au respect de la vie privée, notamment par l'adoption ou la mise en œuvre de cadres législatifs qui assurent une protection adéquate, conforme au droit international. Les États membres devraient vivement encourager l'ensemble des acteurs de l'IA, y compris les entreprises du secteur privé, à appliquer les normes internationales en vigueur et en particulier à conduire des évaluations de l'impact sur la vie privée, dans le cadre des études d'impact en matière d'éthique, qui tiennent compte de l'impact socioéconomique global du traitement prévu des données, ainsi qu'à appliquer le principe du respect de la vie privée dès la conception de leurs systèmes. Le droit à la vie privée devrait être respecté, protégé et encouragé pendant toute la durée de vie des systèmes d'IA.

74. Les États membres devraient veiller à ce que les personnes conservent des droits sur leurs données personnelles et soient protégées par un cadre qui prévoit notamment la transparence, des garanties appropriées pour le traitement des données sensibles, le plus haut niveau de sécurité des données, des mécanismes et des systèmes de responsabilisation efficaces et cohérents, la pleine jouissance des droits des personnes concernées, en particulier le droit d'accès et le droit d'effacement de leurs données personnelles dans les systèmes d'IA, un niveau approprié de protection lorsque les données sont utilisées à des fins commerciales – par exemple pour permettre la publicité microciblée – ou transférées à l'étranger, et une surveillance indépendante efficace dans le cadre d'un mécanisme de gouvernance des données qui respecte la souveraineté des données et la met en balance avec les avantages d'une libre circulation des informations au niveau international, y compris l'accès aux données.

75. Les États membres devraient établir leurs politiques en matière de données ou leurs cadres équivalents, ou renforcer les politiques et cadres déjà existants, afin d'assurer une sécurité accrue pour les données personnelles et sensibles qui, si elles étaient divulguées, risqueraient de causer des dommages, des blessures ou des difficultés exceptionnelles à une personne. Il s'agit par exemple des données relatives aux infractions, aux poursuites pénales

et aux condamnations, ainsi qu'aux mesures de sécurité qui y sont liées ; des données biométriques et génétiques ; des données personnelles concernant l'origine ethnique ou sociale, les opinions politiques, l'appartenance à un syndicat, les croyances religieuses ou autres, la santé et la vie sexuelle.

76. Les États membres devraient utiliser les systèmes d'IA pour améliorer l'accès à l'information et au savoir, notamment à leurs banques de données, et combler les lacunes en matière d'accès tout au long du cycle de vie des systèmes d'IA. Ils peuvent notamment mettre en place des initiatives visant à aider les chercheurs et les développeurs à améliorer la liberté d'expression et l'accès à l'information, et à accroître la divulgation proactive des données et informations officielles. Les États membres devraient également promouvoir les données ouvertes, notamment en créant des dépôts en libre accès pour les données et les codes sources publics ou financés par des fonds publics.

77. Les États membres devraient s'assurer de la qualité globale et de la solidité des ensembles de données pour l'IA, et faire preuve d'une extrême vigilance dans la supervision de leur collecte et de leur utilisation. Lorsque c'est possible et faisable, cela pourrait supposer d'investir dans la création d'ensembles de données de référence, incluant des ensembles de données ouverts et fiables, diversifiés, établis sur une base légale valable, notamment avec le consentement des personnes concernées, lorsque le consentement est requis par la loi. L'élaboration de normes pour l'annotation des ensembles de données devrait être encouragée, afin de pouvoir déterminer facilement comment un ensemble de données est constitué et quelles sont ses propriétés.

78. Comme l'indique également le rapport du Groupe de haut niveau du Secrétaire général de l'Organisation des Nations Unies sur la coopération numérique, les États membres devraient, avec le soutien de l'Organisation des Nations Unies et de l'UNESCO, adopter une approche d'« espace commun numérique » des données, le cas échéant, accroître l'interopérabilité des outils et des ensembles de données ainsi que des interfaces des systèmes hébergeant les données, et encourager les entreprises du secteur privé à partager les données qu'elles collectent au profit de la recherche ou de l'intérêt public, selon les besoins. Ils devraient également promouvoir les efforts publics et privés visant à créer des plates-formes de collaboration pour partager des données de qualité dans des espaces de données fiables et sécurisés.

#### **DOMAINE STRATÉGIQUE 4 : DÉVELOPPEMENT ET COOPÉRATION INTERNATIONALE**

79. Les États membres et les sociétés transnationales devraient accorder une attention prioritaire à l'éthique de l'intelligence artificielle en examinant les questions éthiques relatives à l'IA au sein des instances internationales, intergouvernementales et multipartites concernées.

80. Les États membres devraient veiller à ce que l'utilisation de l'IA dans les domaines liés au développement – les soins de santé, l'agriculture/l'approvisionnement alimentaire, l'éducation, les médias, la culture, l'environnement, la gestion de l'eau, la gestion des infrastructures, la planification et la croissance économiques, entre autres – respecte les valeurs et les principes énoncés dans la présente Recommandation.

81. Les États membres devraient s'employer, au sein des organisations internationales, à mettre en place des cadres de coopération internationale dans le domaine de l'IA au service du développement, notamment en mettant à disposition des compétences techniques, des financements, des données, des connaissances spécialisées, des infrastructures, et en facilitant la collaboration entre experts techniques et commerciaux pour résoudre les problèmes complexes en matière de développement, en particulier pour les PRITI, notamment les PMA, les PDSL et les PEID.

82. Les États membres devraient s'efforcer de promouvoir la collaboration internationale en matière de recherche et d'innovation sur l'IA, notamment au sein des réseaux et des centres de recherche et d'innovation qui favorisent une participation accrue et un plus grand leadership des chercheurs venant des PRITI et d'autres régions, notamment des PMA, des PDSL et des PEID.

83. Les États membres devraient soutenir les recherches sur l'éthique de l'IA menées par les organisations internationales et les établissements de recherche, ainsi que les sociétés transnationales, ces travaux pouvant servir de base à l'utilisation éthique des systèmes d'IA par des entités publiques et privées, notamment les recherches sur l'applicabilité de cadres éthiques adaptés selon les cultures et les contextes, et sur les possibilités d'associer ces cadres à des solutions réalisables sur le plan technologique.

84. Les États membres devraient encourager la coopération et la collaboration internationales dans le domaine de l'IA afin d'éliminer les clivages géotechnologiques. Des échanges/consultations sur les questions d'ordre technologique devraient avoir lieu entre les États membres et leurs populations, entre les secteurs public et privé, ainsi qu'entre les États membres du Nord et du Sud et au sein d'entre eux.

85. Les États membres devraient élaborer et mettre en œuvre un cadre juridique international pour encourager la coopération internationale entre les États et les autres parties prenantes, en accordant une attention particulière à la situation des PRITI, notamment les PMA, les PDSL et les PEID.

## **DOMAINE STRATÉGIQUE 5 : ENVIRONNEMENT ET ÉCOSYSTÈMES**

86. Les États membres devraient évaluer l'impact environnemental direct et indirect du système d'IA tout au long de son cycle de vie, notamment, sans s'y limiter, son empreinte carbone, sa consommation d'énergie et l'impact environnemental de l'extraction des matières premières nécessaires à la fabrication des technologies d'IA. Ils devraient garantir que tous les acteurs du domaine de l'IA respectent les législations, les politiques et les pratiques relatives à l'environnement.

87. Les États membres devraient mettre en place, selon les besoins, des mesures d'incitation permettant d'assurer l'élaboration et l'adoption de solutions fondées sur les droits et sur une IA éthique en faveur de la résilience face aux risques de catastrophe ; de la surveillance, de la protection et de la régénération de l'environnement et des écosystèmes ; et de la préservation de la planète. Ces systèmes d'IA devraient reposer sur la participation des communautés locales et autochtones tout au long de leur cycle de vie et favoriser des approches telles que l'économie circulaire et les modes de consommation et de production durables. Les systèmes d'IA peuvent ainsi, selon les besoins et s'il y a lieu, être utilisés pour :

- (a) soutenir la protection, le suivi et la gestion des ressources naturelles ;
- (b) soutenir la prévention, le contrôle et la gestion des problèmes liés au climat ;
- (c) favoriser un écosystème alimentaire plus efficient et plus durable ;
- (d) contribuer à accélérer l'accès aux énergies durables et leur adoption par le plus grand nombre ;
- (e) permettre et promouvoir la généralisation d'infrastructures durables, de modèles économiques durables et de financements durables au service du développement durable ;

- (f) détecter les polluants ou prévoir les niveaux de pollution et aider ainsi les parties prenantes concernées à définir, planifier et mettre en place des interventions ciblées pour prévenir et réduire la pollution et l'exposition à celle-ci.

88. Au moment de sélectionner une méthode d'IA, compte tenu de l'importance des besoins en données et en ressources de certaines d'entre elles et de l'impact sur l'environnement qui en résulte, les États membres devraient veiller à ce que les acteurs de l'IA, suivant le principe de proportionnalité, privilégient les méthodes d'IA économes en données, en énergie et en ressources. Il convient de fixer des conditions visant à garantir que l'on dispose des éléments probants appropriés pour prouver qu'une application d'IA aura l'effet escompté, ou que cette application est assortie de garanties qui permettent de la justifier.

## **DOMAINE STRATÉGIQUE 6 : ÉGALITÉ DES GENRES**

89. Les États membres devraient veiller à ce que les technologies numériques et l'intelligence artificielle contribuent pleinement à l'égalité des genres et qu'il ne soit porté atteinte aux droits et aux libertés fondamentales des filles et des femmes, notamment à leur sécurité et à leur intégrité, à aucun stade du cycle de vie du système d'IA. En outre, les études d'impact en matière d'éthique devraient prendre en compte les questions de genre de manière transversale.

90. Les États membres devraient allouer des fonds publics au financement des dispositifs relatifs à l'égalité des genres, veiller à ce que les politiques nationales relatives au numérique prévoient un plan d'action pour l'égalité des genres et élaborer des politiques spécifiques, par exemple sur la préparation à la vie active, en vue de soutenir les filles et les femmes et de faire en sorte que celles-ci ne soient pas exclues de l'économie numérique reposant sur l'IA. Certains investissements spéciaux devraient être envisagés et réalisés afin de proposer des programmes ciblés et un langage respectueux de l'égalité des genres, de façon à élargir les possibilités de participation des filles et des femmes aux sciences, à la technologie, à l'ingénierie et aux mathématiques (STEM), notamment aux disciplines liées aux technologies de l'information et de la communication (TIC), et à améliorer la préparation, l'employabilité, l'évolution de carrières et l'épanouissement professionnel des filles et des femmes.

91. Les États membres devraient faire en sorte qu'il soit tiré parti du potentiel des systèmes d'AI en faveur de l'égalité des genres. Ils devraient garantir que ces technologies ne contribueront pas à accentuer les inégalités entre les genres déjà marquées qui existent dans plusieurs domaines du monde analogique. Il s'agit notamment des inégalités salariales entre hommes et femmes ; de l'écart de représentation dans certaines professions et activités ; de la faible représentation aux postes de direction, aux conseils d'administration ou dans les équipes de recherche du domaine de l'IA ; des disparités dans l'accès à l'éducation ; des inégalités en matière d'accessibilité, d'adoption, d'utilisation et de coût abordable du numérique/de l'IA ; de la répartition inégale du travail non rémunéré et des responsabilités en matière de soins dans nos sociétés.

92. Les États membres devraient veiller à ce que les stéréotypes fondés sur le genre et les préjugés discriminatoires ne soient pas transposés dans les systèmes d'IA. Des efforts sont nécessaires pour éviter les effets négatifs cumulés des fractures technologiques dans la réalisation de l'égalité des genres et la prévention de la violence à l'égard des filles et des femmes, et de tous les autres types d'identités de genre.

93. Les États membres devraient encourager l'esprit d'entreprise, la participation et l'engagement des femmes à tous les stades du cycle de vie d'un système d'IA en proposant et en soutenant des incitations économiques et réglementaires, entre autres mesures d'encouragement et de soutien, ainsi que des politiques visant à assurer une participation équilibrée entre les genres à la recherche sur l'IA au sein des universités, et la représentation

des différents genres aux postes de direction ainsi que dans les conseils d'administration et les équipes de recherche des entreprises du secteur numérique ou de l'IA. Les autorités publiques devraient veiller à ce que les fonds publics (en faveur de l'innovation, de la recherche et de la technologie) soient affectés à des programmes et à des entreprises inclusifs, avec un équilibre clair entre les genres, et à ce que les fonds privés soient encouragés suivant les principes de discrimination positive. Par ailleurs, des politiques favorisant des environnements exempts de toute forme de harcèlement devraient être élaborées et appliquées, tout en encourageant la diffusion des bonnes pratiques concernant les moyens de promouvoir la diversité tout au long du cycle de vie du système d'IA.

94. L'UNESCO peut aider à constituer un recueil des meilleures pratiques visant à encourager la participation des femmes et des groupes sous-représentés à tous les stades du cycle de vie de l'IA.

## **DOMAINE STRATÉGIQUE 7 : CULTURE**

95. Les États membres sont encouragés à établir, s'il y a lieu, des systèmes d'IA dans les domaines de la conservation, de l'enrichissement, de la compréhension, de la promotion et de l'accessibilité du patrimoine culturel matériel, documentaire et immatériel, en particulier les langues en péril et les langues et savoirs autochtones, en mettant en place ou en actualisant par exemple les programmes éducatifs concernant l'application de systèmes d'IA dans ces domaines, selon les besoins et suivant une approche participative, à l'intention des institutions et du public.

96. Les États membres sont encouragés à examiner et à traiter les répercussions des systèmes d'IA sur la culture, en particulier celles des applications de traitement du langage naturel, comme la traduction automatique et les assistants vocaux, sur les nuances de l'expression et du langage humains. Ces évaluations devraient permettre de contribuer à la conception et à la mise en œuvre de stratégies visant à tirer parti au maximum des avantages de ces systèmes en réduisant les inégalités culturelles et en améliorant la compréhension humaine, ainsi que de comprendre leurs effets négatifs, tels qu'un recul de l'usage qui pourrait entraîner la disparition de langues en péril, de dialectes locaux et de variations tonales et culturelles associées à l'expression et au langage humains.

97. Les États membres devraient promouvoir l'éducation à l'IA et des formations numériques à l'intention des artistes et des professionnels de la création afin qu'ils soient capables d'évaluer la pertinence de l'utilisation des technologies d'IA dans leurs métiers, les technologies d'IA étant actuellement utilisées pour créer, produire, distribuer et diffuser toute une variété de biens et services culturels, en gardant à l'esprit l'importance de la préservation du patrimoine culturel, de la diversité et de la liberté artistique.

98. Les États membres devraient faire mieux connaître les outils d'IA et promouvoir leur évaluation parmi les industries culturelles locales et les petites et moyennes entreprises travaillant dans le domaine de la culture, afin d'éviter le risque de concentration sur le marché culturel.

99. Les États membres devraient faire appel aux grandes entreprises spécialisées dans les technologies et à d'autres parties prenantes afin de favoriser une offre diversifiée et une pluralité d'accès en matière d'expressions culturelles, en faisant notamment en sorte que la recommandation algorithmique améliore la visibilité des contenus locaux et les rende plus faciles à trouver.

100. Les États membres devraient encourager les nouvelles recherches à la croisée de l'IA et de la propriété intellectuelle, afin de déterminer par exemple qui sont les détenteurs des

droits sur les œuvres créées au moyen de technologies d'IA parmi les différentes parties prenantes intervenant tout au long du cycle de vie de l'IA.

101. Les États membres devraient encourager les musées, les galeries, les bibliothèques et les archives à l'échelle nationale à mettre au point et à utiliser des systèmes d'IA afin de valoriser leurs collections, d'étoffer leurs bases de données et de permettre à leurs usagers d'y accéder.

## **DOMAINE STRATÉGIQUE 8 : ÉDUCATION ET RECHERCHE**

102. Les États membres devraient collaborer avec des organisations internationales et des organismes privés et non gouvernementaux, afin de transmettre au grand public de tous les pays les connaissances nécessaires à la maîtrise de l'IA, de façon à autonomiser la population et à réduire la fracture numérique et les inégalités d'accès dans le domaine du numérique découlant de l'adoption à grande échelle de systèmes d'IA.

103. Les États membres devraient encourager l'acquisition de « compétences préalables » à l'éducation à l'IA, telles que les compétences de base en lecture, en écriture, en calcul, en programmation et en technologie numérique, l'éducation aux médias et à l'information, ainsi que les compétences relatives à l'esprit critique, au travail en équipe et à la communication, les compétences socioémotionnelles et les compétences en matière d'éthique de l'IA, en particulier dans les pays qui présentent des lacunes notables dans l'enseignement de toutes ces compétences.

104. Les États membres devraient promouvoir des programmes généraux de sensibilisation aux avancées de l'IA, notamment concernant les possibilités et les défis découlant des technologies de l'IA. Ces programmes devraient être accessibles aux spécialistes comme aux non-spécialistes.

105. Les États membres devraient encourager les initiatives de recherche portant sur l'utilisation responsable des technologies d'IA dans l'enseignement, la formation des enseignants et l'apprentissage à distance, entre autres sujets, de manière à élargir les possibilités et à atténuer les problèmes et les risques qui existent dans ce domaine. Ces initiatives devraient s'accompagner d'une évaluation adéquate de la qualité de l'éducation et de l'impact sur les élèves et les enseignants de l'utilisation des technologies d'IA. Les États membres devraient également faire en sorte que les technologies d'IA autonomisent les élèves et les enseignants et améliorent leur expérience, tout en gardant à l'esprit que les dimensions affectives et sociales et la valeur des formes traditionnelles d'enseignement sont essentielles dans les relations enseignant/élève et élève/élève, et qu'elles devraient être prises en compte lorsque l'on examine l'adoption de technologies d'IA dans l'éducation.

106. Les États membres devraient favoriser la participation des filles et des femmes, des personnes d'origines ethniques et culturelles diverses ainsi que des personnes handicapées aux programmes d'éducation à l'IA à tous les niveaux, ainsi que le suivi et le partage des meilleures pratiques à cet égard avec d'autres États membres.

107. Les États membres devraient mettre au point, conformément à leurs traditions et à leurs programmes d'éducation nationale, des programmes d'enseignement de l'éthique de l'IA pour tous les niveaux, et favoriser la collaboration croisée entre l'enseignement de compétences techniques dans le domaine de l'IA et les aspects humanistes, éthiques et sociaux de l'éducation à l'IA. Des cours en ligne et des ressources numériques concernant l'éducation à l'éthique de l'IA devraient être élaborés dans les langues locales, en particulier dans des formats accessibles aux personnes handicapées.

108. Les États membres devraient promouvoir la recherche sur l'éthique de l'IA, soit en finançant ce type de recherche, soit en prenant des mesures visant à inciter les secteurs public et privé à investir dans ce domaine.

109. Les États membres devraient veiller à ce que les chercheurs en IA possèdent une formation en éthique de la recherche et exiger d'eux qu'ils tiennent compte de considérations éthiques dans leurs conceptions, leurs produits et leurs publications, en particulier dans les analyses d'ensembles de données qu'ils utilisent, la façon dont ces ensembles sont annotés ainsi que la qualité et la portée des résultats.

110. Les États membres devraient encourager les sociétés du secteur privé à faciliter l'accès de la communauté scientifique à leurs données à des fins de recherche, notamment dans les PRITI, en particulier les PMA, les PDSL et les PEID. Cet accès ne doit pas être accordé au détriment de la confidentialité.

111. Les États membres devraient promouvoir la diversité de genre dans la recherche sur l'IA, dans le milieu universitaire et dans l'industrie, en prenant des mesures qui incitent les filles et les femmes à s'engager dans ce domaine, en mettant en place des mécanismes permettant de lutter contre les stéréotypes sexistes et le harcèlement au sein de la communauté des chercheurs en IA, et en encourageant les entités universitaires et privées à partager leurs meilleures pratiques sur la façon de renforcer la diversité de genre.

112. Pour garantir une évaluation critique de la recherche sur l'IA et un suivi adéquat des éventuels mauvais usages et effets néfastes, les États membres devraient veiller à ce que toute future avancée concernant les technologies d'IA soit fondée sur des recherches scientifiques rigoureuses, et promouvoir la recherche interdisciplinaire sur l'IA en faisant appel à des disciplines autres que les sciences, la technologie, l'ingénierie et les mathématiques (STEM), comme les études culturelles, l'enseignement, l'éthique, les relations internationales, le droit, la linguistique, la philosophie et les sciences politiques.

113. Reconnaissant que les technologies d'IA offrent d'importants moyens de faire progresser la connaissance et la pratique scientifiques, notamment dans les disciplines traditionnellement fondées sur des modèles, les États membres devraient encourager les communautés scientifiques à mesurer les avantages, les limites et les risques que comporte leur utilisation. Il s'agit notamment de veiller à ce que les conclusions tirées d'approches axées sur les données soient solides et fiables. Par ailleurs, les États membres devraient accueillir avec satisfaction et soutenir l'apport de la communauté scientifique dans la contribution aux politiques et la prise de conscience des atouts et des faiblesses des technologies d'IA.

## **DOMAINE STRATÉGIQUE 9 : ÉCONOMIE ET TRAVAIL**

114. Les États membres devraient évaluer et traiter l'impact des systèmes d'IA sur le marché du travail ainsi que son incidence sur les besoins en matière d'éducation dans tous les pays, et plus particulièrement dans ceux dont l'économie nécessite une main-d'œuvre importante. Cela peut nécessiter la mise en place, à tous les niveaux d'éducation, d'un éventail plus large de compétences interdisciplinaires « de base » pour donner aux travailleurs actuels et aux nouvelles générations une chance équitable de trouver un emploi sur un marché à l'évolution rapide, et pour garantir leur sensibilisation aux aspects éthiques des systèmes d'IA. En plus des compétences techniques et spécialisées et des tâches requérant peu de qualifications, telles que l'étiquetage des ensembles de données, il convient d'enseigner des compétences comme « apprendre à apprendre », la communication, la réflexion critique, le travail d'équipe, l'empathie et la capacité de transférer ses connaissances d'un domaine à l'autre. Il est essentiel de faire preuve de transparence quant aux compétences recherchées et d'actualiser les programmes d'études en fonction de celles-ci.

115. Les États membres devraient soutenir les accords de coopération entre les autorités, les établissements universitaires, les entreprises, les organisations de travailleurs et la société civile afin de combler les lacunes en matière de compétences exigées et d'aligner les stratégies et programmes de formation sur les futurs enjeux du travail et sur les besoins des employeurs. Il convient de promouvoir des approches de l'enseignement et l'apprentissage de l'IA qui soient basées sur des projets, ce qui permet d'établir des partenariats entre les entreprises du secteur privé, les universités et les centres de recherche.

116. Les États membres devraient coopérer avec des entreprises du secteur privé, des organisations de la société civile et d'autres parties prenantes, y compris les travailleurs et les syndicats, pour assurer une transition équitable aux employés menacés. Cela suppose de mettre en place des programmes de perfectionnement et de reconversion, de concevoir des mécanismes efficaces visant à retenir les employés pendant ces périodes de transition, et d'envisager des programmes de « couverture sociale » pour ceux qui ne peuvent pas se reconverter. Les États membres devraient élaborer et mettre en œuvre des programmes destinés à étudier les défis recensés et à y répondre, notamment le perfectionnement et la reconversion, le renforcement de la protection sociale, la mise en œuvre de politiques et d'interventions sectorielles énergiques, ou encore l'introduction d'avantages fiscaux et de nouvelles modalités d'imposition. Les régimes fiscaux et d'autres réglementations pertinentes devraient être examinés avec soin et modifiés, si nécessaire, de façon à contrer les conséquences du chômage engendré par l'automatisation basée sur l'IA.

117. Les États membres devraient encourager et aider les chercheurs à se pencher sur l'impact des systèmes d'IA sur l'environnement de travail local afin d'anticiper les tendances et défis à venir. Il conviendrait d'analyser les répercussions des systèmes d'IA sur les secteurs économiques, sociaux et géographiques, ainsi que sur les interactions entre humains et robots et les relations entre êtres humains, dans le but de formuler des conseils sur les pratiques exemplaires en matière de reconversion et de réaffectation.

118. Les États membres devraient concevoir des mécanismes destinés à empêcher la monopolisation des systèmes d'IA tout au long de leur cycle de vie et les inégalités qui en résulteraient, que les monopoles portent sur les données, la recherche, la technologie, le marché ou d'autres domaines. Les États membres devraient analyser les marchés concernés et réglementer et intervenir si de tels monopoles existent, en tenant compte du fait que, en raison d'un manque d'infrastructures, de capacités humaines et de réglementation, les PRITI, et plus particulièrement les PMA, les PDSL et les PEID, sont plus exposés et plus vulnérables à l'exploitation par les grandes entreprises spécialisées dans la technologie.

## **DOMAINE STRATÉGIQUE 10 : SANTÉ ET BIEN-ÊTRE SOCIAL**

119. Les États membres devraient s'efforcer d'utiliser des systèmes d'IA efficaces pour améliorer la santé humaine et protéger le droit à la vie, tout en développant et maintenant la solidarité internationale en vue de faire face aux risques et incertitudes sanitaires mondiaux. Ils devraient en outre s'assurer que le déploiement des systèmes d'IA dans le domaine des soins de santé se fait dans le respect du droit international et des règles, normes et principes du droit international des droits de l'homme. Les États membres devraient veiller à ce que les acteurs impliqués dans les systèmes d'IA relatifs aux soins de santé tiennent compte de l'importance de la relation des patients avec leur famille et le personnel soignant.

120. Les États membres devraient réglementer le développement et le déploiement des systèmes d'IA relatifs à la santé en général et à la santé mentale en particulier, afin de garantir leur sécurité, leur efficacité et leur fiabilité sur les plans scientifique et médical. Par ailleurs, dans le domaine connexe des interventions en cybersanté, les États membres sont vivement encouragés à faire participer activement les patients et leurs représentants à toutes les étapes pertinentes du développement du système.

121. Les États membres devraient accorder une attention particulière à la réglementation des solutions de prévision, de détection et de thérapie médicales contenues dans les applications d'IA, par les moyens suivants :

- (a) assurer une surveillance afin de réduire les biais ;
- (b) garantir que le professionnel, le patient, le soignant ou l'utilisateur du service est intégré à l'équipe en tant que « spécialiste du domaine » lors de l'élaboration des algorithmes ;
- (c) tenir compte du respect de la vie privée, étant donné qu'une surveillance constante peut être nécessaire ;
- (d) veiller à ce que les personnes dont les données sont analysées sachent que leurs données sont surveillées et analysées, et à ce qu'elles donnent leur consentement éclairé à ce sujet ;
- (e) garantir que les soins humains et la prise de décisions définitives relatives au diagnostic et au traitement sont assurés par des êtres humains, tout en reconnaissant que les systèmes d'IA peuvent les y aider.

122. Les États membres devraient mener des recherches sur les effets et la réglementation des dégâts potentiels des systèmes d'IA sur la santé mentale, comme l'augmentation de la dépression, de l'anxiété, de l'isolement social, de la dépendance, du trafic et de la radicalisation, ou encore de la désinformation.

123. Les États membres devraient élaborer des principes directeurs sur les interactions entre humains et robots et leurs conséquences sur les relations entre êtres humains, qui reposeraient sur la recherche et seraient destinés au futur développement des robots, et accorderaient une attention particulière à la santé mentale et physique des êtres humains, notamment pour les robots infirmiers et d'assistance aux personnes âgées et aux personnes handicapées, et pour les robots à usage éducatif, ludique, conversationnel et de compagnie destinés aux enfants et aux adultes. En outre, il convient de faire appel aux technologies d'IA pour améliorer la sécurité et l'ergonomie des robots, y compris dans les environnements de travail associant humains et robots.

124. Les États membres devraient garantir que les interactions entre humains et robots se font dans le respect des valeurs et principes applicables à tout autre système d'IA, ce qui inclut les droits de la personne, la promotion de la diversité dans les relations et la protection des groupes vulnérables.

125. Les États membres devraient protéger le droit des utilisateurs de déterminer facilement s'ils interagissent avec un être vivant ou avec un système d'IA imitant des caractéristiques humaines ou animales.

126. Les États membres devraient instaurer des politiques visant à sensibiliser à l'anthropomorphisation des technologies d'IA, notamment en ce qui concerne les termes utilisés pour les désigner, et évaluer les manifestations, les conséquences éthiques et les possibles limites de ce phénomène, en particulier dans le contexte des interactions entre robots et humains, et plus spécialement lorsqu'il s'agit d'enfants.

127. Les États membres devraient encourager et promouvoir la recherche collaborative sur les effets d'une utilisation à long terme des systèmes d'IA par les individus, en prêtant une attention spéciale aux conséquences psychologiques et cognitives potentielles de ces systèmes sur les enfants et les jeunes. Cette recherche devrait se fonder sur de multiples

normes, principes, protocoles, approches disciplinaires, sur une analyse de la modification des comportements et des habitudes, ainsi que sur une évaluation minutieuse des impacts culturels et sociétaux en aval.

128. Les États membres et les parties prenantes devraient établir des mécanismes visant à faire véritablement participer les enfants et les jeunes aux discussions, débats et prises de décisions relatifs aux répercussions des systèmes d'IA sur leur vie et leur avenir.

129. Les États membres devraient favoriser l'utilisation responsable des systèmes d'IA pour lutter contre le discours de haine en ligne et la désinformation, et veiller à ce qu'ils ne soient pas utilisés pour produire et diffuser ce type de contenus, en particulier en période électorale.

130. Les États membres devraient créer des environnements favorables aux médias, dans lesquels ces derniers disposeraient des droits et ressources nécessaires pour rendre compte efficacement des avantages et des inconvénients des systèmes d'IA, et utiliseraient ces systèmes dans leur travail.

## **V. SUIVI ET ÉVALUATION**

131. Les États membres – en fonction de leur situation, de leur mode de gouvernement et de leur Constitution – devraient assurer de manière crédible et transparente le suivi et l'évaluation des politiques, programmes et mécanismes relatifs à l'éthique de l'IA en combinant des approches quantitatives et qualitatives. L'UNESCO peut les aider en :

- (a) élaborant une méthodologie d'évaluation des conséquences éthiques (ECE) des technologies d'IA mondialement reconnue, notamment des lignes directrices concernant sa mise en œuvre à toutes les étapes du cycle de vie des systèmes d'IA, en s'appuyant sur des travaux scientifiques rigoureux ;
- (b) mettant au point une méthodologie de préparation pour aider les États membres à déterminer leur état d'avancement, à divers égards, à des moments précis de leur parcours de préparation ;
- (c) concevant une méthodologie mondialement reconnue pour évaluer en amont et en aval l'efficacité et l'efficience des politiques et des mesures d'incitation relatives à l'éthique de l'IA par rapport aux objectifs fixés ;
- (d) renforçant l'analyse fondée sur la recherche et sur des éléments factuels des politiques relatives à l'éthique de l'IA ainsi que l'établissement de rapports à ce sujet, notamment par la publication d'un indice comparatif ;
- (e) collectant et diffusant des informations sur les progrès et les innovations réalisés, des rapports de recherche, des publications scientifiques, des données et des statistiques concernant l'éthique de l'IA, afin d'encourager l'échange de pratiques exemplaires et l'apprentissage mutuel et de contribuer à l'application de la présente Recommandation.

132. Les processus de suivi et d'évaluation devraient assurer une large participation des parties prenantes concernées, notamment, mais pas exclusivement, des personnes des différents groupes d'âge, des filles et des femmes, des personnes handicapées, des populations défavorisées, marginalisées et vulnérables, des communautés autochtones, ainsi que de personnes issues de milieux socioéconomiques divers. Il convient de garantir une diversité sur les plans social, culturel et du genre, en vue d'améliorer les processus d'apprentissage et de renforcer les liens entre constatations, prises de décisions, transparence et obligation de rendre des comptes concernant les résultats.

133. Dans le but de promouvoir les politiques et pratiques exemplaires en matière d'éthique de l'IA, il conviendrait d'élaborer des outils et indicateurs appropriés pour évaluer leur efficacité et leur efficience par rapport aux normes, priorités et cibles convenues – y compris des cibles spécifiques pour les personnes appartenant aux groupes défavorisés, marginalisés et vulnérables –, ainsi que les conséquences des systèmes d'IA aux niveaux individuel et sociétal. Le suivi et l'évaluation de l'impact des systèmes d'IA et des politiques et pratiques connexes relatives à l'éthique de l'IA devraient être assurés en permanence et de façon systématique. Ce processus devrait s'appuyer sur des cadres convenus au niveau international et s'accompagner d'évaluations des établissements, prestataires et programmes privés et publics, y compris des auto-évaluations, ainsi que des enquêtes de suivi et la mise au point d'une batterie d'indicateurs. La collecte et le traitement des données devraient être menés conformément à la législation nationale sur la protection et la confidentialité des données.

134. Les mécanismes possibles de suivi et d'évaluation peuvent inclure un observatoire de l'éthique de l'IA, ou des contributions aux initiatives déjà en place pour renforcer la conformité aux principes éthiques dans l'ensemble des domaines de compétence de l'UNESCO, un mécanisme de partage d'expériences permettant aux États membres de réagir à leurs initiatives respectives, des bacs à sable réglementaires sur l'IA, et un guide d'évaluation permettant à tous les acteurs de l'IA de déterminer à quel point ils respectent les recommandations stratégiques mentionnées dans le présent document.

## **VI. UTILISATION ET MISE EN ŒUVRE DE LA PRÉSENTE RECOMMANDATION**

135. Les États membres et toutes les autres parties prenantes identifiées dans la présente Recommandation doivent respecter, promouvoir et protéger les principes et normes éthiques relatifs à l'IA qui y sont énoncés, et devraient prendre toutes les mesures en leur pouvoir pour donner effet aux recommandations stratégiques qu'elle contient.

136. Les États membres devraient s'efforcer d'élargir et de compléter leur propre action en ce qui concerne la présente Recommandation en coopérant avec toutes les organisations nationales et internationales, gouvernementales et non gouvernementales, les sociétés transnationales et les organisations scientifiques dont les activités sont en rapport avec le champ d'application et les objectifs de la présente Recommandation. L'élaboration d'une méthodologie d'évaluation des conséquences éthiques mondialement reconnue et la création de commissions nationales d'éthique des technologies peuvent constituer d'importants outils à cet égard.

## **VII. PROMOTION DE LA PRÉSENTE RECOMMANDATION**

137. L'UNESCO a vocation à être la principale institution des Nations Unies chargée de promouvoir et de diffuser la présente Recommandation et, par conséquent, devrait travailler en collaboration avec d'autres entités du système des Nations Unies, notamment le Groupe de haut niveau du Secrétaire général de l'ONU sur la coopération numérique, la Commission mondiale d'éthique des connaissances scientifiques et des technologies (COMEST), le Comité international de bioéthique (CIB), le Comité intergouvernemental de bioéthique (CIGB), l'Union internationale des télécommunications (UIT), l'Organisation internationale du Travail (OIT), l'Organisation mondiale de la propriété intellectuelle (OMPI), le Fonds des Nations Unies pour l'enfance (UNICEF), ONU-Femmes, l'Organisation des Nations Unies pour le développement industriel (ONUDI), l'Organisation mondiale du commerce (OMC) et d'autres entités compétentes des Nations Unies concernées par l'éthique de l'IA.

138. L'UNESCO devrait également travailler en collaboration avec d'autres organisations internationales et régionales, notamment l'Union africaine (UA), l'Alianza del Pacífico, l'Association des universités africaines (AUA), l'Association des nations de l'Asie du Sud-Est

(ASEAN), la Communauté des Caraïbes (CARICOM), l'Union des télécommunications des Caraïbes, l'Association des services publics des Caraïbes, le Marché commun de l'Afrique orientale et australe (COMESA), la Communauté des États latino-américains et caribéens (CELAC), le Conseil de l'Europe (COE), la Communauté économique des États de l'Afrique de l'Ouest (CEDEAO), l'Union économique eurasiatique (UEE), l'Union européenne (UE), l'Association internationale des universités (AIU), l'Organisation de coopération et de développement économiques (OCDE), l'Organisation pour la sécurité et la coopération en Europe (OSCE), l'Association sud-asiatique de coopération régionale (ASACR), la Communauté de développement de l'Afrique australe (SADC), le Marché commun du Sud (MERCOSUR), ou encore l'Institute of Electrical and Electronic Engineers (IEEE), l'Organisation internationale de normalisation (ISO) et des institutions financières internationales telles que la Banque mondiale, la Banque interaméricaine de développement et la Banque africaine de développement.

139. Bien qu'à l'UNESCO, le mandat de promouvoir et protéger relève de la compétence des gouvernements et des organes intergouvernementaux, la société civile jouera un rôle important pour défendre les intérêts du secteur public. L'UNESCO doit par conséquent asseoir et mettre en avant sa légitimité.

### **VIII. DISPOSITIONS FINALES**

140. La présente Recommandation doit s'entendre comme un tout, et les valeurs et principes fondamentaux comme étant complémentaires et interdépendants.

141. Aucune disposition de la présente Recommandation ne peut être interprétée comme autorisant un État, tout autre acteur de la vie sociale, un groupe ou un individu à se livrer à une activité ou à accomplir un acte contraire aux droits de l'homme, aux libertés fondamentales, à la dignité humaine et au souci de l'environnement et des écosystèmes.