

GOUVERNER L'IA AU BÉNÉFICE DE L'HUMANITÉ



United
Nations



AI
Advisory
Body



Gouverner l'IA au bénéfice de l'humanité : Rapport final

Copyright © Nations Unies, 2024
Tous droits réservés.

La présente publication dans son intégralité ne peut être reproduite ni transmise sous quelque forme ou par quelque moyen que ce soit, électronique ou mécanique, y compris la photocopie, l'enregistrement ou tout autre système informatisé de mise en mémoire ou de recherche des données connu ou encore à inventer, sans la permission écrite de l'éditeur.

Les demandes de reproduction d'extraits ou de photocopie doivent être adressées au Copyright Clearance Center, à l'adresse copyright.com.

Toute autre demande de renseignement sur les droits et licences, y compris les droits dérivés, doit être envoyée à l'adresse suivante : United Nations Publications, 405 East 42nd Street, S-011FW001, New York, NY 10017, États-Unis d'Amérique.
Courriel : permissions@un.org ; site Web : shop.un.org.

Les appellations employées dans cette publication et la présentation des données qui y figurent n'impliquent de la part du Secrétariat de l'Organisation des Nations Unies aucune prise de position quant au statut juridique des pays, territoires, villes ou zones ou de leurs autorités, ni quant au tracé de leurs frontières ou limites.

eISBN: 9789211067880

Gouverner l'IA au bénéfice de l'humanité

Rapport final



**United
Nations**



**AI
Advisory
Body**

À propos de l'Organe consultatif de haut niveau sur l'intelligence artificielle

Initialement proposé en 2020 dans le cadre du Plan d'action de coopération numérique (A/74/821), l'Organe consultatif de haut niveau multipartite sur l'intelligence artificielle a été créé en octobre 2023 en vue d'analyser la gouvernance internationale de l'intelligence artificielle et de formuler des recommandations.

Les membres de l'Organe sont nommés à titre individuel et ne représentent pas leur organisation. Le présent rapport d'étape est l'expression d'un consensus majoritaire ; aucun membre n'est censé approuver chacun des points contenus dans le document. Les membres affirment qu'ils approuvent de manière générale, mais non unilatérale, les conclusions et recommandations formulées. Les termes utilisés dans le présent rapport n'impliquent pas l'approbation institutionnelle des différentes organisations dont sont issus les membres.

Table des matières

À propos de l'Organe consultatif de haut niveau sur l'intelligence artificielle	4
Résumé	7
1. La nécessité d'une gouvernance mondiale	7
2. Lacunes dans la gouvernance mondiale de l'IA	8
3. Renforcement de la coopération mondiale	9
A. Vision commune	10
B. Socle commun	11
C. Avantages communs	15
D. Stratégie cohérente	22
E. Réflexions sur les modèles institutionnels	24
4. Appel à l'action	24
1. Introduction	26
A. Des débouchés et des catalyseurs	28
B. Les catalyseurs de l'IA au service de l'humanité	32
C. La gouvernance : un catalyseur essentiel	32
D. Risques et difficultés	32
E. Risques liés à l'IA	32
F. Difficultés à résoudre	38
2. La nécessité d'une gouvernance mondiale	42
A. Principes directeurs et fonctions de la gouvernance internationale de l'IA	43
B. Paysage international émergent de la gouvernance de l'IA	46

3. Lacunes dans la gouvernance mondiale de l'IA	48
A. Lacunes en matière de représentation	48
B. Lacunes en matière de coordination	50
C. Lacunes en matière d'exécution	52
4. Renforcement de la coopération mondiale	54
A. Vision commune	55
Groupe scientifique international sur l'IA	55
B. Socle commun	60
Concertation sur les mesures à prendre en matière de gouvernance de l'IA	52
Centre d'échange de normes en matière d'IA	64
C. Avantages communs	66
Réseau de développement des capacités	74
Fonds mondial pour l'IA	76
Cadre mondial de données sur l'IA	78
D. Stratégie cohérente	82
Bureau de l'AI au Secrétariat	82
E. Réflexions sur les modèles institutionnels	85
Un organisme international de l'IA ?	86
5. Conclusion : appel à l'action	90
Annexes	92
Annexe A : membres de l'Organe consultatif de haut niveau sur l'intelligence artificielle	92
Annexe B : mandat de l'Organe consultatif de haut niveau sur l'intelligence artificielle	93
Annexe C : liste des consultations tenues en 2024	94
Annexe D : liste des études approfondies	95
Annexe E : réponses à l'enquête sur les risques	96
Annexe F : réponses à l'analyse des opportunités	106
Annexe G : liste des abréviations	112

Résumé

- i** L'intelligence artificielle (IA) est en train de transformer notre monde. Les avantages que ce dispositif technologique peut procurer sont colossaux, qu'il s'agisse d'ouvrir de nouveaux domaines de recherche scientifique, d'optimiser l'utilisation des réseaux énergétiques, d'améliorer la santé publique et l'agriculture, ou encore de promouvoir plus largement la réalisation des objectifs de développement durable.
- ii** Toutefois, en l'absence de gouvernance, les possibilités offertes par l'IA risquent de ne pas se concrétiser ou de ne pas être réparties équitablement. Si la fracture numérique continue de se creuser, seule une poignée d'États, d'entreprises et de personnes pourraient bénéficier de ces avantages. Les utilisations manquées, c'est-à-dire le fait de ne pas exploiter ou diffuser les avantages liés à l'IA en raison d'un manque de confiance ou de l'absence de catalyseurs tels que le manque de capacités et une gouvernance inefficace, pourraient limiter le champ des possibilités.
- iii** L'IA présente également d'autres risques. Outre les biais et la surveillance de l'IA, de nouvelles préoccupations se font jour telles que les confabulations (ou « hallucinations ») des grands modèles de langage, la création et la diffusion de désinformation renforcée par l'IA, les risques pour la paix et la sécurité, et la consommation d'énergie des systèmes d'IA dans un contexte marqué par la crise climatique.
- iv** Les systèmes d'IA rapides, opaques et autonomes posent des problèmes pour les systèmes réglementaires traditionnels, tandis que des systèmes toujours plus puissants pourraient bouleverser le monde du travail. Les utilisations

de l'IA liées aux armes autonomes et à la sécurité publique soulèvent d'importantes questions juridiques, sécuritaires et humanitaires.

- v** Il existe aujourd'hui un manque de gouvernance mondiale en matière d'IA. Malgré de nombreux débats sur l'éthique et les principes, la mosaïque de normes et d'organismes est encore peu développée et présente de grandes lacunes. Le principe de responsabilité brille souvent par son absence, notamment lorsqu'il s'agit de déployer des systèmes d'IA incompréhensibles qui ont des effets sur autrui. Le respect des règles repose souvent sur le volontariat ; la pratique dément la rhétorique.
- vi** Comme indiqué dans notre rapport d'étape¹, la gouvernance de l'IA est cruciale, non seulement pour traiter les difficultés et les risques, mais aussi pour veiller à exploiter ses atouts sans laisser personne de côté.

1. La nécessité d'une gouvernance mondiale

- vii** La nécessité de mettre en place une gouvernance mondiale, en particulier, est incontestable. Les matières premières de l'IA, qu'il s'agisse de minéraux essentiels ou de données d'entraînement proviennent du monde entier. L'IA à usage général, déployée au-delà des frontières, génère de nombreuses applications à travers le monde. Le développement accéléré de l'IA concentre le pouvoir et la richesse à l'échelle mondiale, ce qui a des conséquences géopolitiques et géoéconomiques.

1 Voir <https://un.org/ai-advisory-body>.

- viii** En outre, personne n'est actuellement en mesure de comprendre tous les rouages de l'IA au point de contrôler pleinement ses résultats ou de prédire son évolution. La responsabilité des décideurs n'est pas non plus engagée dans la mise au point, le déploiement ou l'utilisation de systèmes qu'ils ne maîtrisent pas. Par ailleurs, les répercussions négatives, y compris indirectes, que ces décisions engendrent peuvent également avoir une portée mondiale.
- ix** La mise au point, le déploiement et l'utilisation de la technologie ne peuvent être laissés aux aléas des marchés. Les gouvernements nationaux et les organisations régionales auront un rôle crucial à jouer mais le caractère transfrontière de la structure et de l'application de la technologie appelle une approche globale. La gouvernance peut également être un moteur essentiel de l'innovation en matière d'IA au service de la réalisation des objectifs de développement durable à l'échelle mondiale.
- x** L'IA présente donc des obstacles à surmonter et des chances à saisir qui nécessitent une approche globale pour traiter de manière transversale des domaines tels que la politique, l'économie, la société, l'éthique, les droits humains, la technique et l'environnement. Cette approche peut transformer une mosaïque d'initiatives en constante évolution en un dispositif cohérent et interopérable, fondé sur le droit international et les objectifs de développement durable, adaptable à tous les contextes et à toutes les époques.
- xi** Dans notre rapport d'étape, nous avons exposé des principes² qui devraient orienter la création de nouveaux organismes internationaux de gouvernance de l'IA. Ces principes tiennent compte du fait que la gouvernance de l'IA n'opère pas en vase clos et que le droit international, en particulier le droit international des droits de l'homme, s'applique à l'IA.

2. Lacunes dans la gouvernance mondiale de l'IA

- xii** Les documents et débats consacrés à la gouvernance de l'IA ne manquent pas. Des gouvernements, des entreprises et des consortiums, ainsi que des organisations régionales et internationales ont adopté des centaines de guides, de cadres et de principes.
- xiii** Cependant, aucun de ces dispositifs n'a véritablement de portée mondiale et de couverture complète, ce qui pose des problèmes de représentation, de coordination et de mise en œuvre.
- xiv** S'agissant de la représentation, des régions entières du monde ont été exclues des débats sur la gouvernance internationale de l'IA. La figure a) présente sept initiatives de premier plan en matière d'IA ne relevant pas de l'ONU³. Sept pays participent à toutes les initiatives données en exemple, tandis que 118 ne participent à aucune d'entre elles (principalement dans le monde du Sud).
- xv** L'équité exige qu'un plus grand nombre de voix jouent un rôle majeur dans les décisions portant sur la gouvernance des technologies qui nous concernent. La concentration du pouvoir de décision dans le secteur des technologies de l'IA n'est pas justifiée. Nous devons également reconnaître que par le passé de nombreuses populations ont été totalement exclues de débats sur la gouvernance de l'IA qui les concernaient.

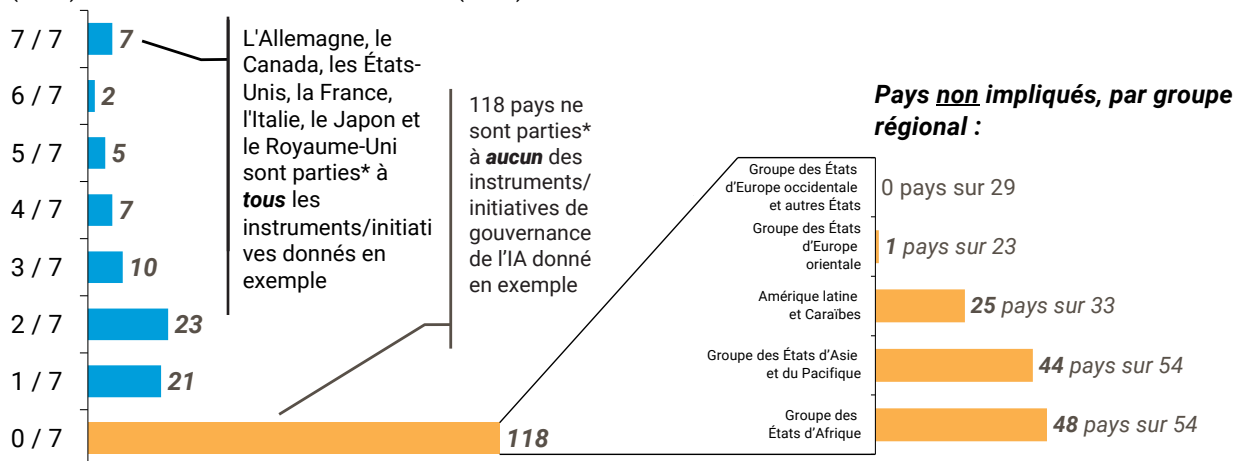
2 Principe directeur 1 : l'IA doit être gouvernée de manière inclusive, par toutes et par tous et dans l'intérêt commun ; principe directeur 2 : l'IA doit être gouvernée dans l'intérêt général ; principe directeur 3 : la gouvernance de l'IA doit être articulée avec la gouvernance des données et la promotion des biens communs en matière de données ; principe directeur 4 : la gouvernance de l'IA doit être universelle, en réseau et ancrée dans une collaboration multipartite adaptative ; principe directeur 5 : la gouvernance de l'IA devrait prendre ses racines dans la Charte des Nations Unies, le droit international des droits de l'homme et d'autres engagements convenus au niveau international, tels que les objectifs de développement durable.

3 À l'exclusion de la Recommandation formulée par l'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO) sur l'éthique de l'intelligence artificielle (2021) et de deux résolutions sur l'IA adoptées par l'Assemblée générale en 2024 : « Saisir les possibilités offertes par des systèmes d'intelligence artificielle sûrs, sécurisés et dignes de confiance pour le développement durable » (78/265) et « Intensifier la coopération internationale en matière de renforcement des capacités dans le domaine de l'intelligence artificielle » (78/311).

Figure a) Représentation de sept initiatives internationales de gouvernance de l'IA ne relevant pas de l'ONU

Exemple : Principes de l'OCDE en matière d'IA (2019), principes du G20 en matière d'IA (2019), groupe de rédaction de la convention du Conseil de l'Europe sur l'IA (2022-2024), Déclaration ministérielle du PMIA (2022), Déclaration des Ministres du G7 (2023), Déclaration de Bletchley (2023) et Déclaration ministérielle de Séoul (2024).

**Interrégional
uniquement, excluant le
niveau régional**



*Par approbation des textes intergouvernementaux pertinents. Les pays ne sont pas considérés comme participant à une initiative plurilatérale du seul fait de leur appartenance à l'Union européenne ou à l'Union africaine. **Abréviations :** Groupe des États d'Afrique ; Groupe des États d'Asie et du Pacifique ; Groupe des États d'Europe orientale ; G20, Groupe des 20 ; G7, Groupe des Sept ; PMIA, Partenariat mondial sur l'intelligence artificielle ; Amérique latine et Caraïbes ; OCDE, Organisation de coopération et de développement économiques ; Groupe des États d'Europe occidentale et autres États.

xvi Les régimes de gouvernance de l'IA doivent également s'étendre à l'ensemble du globe pour être efficaces, notamment éviter les « courses à l'armement de l'IA » ou le nivellement par le bas en matière de sécurité et de droits, détecter et traiter les incidents résultant de décisions prises tout au long du cycle de vie de l'IA qui relèvent de plusieurs pays, stimuler l'apprentissage, encourager l'interopérabilité et mettre en commun les avantages de l'IA. La technologie ne connaît pas de frontières et, à mesure qu'elle se diffuse, l'illusion qu'un État ou un groupe d'États puissent (ou doivent) la contrôler s'estompe.

xvii Le manque de coordination entre les initiatives et les organismes risque de diviser le monde en régimes de gouvernance de l'IA déconnectés et incompatibles. La coordination fait également défaut au sein du système des Nations Unies. Bien que de nombreuses entités des Nations Unies abordent la gouvernance de l'IA, leurs mandats ne leur permettent pas de la traiter dans sa globalité.

xviii Toutefois, la représentation et la coordination ne suffisent pas. La responsabilisation exige de mettre en œuvre les engagements pris en matière de gouvernance mondiale de l'IA afin de produire des résultats concrets, notamment en ce qui concerne le développement des capacités et le soutien aux petites et moyennes entreprises, de sorte que chacun puisse tirer parti des possibilités qui s'offrent à lui. Une grande partie se déroulera aux niveaux national et régional, mais des efforts supplémentaires devront également être déployés à l'échelle mondiale pour faire face aux risques et exploiter les résultats.

3. Renforcement de la coopération mondiale

xix Nos recommandations mettent en avant une vision globale qui favorise une gouvernance de l'IA souple, agile et interconnectée à l'échelle mondiale au bénéfice de l'humanité et qui repose sur une vision

commune, un socle commun et des avantages communs. Seule une approche inclusive et globale de la gouvernance de l'IA peut relever les défis et saisir les avantages multiples et évolutifs de l'IA à l'échelle mondiale et, ce faisant, promouvoir la stabilité internationale et un développement équitable.

xx Guidées par les principes établis dans notre rapport d'étape, nos propositions visent à combler les lacunes et à apporter de la cohérence à l'écosystème de mesure et d'initiatives internationales en matière de gouvernance de l'IA qui évolue rapidement, de manière à éviter la fragmentation et les occasions manquées. Afin de soutenir efficacement ces mesures et d'établir des partenariats utiles avec d'autres organismes, nous proposons une structure légère et agile dans le cadre d'une stratégie cohérente : un bureau de l'IA au Secrétariat de l'ONU, proche du Secrétaire général, qui servirait de « ciment » pour fédérer les initiatives proposées ici de manière efficace et durable.

A. Vision commune

xxi Une approche globale de la gouvernance de l'IA commence par une vision commune de ses capacités, possibilités, risques et incertitudes. Il convient de disposer en temps utile de connaissances et d'informations scientifiques impartiales et fiables sur l'IA afin que les États Membres puissent élaborer une vision fondamentale commune à l'échelle mondiale et d'équilibrer les asymétries d'information entre les entreprises hébergeant des laboratoires d'IA coûteux et le reste du monde, notamment par l'échange d'informations entre les entreprises d'IA et l'ensemble de la communauté de l'IA.

xxii C'est au niveau mondial que la mise en commun des connaissances scientifiques est la plus efficace car elle permet d'investir conjointement dans un bien public mondial et de favoriser la collaboration d'intérêt public dans le cadre de travaux qui seraient autrement fragmentés et redondants.

Groupe scientifique international sur l'intelligence artificielle

xxiii S'inspirant de précédents tels que le Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC) et le Comité scientifique des Nations Unies pour l'étude des effets des rayonnements ionisants (UNSCEAR), un groupe scientifique international multidisciplinaire sur l'IA pourrait rassembler et catalyser des travaux de recherche de pointe afin de fournir aux scientifiques, aux décideurs politiques, aux États membres et à d'autres parties prenantes des éclairages scientifiques sur la technologie de l'IA ou ses applications à partir d'une source impartiale et crédible.

xxiv Un groupe scientifique placé sous les auspices de l'ONU pourrait apporter des connaissances techniques sur les possibilités qu'offre l'IA, notamment faciliter les études approfondies dans les domaines appliqués des objectifs de développement durable, tels que les soins de santé, l'énergie, l'éducation, la finance, l'agriculture, le climat, le commerce et l'emploi.

xxv Les estimations des risques pourraient également s'appuyer sur les travaux d'autres initiatives de recherche sur l'IA, l'ONU fournissant un espace de confiance unique pour que les chercheurs puissent échanger des idées sur les dernières avancées en la matière. En mettant en commun les connaissances au-delà des silos au sein de pays ou d'entreprises qui pourraient autrement ne pas participer ou être mobilisés, un groupe créé sous l'égide de l'ONU peut contribuer à rectifier les perceptions erronées et à renforcer la confiance à l'échelle mondiale.

xxvi Le groupe devrait fonctionner de manière indépendante, avec le soutien d'une équipe transversale du système des Nations Unies composée de représentants du bureau de l'IA proposé ci-après et des organismes des Nations Unies concernés, tels que l'Union internationale des télécommunications (UIT) et l'UNESCO. Il devrait s'associer aux travaux de recherche menés par d'autres organismes internationaux tels que l'Organisation de coopération et de développement économiques (OCDE) et le Partenariat mondial sur l'intelligence artificielle (PMIA).

Recommandation 1

Groupe scientifique international sur l'intelligence artificielle

Nous recommandons la création d'un groupe scientifique international indépendant sur l'IA, composé de divers experts multidisciplinaires du domaine, participant à titre personnel et sur la base du volontariat. Soutenu par le bureau de l'IA dont la création est proposée au Secrétariat et par d'autres organismes des Nations Unies concernés, en partenariat avec d'autres organisations internationales compétentes, le mandat du groupe comprendrait les éléments suivants :

- a) La publication d'un rapport annuel examinant les capacités, les possibilités, les risques et les incertitudes liés à l'IA ainsi que les domaines de consensus scientifique sur les tendances technologiques et les domaines dans lesquels des travaux de recherche supplémentaires doivent être menés ;
- b) La production de dossiers de recherche thématiques trimestriels sur les domaines dans lesquels l'IA pourrait contribuer à la réalisation des objectifs de développement durable, qui seraient axés sur les domaines d'intérêt général susceptibles d'être négligés ;
- c) La publication de rapports ad hoc sur de nouvelles questions, en particulier l'apparition de nouveaux risques ou de lacunes majeures dans le paysage de la gouvernance.

B. Socle commun

xxvii Outre une vision commune de l'IA, il faut mettre en place un socle commun afin d'établir des approches de gouvernance interopérables ancrées dans des normes et des principes mondiaux, dans l'intérêt de tous les pays. Au niveau mondial, cela permet d'éviter le nivellement par le bas et, en même temps, de réduire les frictions réglementaires transfrontalières, maximiser l'apprentissage et l'interopérabilité technique et répondre efficacement aux défis liés au caractère transfrontalier de l'IA.

Concertation sur les mesures à prendre en matière de gouvernance de l'IA

xxviii Il convient d'établir une instance stratégique inclusive permettant à tous les États Membres, sur la base du savoir-faire des parties prenantes, de mettre en commun des pratiques optimales qui sont fondées sur les droits humains et favorisent le développement, qui encouragent les approches de gouvernance interopérables et qui tiennent compte des défis transfrontaliers qui justifient un examen plus approfondi des politiques. Cette démarche ne

permet pas d'assurer une gouvernance mondiale de tous les aspects de l'IA mais elle peut définir le cadre de la coopération internationale et mieux harmoniser les activités des entreprises et des pays avec les normes et les principes mondiaux.

xxix L'institutionnalisation de ces échanges multipartites sous les auspices de l'ONU peut fournir un cadre fiable et inclusif pour examiner les nouvelles pratiques de gouvernance et les mesures stratégiques adaptées. En dépassant les zones de confort, le dialogue entre des pays n'ayant pas les mêmes sensibilités, ainsi qu'entre les États et les parties prenantes, peut favoriser l'apprentissage et jeter les bases d'une plus grande coopération, notamment en ce qui concerne les normes de sûreté et les droits, ainsi qu'en cas de crise mondiale. Le cadre de l'ONU est essentiel pour ancrer ces travaux dans l'ensemble le plus large possible de normes communes.

xxx Associé au développement des capacités (voir les recommandations 4 et 5), le dialogue inclusif sur les approches de gouvernance peut aider les États et les entreprises à actualiser leurs cadres et méthodologies réglementaires pour faire face à l'accélération de l'IA. Les liens avec le groupe scientifique international renforceraient cette dynamique, comparable au lien entre le GIEC et la Convention-cadre des Nations Unies sur les changements climatiques.

xxxi Une concertation pourrait s'engager en marge des réunions qui se tiennent à New York (par exemple à l'Assemblée générale⁴) et à Genève. Les réunions semestrielles pourraient se concentrer davantage sur les possibilités offertes par les

différents secteurs lors de la première et sur les risques lors de la seconde⁵. À l'avenir, ces réunions permettraient d'échanger des informations sur les incidents liés à l'IA, notamment ceux qui mettent à rude épreuve ou dépassent les capacités des organismes existants.

xxxii Une partie de chaque session de dialogue pourrait se concentrer sur les stratégies nationales adoptées par les États Membres, tandis qu'une seconde partie serait consacrée au savoir-faire et aux contributions des principales parties prenantes, en particulier les entreprises technologiques et les représentants de la société civile. Outre les sessions de dialogue officielles, la mobilisation des parties prenantes sur la politique en matière d'IA pourrait s'appuyer sur d'autres dispositifs existants, plus spécialisés, tels que la réunion AI for Good de l'UIT, la réunion annuelle du Forum sur la gouvernance d'Internet, le Forum mondial de l'UNESCO sur l'éthique de l'IA et l'eWeek de la CNUCED.

4 Analogue au forum politique de haut niveau pour le développement durable dans le contexte des objectifs de développement durable placé sous les auspices du Conseil économique et social.

5 Les parties concernées du système des Nations Unies pourraient être mobilisées afin de mettre en évidence les possibilités et les risques, notamment l'UIT pour les normes en matière d'intelligence artificielle ; l'UIT, la CNUCED, le Programme des Nations Unies pour le développement (PNUD) et le Bureau de la coordination des activités de développement sur les applications de l'IA pour la réalisation des objectifs de développement durable ; l'UNESCO sur l'éthique et la capacité de gouvernance ; le Haut-Commissariat des Nations Unies aux droits de l'homme sur la responsabilité en matière de droits humains sur la base des normes et dispositifs existants ; le Bureau des affaires de désarmement sur la réglementation de l'IA dans les systèmes militaires ; le PNUD sur le soutien aux capacités nationales de développement ; le Forum sur la gouvernance d'Internet pour la collaboration et le dialogue entre les parties prenantes ; l'Organisation mondiale de la propriété intellectuelle (OMPI), l'Organisation internationale du Travail (OIT), l'Organisation mondiale de la Santé (OMS), l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO), le Programme alimentaire mondial (PAM), le Haut-Commissariat des Nations Unies pour les réfugiés, l'UNESCO, le Fonds international des Nations Unies pour l'enfance (UNICEF), l'Organisation météorologique mondiale (OMM) sur les applications sectorielles et la gouvernance.

Recommandation 2

Concertation sur les mesures à prendre en matière de gouvernance de l'IA

Nous recommandons le lancement d'une concertation intergouvernementale et multipartite sur la gouvernance de l'IA, deux fois par an, en marge des réunions existantes à l'ONU. Ses objectifs seraient les suivants :

- a) Échanger les meilleures pratiques en matière de gouvernance de l'IA qui encouragent le développement en même temps qu'elles favorisent le respect, la protection et la concrétisation des droits humains, notamment l'exploitation des possibilités et la gestion des risques ;
- b) Promouvoir une vision commune de la mise en œuvre des mesures de gouvernance de l'IA par les développeurs et les utilisateurs des secteurs privé et public afin d'améliorer l'interopérabilité internationale de la gouvernance de l'IA ;
- c) Faire part à titre volontaire des problèmes importants liés à l'IA qui ont mis à rude épreuve ou dépassé la capacité de réaction des organismes publics ;
- d) Examiner les rapports du groupe scientifique international sur l'IA, selon qu'il convient.

Échange de normes sur l'IA

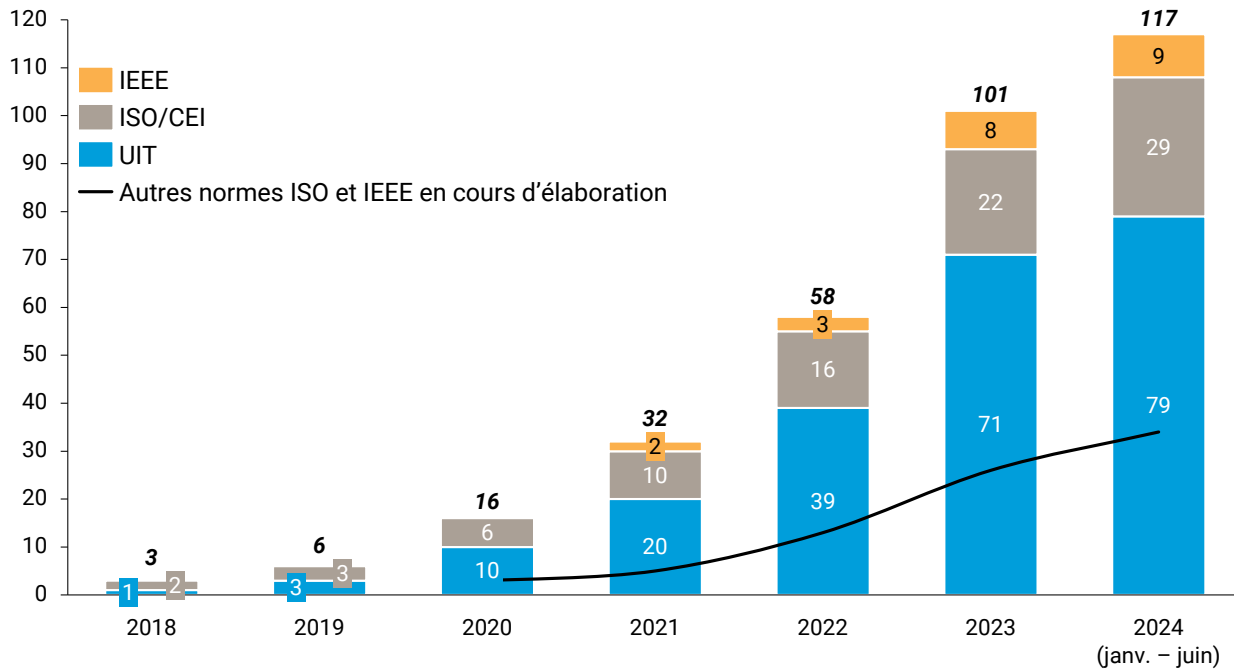
xxxiii Lorsque les systèmes d'IA ont été étudiés pour la première fois, il existait peu de normes pour aider à naviguer ou à mesurer cette nouvelle frontière. Plus récemment, on a assisté à une multiplication des normes. La figure b) illustre l'augmentation du nombre de normes adoptées par l'UIT, l'Organisation internationale de normalisation (ISO), la Commission électrotechnique internationale (CEI) et Institute of Electrical and Electronics Engineers (IEEE).

xxxiv Il n'existe pas de langage commun entre ces organismes de normalisation, et de nombreux termes couramment utilisés en rapport avec l'IA (équité, sûreté, transparence) ne font pas l'objet de définitions convenues. Il existe également des

divergences entre les normes qui ont été adoptées à des fins techniques étroites ou de validation interne, et celles qui visent à intégrer des principes éthiques plus larges. Nous disposons aujourd'hui d'un nouvel ensemble de normes qui ne font pas l'objet d'une définition commune ou qui sont dissociées des valeurs qu'elles étaient censées défendre.

xxxv En s'appuyant sur les compétences du groupe scientifique international et en intégrant des membres des diverses entités nationales et internationales qui ont contribué à l'établissement de normes, ainsi que des représentants d'entreprises technologiques et de la société civile, le système des Nations Unies pourrait servir de centre d'échange de normes en matière d'IA qui s'appliqueraient au niveau mondial.

Figure b) Nombre de normes relatives à l'IA



Sources : IEEE, ISO/CEI, UIT, Coopération mondiale de la normalisation (sur la base de la cartographie de juin 2023, élargie par l'inclusion des normes liées à l'IA).

Recommandation 3

Centre d'échange de normes en matière d'IA

Nous recommandons la création d'un centre d'échange de normes en matière d'IA, réunissant des représentants d'organismes nationaux et internationaux de normalisation, d'entreprises technologiques et de la société civile ainsi que des représentants du groupe scientifique international. Ce centre aurait les fonctions suivantes :

- a) Élaborer et tenir à jour un registre des définitions et des normes applicables à la mesure et à l'évaluation des systèmes d'IA ;
- b) Examiner et évaluer les normes et les processus d'élaboration ;
- c) Recenser les lacunes qui doivent être comblées par de nouvelles normes.

C. Avantages communs

- xxxvi** Le Programme de développement durable à l'horizon 2030 et ses 17 objectifs de développement durable peuvent clarifier l'objectif de la mise au point, du déploiement et des utilisations de l'IA et, ce faisant, axer l'arc des investissements sur les défis mondiaux en matière de développement. Sans une approche globale et inclusive de la gouvernance, l'IA risque de ne pas contribuer positivement à la réalisation des objectifs de développement durable et son déploiement pourrait, par inadvertance, renforcer ou exacerber les disparités et les préjugés.
- xxxvii** L'IA n'est pas une solution miracle aux problèmes de développement durable. Il s'agit d'une composante d'un ensemble plus large de solutions. Afin de tirer pleinement parti de la capacité de l'IA de relever les défis sociétaux, la collaboration entre les gouvernements, le milieu universitaire, l'industrie et la société civile est cruciale pour que les solutions basées sur l'IA soient inclusives et équitables.
- xxxviii** Cela dépend en grande partie de l'accès aux talents, à la puissance de calcul et aux données, selon des modalités qui favorisent l'épanouissement de la diversité culturelle et linguistique. Une infrastructure de base et les ressources nécessaires à son entretien sont également des conditions préalables.
- xxxix** S'agissant des talents, toutes les sociétés n'ont pas besoin de cadres d'informaticiens pour construire leurs propres modèles. Toutefois, que la technologie soit achetée, empruntée ou construite, une capacité sociotechnique de base est nécessaire pour comprendre les capacités et les limites de l'IA et l'utiliser au mieux compte tenu des risques propres au contexte.

- xi** La puissance de calcul est l'un des plus grands obstacles à l'entrée dans le domaine de l'IA. Sur les 100 premiers groupes de calcul haute performance au monde capables d'entraîner de grands modèles d'IA, aucun n'est hébergé dans un pays en développement⁶. Il est irréaliste de promettre un accès à la puissance de calcul que même les pays et les entreprises les plus riches peinent à acquérir. Nous cherchons plutôt à réduire la fracture de l'IA pour ceux qui ne sont pas en mesure de se procurer les outils nécessaires par d'autres moyens, notamment en soutenant les initiatives en faveur de modèles de développement de l'IA distribués et fédérés.
- xli** S'agissant des données, il est courant de parler d'utilisation abusive des données dans le contexte de l'IA (comme les atteintes à la vie privée) ou d'utilisations manquées des données (le fait de ne pas exploiter les ensembles de données existants). Toutefois un problème connexe est celui des données manquantes, ce qui inclut les grandes parties du globe qui sont dépourvues de données. Le fait de ne pas tenir compte de la diversité linguistique et culturelle du monde a été associé à des préjugés dans les systèmes d'IA, mais peut également représenter une occasion manquée pour ces populations d'accéder aux avantages offerts par l'IA.
- xlii** Il convient d'établir un ensemble de ressources communes, notamment des modèles ouverts, afin de soutenir la participation inclusive et efficace de tous les États Membres à l'écosystème de l'IA car, à ce niveau, les approches globales présentent des avantages distincts.

6 Indicateur supplétif car la plupart des groupes de calcul haute performance n'ont pas d'unités de processeur graphique et présentent une utilité limitée pour l'IA avancée.

Réseau de développement des capacités

xliii L'augmentation de la demande des secteurs public et privé en capacités, notamment humaines, dans le domaine de l'IA coïncide avec l'apparition de centres d'excellence nationaux, régionaux et publics-privés de l'IA, qui jouent un rôle dans le développement des capacités au niveau international. Un réseau mondial peut servir de plateforme de correspondance qui élargit l'éventail des partenariats possibles et améliore l'interopérabilité des approches de renforcement des capacités.

xliv Depuis les objectifs du Millénaire pour le développement jusqu'aux objectifs de développement durable, l'ONU s'emploie de longue date à renforcer les capacités des personnes et des organismes⁷. Un réseau d'organismes, affilié à l'ONU, pourrait élargir les options pour les pays qui sont à la recherche de partenariats en matière de capacités. Il pourrait également catalyser de nouveaux centres d'excellence nationaux afin de stimuler le développement d'écosystèmes locaux d'innovation en matière d'IA, sur la base d'approches interopérables conformes aux engagements normatifs de l'ONU.

xlv Le réseau favoriserait un autre paradigme de développement des technologies de l'IA : ascendant, interdisciplinaire, ouvert et collaboratif. Les initiatives nationales pourraient continuer à utiliser des outils de diagnostic tels que la méthode d'évaluation de l'état de préparation à l'IA élaborée par l'UNESCO afin de recenser les lacunes au niveau national, le réseau international aidant à les combler.

⁷ Grâce aux travaux de l'UNESCO, de l'OMPI et d'autres organisations, l'ONU contribue à préserver la grande diversité des cultures et des traditions de production de connaissances dans le monde entier. L'Université des Nations Unies (UNU) s'emploie depuis longtemps à renforcer les capacités par l'enseignement supérieur et la recherche, et l'Institut des Nations Unies pour la formation et la recherche contribue à former des fonctionnaires dans des domaines essentiels au développement durable. La méthode d'évaluation de l'état de préparation élaborée par l'UNESCO est un outil essentiel pour aider les États Membres à appliquer la recommandation de l'Organisation sur l'éthique de l'IA. Il existe également d'autres exemples, tels que l'Académie de l'OMS à Lyon (France), l'Institut virtuel de la CNUCED, le Programme de bourses d'études sur le désarmement du Bureau des affaires de désarmement et les programmes de développement des capacités menés par l'UIT et le PNUD.

Recommandation 4

Réseau de développement des capacités

Nous recommandons la création d'un réseau de développement des capacités en matière d'IA afin de relier un ensemble de centres de développement des capacités affiliés à l'ONU et collaborant entre eux, et mettre ainsi à la disposition des acteurs clés des compétences, la puissance de calcul et les données d'entraînement en matière d'IA. Les objectifs de ce réseau sont les suivants :

- a) Catalyser et harmoniser les initiatives régionales et mondiales liées aux capacités en matière d'IA au moyen de la mise en réseau ;
- b) Renforcer les capacités de gouvernance des fonctionnaires en matière d'IA afin d'encourager le développement et, en même temps, de favoriser le respect, la protection et la concrétisation des droits humains ;
- c) Mettre à disposition des formateurs, une puissance de calcul et des données d'entraînement à l'IA dans plusieurs centres à l'intention des chercheurs et des entrepreneurs sociaux qui visent à appliquer l'IA à des cas d'utilisation d'intérêt général local, notamment par les moyens suivants :
 - i) Des protocoles permettant aux équipes de recherche interdisciplinaires et aux entrepreneurs d'accéder à la puissance de calcul lorsqu'elle fait défaut aux fins de l'entraînement, de la mise au point et de l'application de leurs modèles en fonction des contextes locaux ;
 - ii) Des dispositifs expérimentaux pour mettre à l'essai d'éventuelles solutions d'IA et apprendre par la pratique ;
 - iii) Une série de formations en ligne sur l'IA destinées aux étudiants, aux jeunes chercheurs, aux entrepreneurs sociaux et aux fonctionnaires du secteur public ;
 - iv) Un programme de bourses destiné à des personnes prometteuses afin de leur permettre de passer du temps dans des établissements universitaires ou des entreprises technologiques.

Fonds mondial pour l'IA

xlvi De nombreux pays font face à des contraintes de budget et de ressources qui limitent leur capacité à utiliser l'IA de manière adaptée et efficace. Malgré les efforts de développement des capacités (recommandation 4), il est parfois encore impossible d'accéder à la formation, à la puissance de calcul, aux modèles et aux données d'entraînement sans un soutien international. D'autres efforts de financement risquent également de ne pas pouvoir être déployés sans ce soutien.

xlvii Notre intention, en proposant un fonds, n'est pas de garantir l'accès à des ressources et à des capacités de calcul avancées. La réponse n'est pas toujours une augmentation de la puissance de calcul. Nous avons également besoin de meilleurs moyens pour mettre en relation les talents, la puissance de calcul et les données. L'objectif du fonds est de combler le manque sous-jacent de capacité et de collaboration pour ceux qui ne sont pas en mesure d'accéder aux catalyseurs requis, notamment :

- a. Permettre aux pays qui en ont besoin d'accéder à des outils d'aide à l'IA, ce qui réduirait la fracture en matière d'IA ;
- b. Favoriser la collaboration pour le développement des capacités en matière d'IA afin de créer des habitudes de coopération et d'atténuer la concurrence géopolitique ;
- c. Inciter les pays ayant des approches réglementaires divergentes à élaborer des modèles communs pour régir les données, les modèles et les applications afin de relever les défis sociétaux liés aux objectifs de développement durable et aux avancées scientifiques.

xlvi L'accent étant mis sur l'intérêt général, le fonds vient compléter la proposition de réseau de développement des capacités en matière d'IA, vers lequel il canaliserait également des ressources. Il fournirait une capacité indépendante de suivi de l'impact, et pourrait trouver et partager des contributions en nature, notamment de la part d'entités du secteur privé, afin de mettre à disposition des programmes de formation liés à l'IA, du temps, une puissance de calcul, des modèles et des ensembles de données à un coût inférieur à celui du marché. Nous veillons ainsi à ce que de vastes régions du monde ne soient pas laissées pour compte et aient au contraire les moyens de mettre l'IA au service des objectifs de développement durable dans différents contextes.

xlvi Il est dans l'intérêt de toutes et de tous qu'une coopération soit établie dans le monde tant numérique que physique. Des analogies peuvent être faites avec les efforts de lutte contre les changements climatiques, où les coûts de la transition, de l'atténuation ou de l'adaptation ne sont pas répartis de manière égale, et où l'assistance internationale est essentielle pour que les pays aux ressources limitées participent à l'effort mondial en vue de relever un défi planétaire.

Recommandation 5

Fonds mondial pour l'IA

Nous recommandons la création d'un fonds mondial pour l'IA afin de réduire la fracture en la matière. Géré par une structure de gouvernance indépendante, le fonds recevrait des contributions financières et des ressources en nature de sources publiques et privées, qui seraient allouées, notamment par l'intermédiaire du réseau de développement des capacités, pour faciliter l'accès aux outils de l'IA en vue de favoriser la réalisation des objectifs de développement durable au niveau local, notamment :

- a) Des ressources de calcul mises en commun pour l'entraînement et la mise au point de modèles par les développeurs d'IA des pays qui ne disposent pas des capacités locales adéquates ou des moyens de se les procurer ;
- b) Des dispositifs expérimentaux et des outils d'analyse comparative et d'essai visant à généraliser les meilleures pratiques en matière de développement de modèles sûrs et fiables et de gouvernance des données ;
- c) Des solutions de gouvernance, de sûreté et d'interopérabilité applicables à l'échelle mondiale ;
- d) Des ensembles de données et des travaux de recherche sur la manière dont les données et les modèles pourraient être combinés afin de mettre en œuvre des projets liés aux objectifs de développement durable ;
- e) Un référentiel de modèles d'IA et d'ensembles de données conservés aux fins de la réalisation des objectifs de développement durable.

Cadre mondial de données sur l'IA

I L'accès aux données d'entraînement de l'IA, par le marché ou d'autres dispositifs, est un facteur essentiel pour faire prospérer des écosystèmes locaux d'IA – en particulier dans les pays, les populations, les régions et les groupes démographiques où les données sont « manquantes » (voir la section sur les « avantages communs » ci-dessus).

ii Seule une action collective mondiale peut encourager l'interopérabilité, la gestion, la préservation de la vie privée, l'autonomisation et le renforcement des droits de manière à promouvoir un nivellement par le haut entre les pays pour la protection des droits humains et d'autres engagements convenus, la disponibilité des données et une compensation équitable pour les personnes concernées dans la gouvernance de la collecte, de la création, de l'utilisation et de la monétisation des données d'entraînement de l'IA. Notre proposition de cadre mondial pour les données d'IA s'inscrit dans cet objectif.

lii Le cadre ne créerait pas de nouveaux droits liés aux données. Il permettrait au contraire de traiter les questions de disponibilité, d'interopérabilité et d'utilisation des données d'entraînement de l'IA. Il contribuerait à élaborer une vision commune de la manière d'harmoniser les différents cadres nationaux et régionaux de protection des données. Il pourrait également faire prospérer les écosystèmes locaux d'IA qui soutiennent la diversité culturelle et linguistique et limiter la poursuite de la concentration économique.

liii Les mesures pourraient être complétées par la promotion de communs de données et de dispositions relatives à l'hébergement de référentiels de données dans des domaines en rapport avec les objectifs de développement durable, sur la base de modèles d'accords visant à détenir et à mettre en commun des données de manière juste, sûre et équitable. L'élaboration de ces modèles, ainsi que le stockage et l'analyse des données communes ou fiduciaires, pourraient être soutenus par le réseau de développement des capacités et le fonds mondial pour l'IA proposés (recommandations 4 et 5).

liv L'ONU est particulièrement bien placée pour soutenir l'établissement de principes mondiaux et de dispositions pratiques régissant la gouvernance et l'utilisation des données d'entraînement de l'IA, conformément aux engagements internationaux pris en matière de droits humains, de propriété intellectuelle et de développement durable, en s'appuyant sur les travaux menés de longue date par la communauté des données et en les intégrant aux nouveautés récentes en matière d'éthique et de gouvernance de l'IA. Ce rôle est semblable à celui de la Commission des Nations Unies pour le droit commercial international (CNUDCI), qui fait progresser le commerce international en élaborant des cadres transfrontaliers juridiques et non juridiques.

lv De même, la Commission de la science et de la technique au service du développement et la Commission de statistique ont inscrit à leur ordre du jour l'examen des données pour le développement et les objectifs de développement durable. L'OMPI traite également d'importantes questions relatives au contenu, aux droits d'auteur et à la protection des connaissances autochtones et de l'expression culturelle.

Recommandation 6

Cadre mondial de données sur l'IA

Nous recommandons la création d'un cadre mondial de données sur l'IA, qui s'inscrirait dans un processus lancé par un organisme compétent tel que la CNUDCI, sur la base des travaux menés par d'autres organisations internationales, dans les buts suivants :

- a) Élaborer des définitions et des principes relatifs aux données pour la gouvernance mondiale des données d'entraînement de l'IA, notamment à partir des meilleures pratiques existantes, et promouvoir la diversité culturelle et linguistique ;
- b) Établir des normes communes concernant la provenance et l'utilisation des données d'entraînement de l'IA pour une responsabilité transparente et fondée sur les droits dans tous les pays ;
- c) Instaurer des dispositifs de gestion et d'échange de données qui façonnent le marché afin de faire prospérer des écosystèmes d'IA locaux à l'échelle mondiale, tels que :
 - i) Des référentiels de données ;
 - ii) Des places de marché mondiales bien gérées pour l'échange de données anonymes destinées à l'entraînement de modèles d'IA ;
 - iii) Des accords types visant à faciliter l'accès international aux données et l'interopérabilité mondiale, éventuellement en tant que protocoles technicojuridiques du cadre.

D. Stratégie cohérente

- lvi** Les propositions susmentionnées visent à combler les lacunes en matière de représentation, de coordination et de mise en œuvre recensées dans le nouveau régime international de gouvernance de l'IA. Ces lacunes peuvent être comblées par des partenariats et une collaboration avec les organismes et dispositifs existants afin de promouvoir une vision commune, un socle commun et des avantages communs.
- lvii** Toutefois, sans un point focal à l'ONU chargé de soutenir et favoriser la coordination de ces travaux, il manquera au niveau mondial une approche cohérente, agile, interconnectée et inclusive nécessaire à une gouvernance efficace et équitable de l'IA en tant que technologie transfrontalière, à évolution rapide et à usage général.
- lviii** La mosaïque de normes et d'organismes décrite à la section « Lacunes dans la gouvernance mondiale de l'IA » ci-dessus témoigne de la nécessité d'établir une gouvernance de l'IA au niveau mondial. Il convient de combler cette lacune au moyen d'une stratégie cohérente.

Bureau de l'AI au Secrétariat de l'ONU

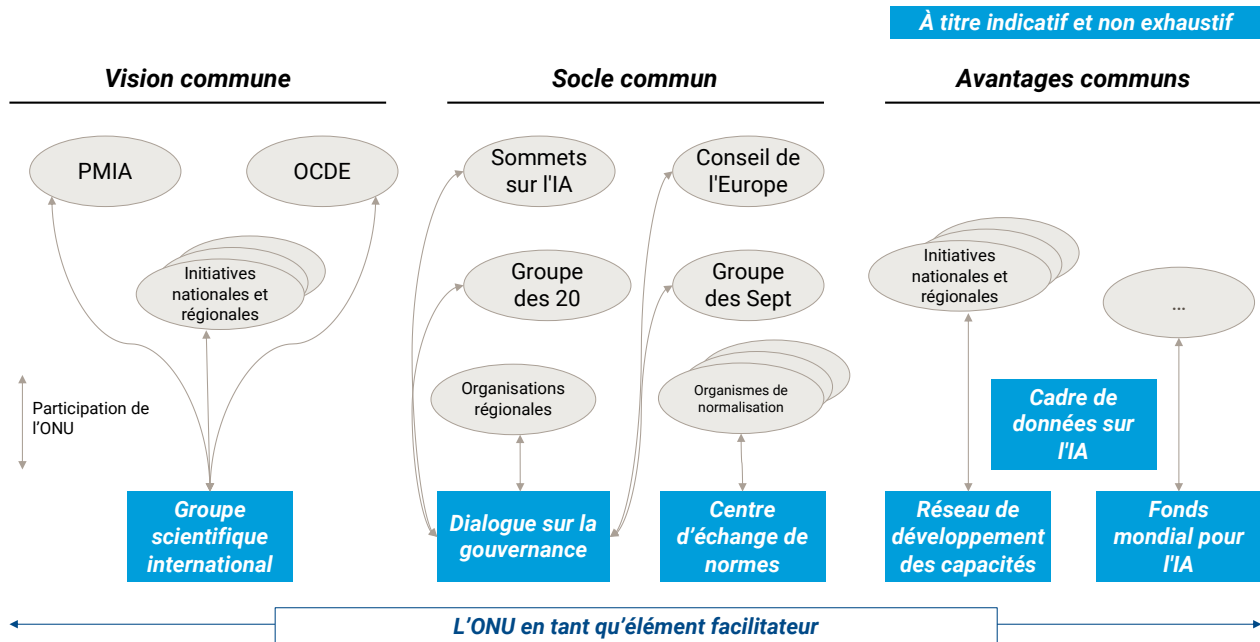
- lix** Nous proposons donc un dispositif léger qui servirait de « ciment » pour soutenir et catalyser les propositions énoncées dans le présent rapport, notamment au moyen de partenariats, tout en permettant au système des Nations Unies de parler d'une seule voix dans l'écosystème en constante évolution de la gouvernance de l'IA.
- lx** Une petite capacité souple, sous la forme d'un bureau de l'IA au Secrétariat de l'ONU, relèverait du Secrétaire général, ce qui lui permettrait de bénéficier de connexions à l'échelle du système, sans être liée à une seule partie. Cela est important en raison de l'avenir incertain de l'IA et de la forte probabilité qu'elle s'infilte dans tous les aspects de l'activité humaine.

lxi L'organe devrait être agile, promouvoir l'inclusion et s'associer rapidement afin d'accélérer la coordination et la mise en œuvre, en s'appuyant en priorité sur les ressources et les fonctions existantes au sein du système des Nations Unies. L'accent devrait être mis sur les applications civiles de l'IA.

lxii Le bureau pourrait être composé en partie de membres du personnel de l'ONU détachés par des organismes spécialisés et d'autres parties du système des Nations Unies, telles que l'UIT, l'UNESCO, le Haut-Commissariat aux droits de l'homme, la CNUCED, l'UNU et le PNUD. Il devrait collaborer avec de nombreuses parties prenantes, notamment les entreprises, la société civile et le milieu universitaire, et nouer des partenariats avec des organisations de premier plan en dehors de l'ONU [voir fig. c)]. L'Organisation serait ainsi en mesure d'établir des liens pour favoriser une vision commune, un socle commun et des avantages communs dans l'écosystème international de la gouvernance de l'IA.

lxiii La recommandation 7 se fonde sur une évaluation lucide des domaines dans lesquels l'ONU peut apporter une valeur ajoutée, y compris un rôle moteur, un rôle de coordination ou un rôle de deuxième plan. Elle se fonde également sur les avantages des dispositions institutionnelles existantes, notamment les financements prénégociés et les processus administratifs qui sont bien établis et compris.

Figure c) Rôle que l'ONU pourrait assurer dans l'écosystème international de gouvernance de l'IA



Abréviations : PMIA, Partenariat mondial sur l'intelligence artificielle ; OCDE, Organisation de coopération et de développement économiques ; organismes de normalisation.

Recommandation 7

Bureau de l'IA au Secrétariat

Nous recommandons la création d'un bureau de l'IA au Secrétariat, qui relèverait du Secrétaire général. Sa structure devrait être légère et agile et s'appuyer, dans la mesure du possible, sur les entités de l'ONU existantes. Agissant comme le « ciment » qui soutient et catalyse les propositions contenues dans le présent rapport, en partenariat et en interface avec d'autres processus et organismes, le bureau aurait le mandat suivant :

- a) Soutenir le groupe scientifique international, la concertation, l'échange de normes, le réseau de développement des capacités et, dans la mesure où cela est nécessaire, le fonds mondial et le cadre mondial de données sur l'IA ;
- b) Sensibiliser diverses parties prenantes, notamment les entreprises technologiques, la société civile et le milieu universitaire, aux questions qui se posent en matière d'IA ;
- c) Conseiller le Secrétaire général sur les questions liées à l'IA, en coordination avec d'autres parties concernées du système des Nations Unies, afin de fournir une réponse à l'échelle du système.

E. Réflexions sur les modèles institutionnels

lxiv Les débats sur l'IA sont souvent polarisés. Lors des consultations que nous avons menées à travers le monde, nous nous sommes entretenus avec ceux qui entrevoient un avenir de biens illimités fournis par des systèmes d'IA toujours moins chers et toujours plus utiles. Nous nous sommes également entretenus avec ceux qui craignent un avenir plus sombre, la division et le chômage, voire l'extinction⁸.

lxv Nous ignorons si l'avenir est plus probable dans sa version utopique ou dystopique. De même, nous sommes conscients que la technologie peut évoluer dans une direction qui supprime cette dualité. Le présent rapport se concentre sur les possibilités et les risques à court terme, sur la base de données scientifiques et factuelles.

lxvi Les sept recommandations susmentionnées constituent notre meilleur espoir de tirer profit des avantages de l'IA, tout en réduisant au minimum et en atténuant les risques, au fur et à mesure que l'IA continue d'évoluer. Nous sommes également conscients des défis pratiques que pose le renforcement des institutions internationales à plus grande échelle. C'est pourquoi nous proposons une approche institutionnelle en réseau, avec un soutien léger et agile. Si ou lorsque les risques s'aggravent et les enjeux liés aux possibilités offertes augmentent, ces calculs peuvent changer.

lxvii Les guerres mondiales ont donné naissance au système international moderne. La mise au point d'armes chimiques, biologiques et nucléaires toujours plus puissantes a conduit à l'établissement de régimes qui limitent leur diffusion et favorisent l'utilisation pacifique des technologies sous-jacentes. L'évolution de la vision de notre humanité commune a conduit au système moderne des droits humains et à notre engagement actuel en faveur des objectifs de développement durable pour tous. Les changements climatiques ne sont plus une préoccupation de niche mais un défi mondial.

lxviii De même, l'évolution de l'IA peut exiger plus de ressources et de limites que ce qui est proposé dans les recommandations susmentionnées, ainsi que des travaux plus complexes d'élaboration de normes, de mise en œuvre, de suivi, de vérification et de validation, d'application, de responsabilisation, de recours en cas de préjudice et d'interventions d'urgence. Il est donc prudent de réfléchir à ces modèles institutionnels. La dernière partie de notre rapport vise à contribuer à ces travaux.

4. Appel à l'action

lxix Nous restons optimistes quant à l'avenir de l'IA et aux effets positifs qu'elle peut engendrer. Cet optimisme dépend toutefois de la réalité des risques et de l'inadéquation des structures et des incitations actuellement en place. La technologie est trop importante et les enjeux sont trop élevés pour que l'on s'en remette uniquement aux forces du marché et à une mosaïque fragmentée de mesures nationales et multilatérales.

lxx L'ONU peut être le vecteur d'un nouveau contrat social pour l'IA qui garantisse l'adhésion mondiale à un régime de gouvernance qui nous protège et nous donne à tous les moyens d'agir. Un tel contrat social garantira que les chances sont équitablement réparties et que les risques ne sont pas reportés sur les plus vulnérables – ou sur les générations futures, comme nous l'avons vu, malheureusement, avec les changements climatiques.

lxxi En tant que groupe et en tant que personnes représentant de nombreux domaines de compétences, organisations et régions du monde, nous nous réjouissons de poursuivre cette réflexion cruciale. En collaboration avec les nombreuses personnes rencontrées au cours de ce voyage, et la communauté mondiale qu'elles représentent, nous espérons que le présent rapport favorisera les travaux que nous menons conjointement en vue de gouverner l'IA au bénéfice de l'humanité.

8 Voir <https://safe.ai/work/statement-on-ai-risk>.

Figure d) L'organe consultatif de haut niveau sur l'IA lors de sa réunion à Singapour, le 29 mai 2024



1. Introduction

- 1 L'Organe consultatif de haut niveau sur l'IA a été créé en vue d'analyser la gouvernance internationale de l'IA et de formuler des recommandations. Nos membres se distinguent par leur diversité, tant sur le plan géographique que du genre, de la discipline et de l'âge. Nous nous appuyons sur les compétences des gouvernements, de la société civile, du secteur privé et du milieu universitaire. Des débats intenses et divers ont permis de dégager un large consensus (tel qu'indiqué dans notre rapport d'étape¹) sur le manque de gouvernance mondiale en matière d'IA. Dans ce rapport, nous avons formulé des principes directeurs pour le rôle et les fonctions qui pourraient être requises au niveau international.
- 2 Au fil des mois qui ont suivi, nous avons exploité les nombreux retours d'information reçus et les vastes consultations organisées. Cela comprend 18 études approfondies menées dans divers domaines par plus de 500 experts, plus de 250 contributions écrites émanant de plus de 150 organisations et 100 personnes de toutes les régions, une enquête sur les risques liés à l'IA à laquelle quelque 350 experts de toutes les régions ont répondu, une analyse des opportunités à laquelle environ 120 experts de toutes les régions ont participé, ainsi que des consultations et des séances d'informations tenues régulièrement avec des États Membres, des entités des Nations Unies et d'autres groupes de parties prenantes, à l'occasion d'une quarantaine d'activités organisées dans toutes les régions². Les membres de l'Organe consultatif ont également participé à de nombreux forums dans le monde entier, tenu plus d'une centaine de débats en ligne et organisé trois réunions plénières en personne à New York, Genève et Singapour.
- 3 Le présent rapport final est donc le fruit du travail de nombreux auteurs. S'il ne peut pas rendre compte de toute la richesse et la diversité des points de vue exprimés, il témoigne de notre volonté commune de veiller à ce que l'IA soit mise au point, déployée et utilisée de manière à bénéficier à l'ensemble de l'humanité, et gouvernée de manière efficace et inclusive au niveau international.
- 4 Les auteurs du présent rapport réaffirment les conclusions du rapport d'étape de l'Organe consultatif sur les possibilités et les catalyseurs, les risques et les défis. Ils réaffirment également la nécessité d'une gouvernance mondiale de l'IA et formulent sept recommandations.
- 5 Ces recommandations portent notamment sur la création d'un groupe scientifique chargé de promouvoir une vision commune des capacités, possibilités, risques et incertitudes liés à l'IA. Sur la base de cette vision commune, il nous faut établir des dispositifs pour trouver un socle commun qui sous-tend la gouvernance de l'IA au niveau international. La réalisation de cet objectif repose sur un dialogue régulier et l'élaboration de normes acceptables et applicables par tous.
- 6 Le présent rapport contient également des recommandations sur les avantages communs pour faire en sorte que les retombées de l'IA soient équitablement réparties, ce qui peut dépendre de l'accès à des modèles ou à des capacités telles que les talents, la puissance de calcul et les données. À cette fin, il convient notamment de mettre en place un réseau de développement des capacités, un fonds mondial pour l'IA et un cadre mondial de données sur l'IA.

1 Voir <https://un.org/ai-advisory-body>.

2 Voir l'annexe C pour un aperçu des consultations.

- 7** Afin de faciliter ces travaux, collaborer avec d'autres initiatives et organismes pour résoudre les préoccupations et saisir les possibilités liées à l'IA et garantir que le système des Nations Unies parle d'une seule voix sur cette question, nous proposons la création d'un bureau de l'IA au Secrétariat de l'ONU.
- 8** Bien que nous envisagions de recommander la création d'un organisme international pour l'IA, nous ne recommandons pas cette mesure pour le moment, mais nous sommes conscients qu'il faut assurer la gouvernance afin de suivre le rythme de l'évolution technologique.
- 9** Au-delà des débats et processus multilatéraux immédiats impliquant les gouvernements, notre rapport est également destiné à la société civile et au secteur privé, aux chercheurs et aux personnes concernées dans le monde entier. Nous sommes tout à fait conscients que la réalisation des objectifs ambitieux que nous avons établis repose obligatoirement sur une participation mondiale multisectorielle.
- 10** Dans l'ensemble, nous pensons que l'avenir de cette technologie reste ouvert, comme indiqué dans notre étude approfondie de l'orientation de la technologie et le débat entre les approches ouvertes et fermées de son développement (voir encadré 9). À l'avenir, il est possible que des modèles plus grands et plus puissants soient mis au point par un nombre de plus en plus restreint d'entreprises, ou que le paysage mondial de l'innovation soit plus diversifié et dominé par des modèles d'IA interopérables de petite et moyenne taille fournissant une multitude d'applications sociétales et économiques. Avec nos recommandations, nous cherchons à rendre cette deuxième hypothèse plus probable, tout en tenant compte des risques qu'elle comporte.
- 11** Depuis sa création, l'ONU s'emploie à promouvoir le progrès économique et social de tous les peuples³. Les objectifs du Millénaire pour le développement comportaient des buts ambitieux pour que tous les habitants de la planète aient accès à des débouchés économiques ; par la suite, les objectifs de développement durable ont cherché à concilier le besoin de développement et les contraintes environnementales de notre planète. Le développement, le déploiement et l'utilisation accrues des outils et des systèmes d'IA constituent le prochain grand défi à relever afin que nous puissions accueillir ensemble notre avenir numérique, au lieu de creuser la fracture numérique.
- 12** La gouvernance inclusive de l'IA est sans doute l'un des défis les plus difficiles à relever par l'ONU en matière de gouvernance. Le rôle dominant du secteur privé dans l'IA n'est pas compatible avec le système westphalien de politique internationale. Les États sont tentés par la puissance et la prospérité que l'IA peut apporter à une époque de concurrence géopolitique intense. De nombreuses sociétés restent en marge du développement, du déploiement et de l'utilisation de l'IA, tandis que d'autres sont saisies d'un enthousiasme mêlé d'inquiétude quant à l'impact transversal de l'IA.
- 13** Malgré les défis, il n'y a pas de retour en arrière possible. Les enjeux sont tout simplement trop élevés pour l'ONU, ses États Membres et la communauté au sens large dont l'Organisation représente les aspirations. Nous espérons que le présent rapport fournira des pistes pour faciliter les travaux menés conjointement en vue de gouverner l'IA au bénéfice de l'humanité.

3 Il s'agit notamment du commerce, de l'investissement étranger direct et du transfert de technologie, qui favorisent le développement à long terme.

A. Des débouchés et des catalyseurs

14 L'IA est en train de transformer notre monde. Les avantages que cet ensemble de technologies⁴ peut procurer sont colossaux, qu'il s'agisse d'ouvrir de nouveaux domaines de recherche scientifique (voir encadré 1), d'optimiser l'utilisation des réseaux énergétiques ou d'améliorer la santé publique et l'agriculture⁵. Si elles se concrétisent, les

possibilités offertes par l'utilisation des outils d'IA pour les personnes, les secteurs de l'économie, la recherche scientifique et d'autres domaines d'intérêt général pourraient jouer un rôle important dans la stimulation de nos économies (voir encadré 2) et la transformation positive de nos sociétés. L'IA d'intérêt général, consistant notamment à anticiper et combattre les pandémies, les inondations, les feux de forêt et l'insécurité alimentaire, pourrait même contribuer à faire progresser la réalisation des objectifs de développement durable.

Encadré 1 : Contribution que l'IA pourrait apporter au progrès scientifique

L'IA pourrait bien constituer le prochain grand bond en avant du progrès scientifique, sur la base de l'héritage transformateur d'Internet. Le Web a facilité l'échange de vastes quantités de données expérimentales, d'articles scientifiques et de documentation entre scientifiques. L'IA s'appuie sur cette base pour analyser de vastes ensembles de données, découvrir des modèles cachés, élaborer de nouvelles hypothèses et associations, et accélérer le rythme des découvertes, notamment au moyen d'expériences à grande échelle avec la robotique automatisée.

Les effets exercés par l'IA sur la science touchent des disciplines majeures. De la biologie à la physique, et des sciences de l'environnement aux sciences sociales, l'IA est intégrée dans les flux de travail de la recherche et accélère la production de connaissances scientifiques. Certaines affirmations sont peut-être exagérées, tandis que d'autres ont été démontrées et leurs effets à long terme semblent prometteurs^a.

En biologie, par exemple, l'IA a permis de relever un défi vieux de 50 ans, à savoir prédire le repliement et la structure des protéines. Elle a notamment permis de prédire la structure de plus de 200 millions de protéines, la base de données en libre accès qui en résulte ayant été utilisée, au moment de la rédaction du présent document, par plus de 2 millions de scientifiques dans plus de 190 pays, dont un grand nombre travaillent sur des maladies négligées. Cette approche a depuis été étendue aux autres biomolécules de la vie, à l'ADN, à l'ARN, aux ligands et à leurs interactions.

a Voir John Jumper et al., « Highly accurate protein structure prediction with AlphaFold », *Nature*, vol. 596 (juill. 2021), p. 583 à 589 ; voir également Josh Abramson et al., « Accurate structure prediction of biomolecular interactions with AlphaFold 3 », *Nature*, vol. 630, p. 493 à 500 (mai 2024).

4 Selon la définition de l'OCDE, un système d'intelligence artificielle est un système automatisé qui, pour des objectifs explicites ou implicites, déduit, à partir d'entrées reçues, comment générer des résultats en sortie tels que des prévisions, des contenus, des recommandations ou des décisions qui peuvent influencer sur des environnements physiques ou virtuels. Différents systèmes d'IA présentent des degrés variables d'autonomie et d'adaptabilité après déploiement (voir <https://oecd.ai/en/wonk/ai-system-definition-update>).

5 Toutefois, nous estimons que des experts en la matière doivent procéder à une évaluation rigoureuse des avantages que l'on attribue à l'IA. La quête des bienfaits de l'IA devrait se fonder sur des données scientifiques et sur une évaluation approfondie des compromis et des solutions de rechange. Outre la recherche scientifique, les sciences sociales sont également en train de se transformer.

Encadré 1 : Contribution que l'IA pourrait apporter au progrès scientifique (suite)

Pour les maladies d'Alzheimer, de Parkinson et la sclérose latérale amyotrophique, des experts utilisent l'IA en vue de détecter les biomarqueurs de la maladie et prédire les réponses au traitement, améliorant ainsi nettement la précision et la rapidité du diagnostic et de la mise au point du traitement^b. D'une manière générale, l'IA contribue à faire progresser la médecine de précision (par exemple pour les maladies neurodégénératives) car elle permet d'adapter les traitements en fonction des profils génétiques et cliniques. Elle contribue également à accélérer la découverte et la mise au point de nouveaux composés chimiques^c.

En radioastronomie, la vitesse et la quantité de données recueillies au moyen d'instruments modernes tels que le Square Kilometer Array peuvent dépasser celles obtenues les méthodes traditionnelles. L'IA peut changer la donne, notamment en aidant à sélectionner la partie des données sur laquelle se concentrer afin d'obtenir de nouvelles informations. Grâce au « regroupement non supervisé », l'IA peut repérer des tendances dans les données sans qu'on lui dise ce qu'elle doit rechercher en particulier^d. L'application de l'IA à la recherche en sciences sociales pourrait également permettre de mieux comprendre les dynamiques humaines complexes, en améliorant notre compréhension des tendances sociétales et des développements économiques.

À terme, en permettant des niveaux d'interdisciplinarité sans précédent, l'IA pourrait être conçue et déployée de manière à donner naissance à de nouveaux domaines scientifiques, tout comme la bio-informatique et la neuro-informatique sont nées de l'intégration de techniques informatiques avec la recherche biologique et neurologique. La capacité de l'IA à intégrer et à analyser divers ensembles de données dans des domaines tels que les changements climatiques, la sécurité alimentaire et la santé publique pourrait ouvrir des voies de recherche reliant ces domaines habituellement distincts, à condition que cela soit fait de manière responsable.

L'IA peut également renforcer la contribution de la recherche scientifique aux politiques publiques car elle permet de valider des hypothèses complexes, par exemple de combiner des modèles climatiques avec des données agricoles afin de prévoir les risques liés à la sécurité alimentaire et de relier ces connaissances aux résultats en matière de santé publique. Une autre projection est la stimulation de la science citoyenne et l'exploitation des connaissances et des données locales en vue de relever des défis mondiaux.

b Isaias Ghebrehwet et al., « Revolutionizing personalized medicine with generative AI : a systematic review », *Artificial Intelligence Review*, vol. 57, n° 127 (avril 2024).

c Amil Merchant et al., « Scaling deep learning for materials discovery », *Nature*, vol. 624, p. 80 à 85 (nov. 2023).

d Zack Savitsky, « Astronomers are enlisting AI to prepare for a data downpour », *MIT Technology Review*, 20 mai 2024.

Encadré 2 : Débouchés économiques de l'IA

Depuis la révolution industrielle, une poignée d'innovations ont considérablement accéléré le progrès économique. Les anciennes « technologies à usage général » ont remodelé de nombreux secteurs et industries. Le dernier grand changement est intervenu avec les ordinateurs et l'ère numérique. Ces technologies ont transformé les économies et augmenté la productivité dans le monde entier, mais il a fallu des décennies pour que leurs effets se fassent pleinement sentir.

L'IA générative rompt avec la tendance à l'adoption lente. Les experts estiment que ses effets transformateurs seront visibles au cours de la décennie en cours. Cette intégration rapide implique que les récents progrès en matière d'IA pourraient rapidement remodeler les industries, modifier les processus de travail et augmenter la productivité. L'adoption rapide de l'IA pourrait donc transformer nos économies et nos sociétés comme jamais auparavant.

Les avantages économiques procurés de l'IA peuvent être colossaux. Bien qu'il soit difficile de prévoir toutes les ramifications de l'IA sur nos économies complexes, les projections indiquent que l'IA pourrait augmenter nettement le produit intérieur brut mondial, avec des effets majeurs dans presque tous les secteurs. Pour les entreprises, en particulier les microentreprises et les petites et moyennes entreprises, l'IA peut fournir un accès à des outils d'analyse et d'automatisation avancés qui n'étaient auparavant accessibles qu'aux grandes entreprises. La large applicabilité de l'IA laisse penser qu'elle pourrait être une technologie à usage général. À ce titre, l'IA pourrait favoriser la productivité des particuliers, des petites et grandes entreprises et d'autres organisations dans des secteurs aussi divers que la vente au détail, la fabrication et les opérations, les soins de santé et le secteur public, dans les économies développées et en développement^a. Cela nécessitera une large adoption au sein des secteurs et entre eux ; l'application à des fins d'amélioration de la productivité ; et une IA qui rend les travailleurs plus productifs et introduit de nouvelles activités économiques à grande échelle. Cela nécessitera également des investissements et un renforcement du capital, des co-innovations, des changements de processus et d'organisation, une préparation de la main-d'œuvre et des politiques de soutien.

Néanmoins, si l'IA peut améliorer la productivité, stimuler le commerce international et augmenter les revenus, elle devrait également avoir des effets sur le travail. Des travaux de recherche montrent que l'IA pourrait, dans certains cas, aider les travailleurs et, dans d'autres, entraîner des suppressions d'emplois^b. Ceux menés notamment par l'OIT, montrent que dans un avenir proche, l'IA sera davantage susceptible d'assister les travailleurs que de les remplacer^c.

Des travaux de recherche ont également montré que, le cas échéant, la suppression d'emplois devrait se produire différemment dans des économies à différents stades de développement^d. Si les économies avancées sont plus exposées, elles sont également davantage en mesure d'exploiter l'IA et de constituer un complément à la main-d'œuvre. Les pays à faible revenu et à revenu intermédiaire peuvent être moins à même de tirer parti de cette technologie. En outre, l'intégration de l'IA dans la main-d'œuvre peut affecter de manière disproportionnée certaines catégories démographiques, les femmes étant plus susceptibles d'être touchées par la suppression d'emplois dans certains secteurs.

a James Manyika et Michael Spence, « The coming AI economic revolution : can artificial intelligence reverse the productivity slowdown ? », *Affaires étrangères*, 24 octobre 2023.

b Erik Brynjolfsson et al., « Generative AI at work », National Bureau of Economic Research, document de travail 31161 (2023) ; voir également Shakked Noy et Whitney Zhang, « Experimental evidence on the productivity effects of generative artificial intelligence », *Science*, vol. 381, n° 6654, p. 187-192 (juill. 2023).

c Pawel Gmyrek et al., *Generative AI and Jobs : A Global Analysis of Potential Effects on Job Quantity and Quality* (Genève : OIT, 2023).

d Mauro Cazzaniga et al., « Gen-AI : artificial intelligence and the future of work », staff discussion note SDN2024/001 [Washington : Fonds monétaire international (FMI), 2024].

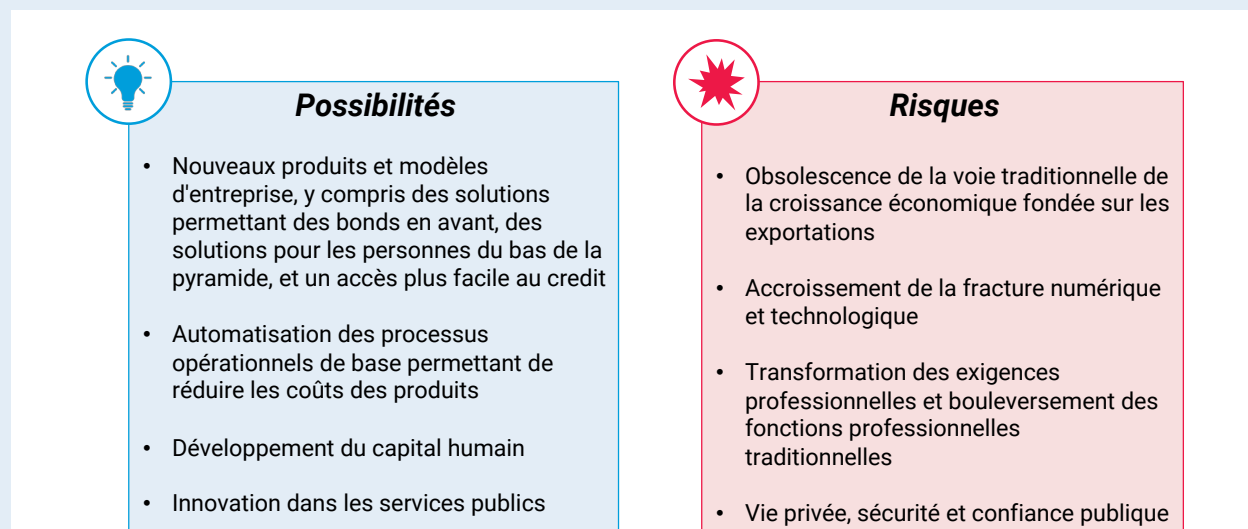
Encadré 2 : Débouchés économiques de l'IA (suite)

Des efforts ciblés et coordonnés sont indispensables pour réduire la fracture numérique, sans quoi la capacité de l'IA à soutenir le développement durable et la réduction de la pauvreté resterait inexploitée et, partant, de larges segments de la population mondiale seront désavantagés dans un environnement technologique en rapide mutation et les inégalités se creuseront.

Afin d'intégrer l'IA dans l'économie mondiale, nous avons besoin d'une gouvernance efficace qui gère les risques et garantit des résultats équitables. Il s'agit notamment de créer des dispositifs expérimentaux réglementaires pour mettre à l'essai les systèmes d'IA, promouvoir la coopération internationale en matière de normes et mettre en place des mécanismes permettant d'évaluer en permanence l'incidence de l'IA sur les marchés du travail et la société. Outre des stratégies nationales solides en matière d'IA et un soutien international, cela nécessite les mesures suivantes :

- Développement des compétences : mettre en œuvre des programmes d'éducation et de formation pour développer les compétences en matière d'IA sur le lieu de travail, allant de la culture numérique de base aux compétences techniques avancées, afin de préparer les travailleurs à un avenir renforcé par l'IA.
- Infrastructure numérique : mobiliser des investissements massifs dans l'infrastructure numérique, en particulier dans les pays en développement, afin de combler la fracture de l'IA et de faciliter l'adoption généralisée de l'IA.
- Intégration sur le lieu de travail : tirer parti du dialogue social et des partenariats public-privé afin de gérer l'intégration de l'IA sur le lieu de travail, garantir la participation des travailleurs au processus et protéger les droits du travail.
- Considérations relatives à la chaîne de valeur : garantir des conditions de travail décentes tout au long de la chaîne de valeur de l'IA, notamment dans des domaines souvent négligés tels que l'annotation de données et la modération de contenu, aux fins d'un développement équitable de l'IA.

Figure 1 : Quelques débouchés et risques liés à l'IA dans les nouveaux marchés



Fuente: Corporación Financiera Internacional.

B. Les catalyseurs de l'IA au service de l'humanité

15 Les avantages que pourraient procurer la mise au point et l'utilisation de l'IA ne seront pas nécessairement concrétisées ou exploitées de manière équitable. En mai 2024, une analyse sur le financement des projets d'IA visant à faire progresser la réalisation des objectifs de développement durable a révélé que seulement 10 % des subventions avaient été allouées à des organisations basées dans des pays à faible revenu ou à revenu intermédiaire ; s'agissant de capitaux privés, le chiffre est de 25 % (dont plus de 90 % en Chine)⁶.

C. La gouvernance : un catalyseur essentiel

16 Il convient de mettre en place des catalyseurs à l'échelle mondiale afin que les avantages procurés par l'IA soient pleinement exploités et ne profitent pas seulement à quelques personnes dans un petit nombre de pays. Pour que l'IA soit mise au service du bien commun et que ses retombées soient réparties équitablement, il faudra que les gouvernements et les organisations intergouvernementales encouragent la participation du secteur privé, du milieu universitaire et de la société civile. Tout cadre de gouvernance devrait orienter des mesures à l'échelle mondiale visant à promouvoir la réalisation d'objectifs plus vastes et inclusifs et contribuer à détecter et à prendre en compte les compromis.

D. Risques et difficultés

17 La mise au point, le déploiement et l'utilisation de l'IA comportent des risques, qui peuvent toucher de nombreux domaines à la fois. Nous conceptualisons les risques liés à l'IA en rapport avec les vulnérabilités, ce qui permet de définir les programmes stratégiques en fonction des vulnérabilités.

18 La vitesse, l'opacité et l'autonomie de l'IA posent des défis aux systèmes réglementaires traditionnels. Par l'accélération de sa mise au point et de son déploiement techniques, l'IA soulève également des enjeux pour la gouvernance internationale, son caractère polyvalent ayant des implications transfrontalières dans plusieurs domaines à la fois.

E. Risques liés à l'IA

19 On constate de plus en plus de problèmes tels que la partialité des systèmes d'IA et la surveillance abusive basée sur l'IA. D'autres risques sont associés à l'utilisation de l'IA avancée, tels que les confabulations de grands modèles de langage, la consommation élevée de ressources et les risques pour la paix et la sécurité. La désinformation générée par l'IA menace les institutions démocratiques.

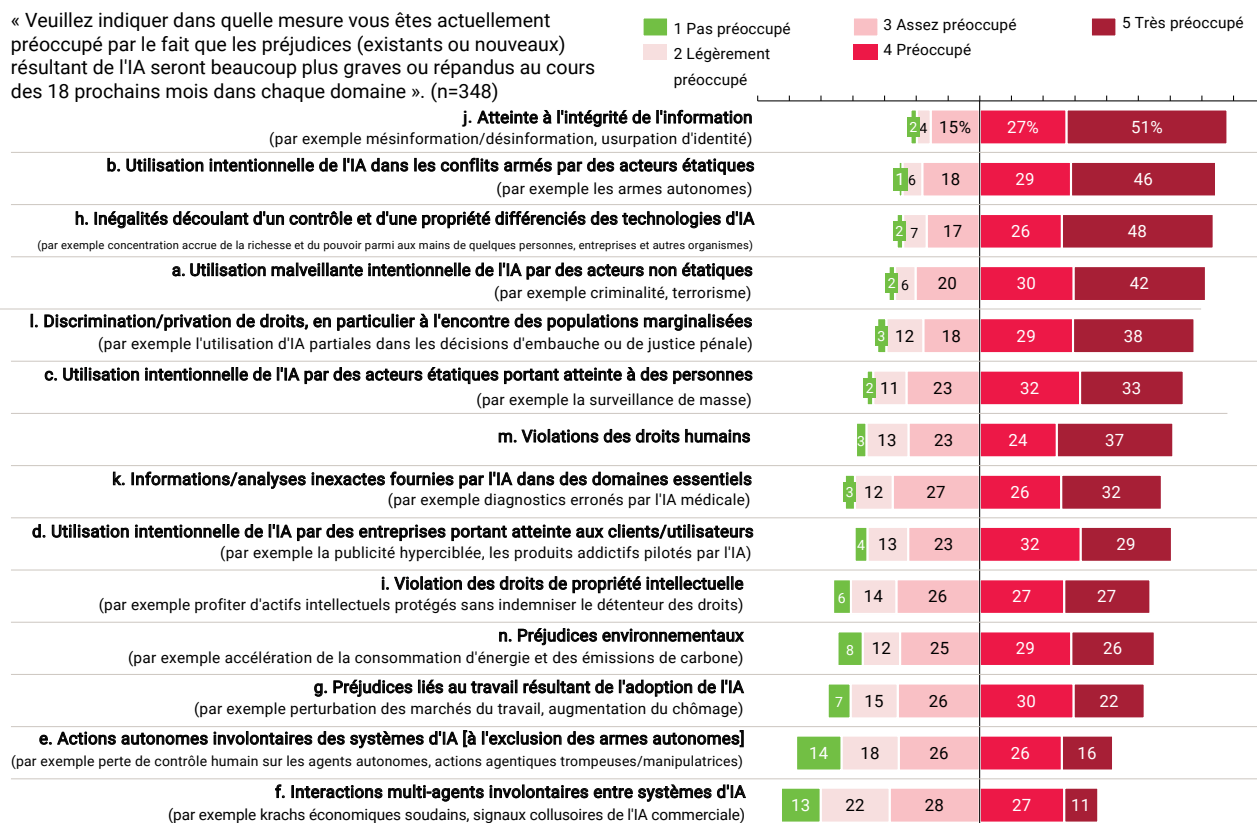
20 Dresser une liste exhaustive et définitive des risques liés à l'IA est une perte de temps. Compte tenu de l'omniprésence et de l'évolution rapide de l'IA et de ses utilisations, nous estimons qu'il est plus utile d'examiner les risques du point de vue des populations vulnérables et des biens communs (voir par. 26 à 28 ci-après).

21 Les résultats d'un tour d'horizon prospectif effectué dans le cadre de nos travaux (Enquête sur les risques liés à l'IA ; voir annexe E) illustrent la perception actuelle des risques par les experts. Ce

⁶ « Une analyse de la localisation du siège des bénéficiaires de subventions à partir d'une base de données de fondations à majorité américaine révèle qu'entre 2018 et 2023, seuls 10 % des subventions allouées à des initiatives d'IA portant sur un ou plusieurs des objectifs de développement durable ont été consacrées à des organisations basées dans des pays à faible revenu ou à revenu intermédiaire [...] L'analyse des capitaux privés montre que 36 % des 9 000 entreprises qui contribuent à réaliser les objectifs de développement durable ont leur siège aux États-Unis, mais que ces entreprises ont reçu 54 % du financement total. Nous avons également constaté que si 20 % des 9 000 entreprises qui contribuent à réaliser les objectifs de développement durable ont leur siège dans des pays à faible revenu ou à revenu intermédiaire, elles ont reçu une proportion plus élevée (25 %) du financement total. Cela s'explique notamment par le fait que les entreprises chinoises bénéficient d'une part importante des investissements [...] Les autres pays en développement de l'échantillon n'ont reçu que 3 % du financement alors qu'ils représentent 7 % de l'échantillon (Medha Bankhwal et al., « AI for social good: improving lives and protecting the planet » ; McKinsey and Company, mai 2024).

Figure 2 : Niveaux de préoccupation des experts concernant les risques liés à l'IA dans différents domaines

« Veuillez indiquer dans quelle mesure vous êtes actuellement préoccupé par le fait que les préjudices (existants ou nouveaux) résultant de l'IA seront beaucoup plus graves ou répandus au cours des 18 prochains mois dans chaque domaine ». (n=348)



Note : Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc.
Source : OSET Poulos des risques de l'IA, 13-25 mai 2024.

sondage a permis de recueillir les perceptions des tendances et des risques liés à l'IA auprès de 348 experts de l'IA de toutes les disciplines et de 68 pays de toutes les régions du monde⁷. Dans l'ensemble, 7 experts interrogés sur 10 se disent préoccupés ou très préoccupés par le fait que les préjudices (existants ou nouveaux) résultant de l'IA seront beaucoup plus graves ou répandus au cours des 18 prochains mois (voir annexe E).

22 Sur la base d'une liste d'exemples de risques liés à l'IA⁸, un grand nombre d'experts se sont dit préoccupés ou très préoccupés par les préjudices suivants (voir également fig. 2) :

- Implications sociétales de l'IA : **78 %** concernant l'atteinte à l'intégrité de l'information [question j.], **74 %** concernant les inégalités telles que la concentration des richesses et du pouvoir entre quelques mains [question h.], **67 %** concernant la discrimination/la privation des droits, en particulier parmi les populations marginalisées [question l] ;
- Utilisation intentionnelle de l'IA qui nuit à autrui : **75 %** concernent un conflit armé par des acteurs étatiques [questions b.], **72 %** concerne une utilisation malveillante par des acteurs non étatiques [question a.] et **65 %** concerne une utilisation par des acteurs étatiques qui nuit à des personnes [question c.].

7 La liste des invités a été établie à partir des réseaux du Bureau de l'Envoyé du Secrétaire général pour les technologies et de l'Organe consultatif, et comprend les participants aux études approfondies. D'autres experts ont été invités régulièrement au cours de la période de consultation afin d'améliorer la représentation. Le nombre final n=348 représente un échantillon global solide et équilibré de répondants disposant des compétences nécessaires pour donner un avis éclairé sur les risques liés à l'IA (voir la méthode complète à l'annexe E).

8 Établie sur la base de la catégorisation des risques fondée sur la vulnérabilité, indiquée à l'encadré 4, dont une version antérieure figurait également dans notre rapport d'étape.

23 La plupart des experts de l'IA interrogés se sont dits préoccupés ou très préoccupés par la matérialisation des préjudices dans tous les domaines de risque, à l'exception de deux. Bien que moins de la moitié des experts aient exprimé une telle inquiétude concernant les préjudices involontaires causés par l'IA [questions e et f], 1 personne sur 6 a indiqué qu'elle s'attendait à ce que les systèmes agentiques aient des impacts très surprenants ou majeurs sur les risques liés à l'IA d'ici à 2025⁹.

24 Les perceptions des experts ont varié, notamment en fonction de la région et du genre (voir des résultats plus détaillés à l'annexe E). Cela montre

l'importance d'une représentation inclusive dans les exercices concernant la définition des risques partagés. Malgré les différences, les résultats ont mis en évidence des préoccupations communes concernant les préjudices causés par l'IA au cours de l'année suivante, les experts estimant qu'il était urgent de traiter les risques dans de nombreux domaines et vulnérabilités dans un avenir proche.

25 En outre, les armes autonomes dans les conflits armés, la criminalité ou le terrorisme, et l'utilisation de l'IA en matière de la sécurité publique en particulier soulèvent d'importantes questions juridiques, sécuritaires et humanitaires (voir encadré 3)¹⁰.

Encadré 3 : L'IA et la sécurité nationale et internationale

De nombreuses technologies d'IA ne sont pas simplement à double usage, mais intrinsèquement réutilisables. Les applications de l'IA aux fins du maintien de l'ordre et des contrôles aux frontières se multiplient et soulèvent des inquiétudes quant au respect des procédures, à la surveillance et à l'absence de responsabilité concernant les engagements des États à l'égard des normes relatives aux droits humains inscrites dans la Déclaration universelle des droits de l'homme et dans d'autres instruments.

Parmi les défis posés par l'utilisation de l'IA dans le domaine militaire, il y a les nouvelles courses aux armements, l'abaissement du seuil de conflit, le brouillage des frontières entre la guerre et la paix, la multiplication des acteurs non étatiques et les dérogations aux principes établis de longue date du droit international humanitaire, tels que la nécessité militaire, la distinction, la proportionnalité et la limitation des souffrances inutiles. Pour des raisons juridiques et morales, les décisions de mort ne devraient pas être prises de manière automatisée par l'IA. Les États devraient s'engager à s'abstenir de déployer et d'utiliser les applications militaires de l'IA dans les conflits armés d'une manière qui ne soit pas pleinement conforme au droit international, notamment au droit international humanitaire et au droit des droits humains.

Actuellement, 120 États Membres de l'ONU soutiennent un nouveau traité sur les armes autonomes, et le Secrétaire général ainsi que la Présidente du Comité international de la Croix-Rouge ont tous deux demandé que les négociations relatives à ce traité soient achevées d'ici à 2026. L'Organe consultatif invite instamment les États Membres à donner suite à cet appel.

L'Organe consultatif estime qu'il est essentiel de définir des lignes rouges claires délimitant les cas d'utilisation illicite, notamment le recours à l'IA en vue de sélectionner et prendre des cibles à partie de manière autonome. Sur la base des engagements pris en matière d'examen des armes dans le cadre du droit international humanitaire, les États devraient exiger, notamment au moyen d'obligations contractuelles, que les fabricants d'armes procèdent à des examens juridiques et techniques afin d'empêcher la conception et la mise au point d'applications militaires de l'IA contraires à l'éthique. Ils devraient également procéder à des examens juridiques et techniques du recours à des armes et moyens de guerre utilisant l'IA et mettre en commun les meilleures pratiques en la matière.

9 Question : « Quelles sont les nouvelles tendances qui, selon vous, pourraient avoir les effets les plus surprenants ou les plus importants sur les risques liés à l'IA au cours des 18 prochains mois ? »

10 Cette liste n'a qu'une valeur indicative et ne traite que quelques-uns des risques auxquels les personnes et les sociétés font face.

Encadré 3 : L'IA et la sécurité nationale et internationale (suite)

En outre, les États devraient élaborer une vision commune des dispositifs d'essai, d'évaluation, de vérification et de validation de l'IA dans le domaine de la sécurité et de la défense. Ils devraient coopérer en vue de renforcer les capacités et mettre en commun les connaissances en échangeant les bonnes pratiques et en favorisant une gestion responsable du cycle de vie des applications de l'IA dans le domaine de la sécurité et de la défense. Afin d'empêcher l'acquisition de systèmes d'IA puissants et potentiellement autonomes par des acteurs non étatiques dangereux, tels que des groupes criminels ou terroristes, les États devraient mettre en place des contrôles et des processus appropriés tout au long du cycle de vie des systèmes d'IA, notamment la gestion de la fin du cycle de vie (à savoir le déclassé) des applications militaires de l'IA.

Aux fins de la transparence, des « conseils consultatifs » pourraient être mis en place afin de fournir des conseils techniques indépendants et d'assurer un suivi tout au long du cycle de vie des applications sécuritaires et militaires de l'IA. L'industrie et d'autres acteurs devraient envisager de mettre en place des dispositifs visant à empêcher l'utilisation abusive de la technologie de l'IA à des fins militaires malveillantes ou involontaires.

26 Toutefois, la gestion des risques ne se limite pas à dresser une liste des risques ou à les classer par ordre de priorité. La définition des risques sur la base des vulnérabilités peut axer les programmes stratégiques, non plus sur le type de risque (par exemple, « risque pour la sécurité »), mais sur la personne exposée, le lieu et le responsable dans chaque cas.

27 Il s'agit d'un point important car les risques évolutifs se manifestent différemment selon les personnes et les sociétés. Une approche fondée sur la vulnérabilité, également proposée dans notre rapport d'étape, fournit un cadre ouvert permettant de se concentrer sur les personnes auxquelles l'IA pourrait nuire, ce qui peut servir de base à une gestion dynamique des risques (voir encadré 4).

Encadré 4 : Catégorisation des risques liés à l'IA en fonction de la vulnérabilité existante ou potentielle

Personnes

- Dignité humaine, valeur ou moyens d'action (par exemple manipulation, tromperie, incitation, condamnation, exploitation, discrimination, égalité de traitement, poursuites, surveillance, perte d'autonomie humaine et ciblage assisté par l'IA).
- Intégrité physique et mentale, santé, sûreté et sécurité (par exemple, incitation, solitude et isolement, neurotechnologie, armes létales autonomes, voitures autonomes, diagnostics médicaux, accès aux soins de santé, interaction avec des systèmes chimiques, biologiques, radiologiques ou nucléaires).
- Chances à saisir dans la vie (par exemple éducation, emploi et logement).
- (Autres) Droits humains et libertés civiles, tels que les droits à la présomption d'innocence (par exemple police prédictive) et les droits à un procès équitable (par exemple prédiction de la récidive, culpabilité, récidive, prédiction et procès autonomes), à la liberté d'expression et d'information (par exemple incitation, informations personnalisées, info-bulles), à la protection de la vie privée (par exemple technologie de reconnaissance faciale) ainsi qu'à la liberté de réunion et de circulation (par exemple technologie de traçage dans les espaces publics).

Encadré 4 : Catégorisation des risques liés à l'IA en fonction de la vulnérabilité existante ou potentielle (suite)

Politique et société

- Discrimination et traitement inéquitable de groupes, y compris sur la base de caractéristiques individuelles ou collectives telles que le genre, l'isolement et la marginalisation de groupes.
- Impact différentiel sur les enfants, les personnes âgées, les personnes handicapées et les groupes vulnérables.
- Sécurité internationale et nationale (par exemple armes autonomes, maintien de l'ordre et contrôle des frontières vis-à-vis des migrants et des réfugiés, criminalité organisée, terrorisme, multiplication et escalade des conflits).
- Démocratie (par exemple élections et confiance).
- Intégrité de l'information (par exemple mésinformation ou désinformation, hypertrucages et nouvelles personnalisées).
- État de droit (par exemple fonctionnement des institutions, des services de maintien de l'ordre et du système judiciaire, et confiance à leur égard).
- Diversité culturelle et évolution des relations humaines (par exemple homogénéité et faux amis).
- Cohésion sociale (par exemple bulles de filtres, baisse de la confiance dans les institutions et sources d'information).
- Valeurs et normes (par exemple éthiques, morales, culturelles et juridiques).

Économie

- Concentration du pouvoir.
- Dépendance technologique.
- Inégalité des débouchés économiques, de l'accès au marché, ainsi que de la distribution et de l'allocation des ressources.
- Sous-utilisation de l'IA.
- Surutilisation de l'IA ou « solutionnisme technique ».
- Stabilité des systèmes financiers, des infrastructures essentielles et des institutions.
- Protection de la propriété intellectuelle.

Environnement

- Consommation excessive d'énergie, d'eau et de ressources matérielles (y compris les minéraux rares et d'autres ressources naturelles).

28 L'examen des considérations relatives à la gouvernance de l'IA du point de vue d'un groupe vulnérable particulier, tel que les enfants, illustre la pertinence stratégique d'une approche des risques liés à l'IA fondée sur la vulnérabilité (voir encadré 5).

29 Les personnes, groupes ou entités concernés recensés au moyen d'une analyse des risques liés à l'IA fondée sur la vulnérabilité – et les stratégies qui en découlent – peuvent varier. Dans le cadre

de l'enquête sur les risques de l'IA, on a également demandé aux experts à quels personnes, groupes, sociétés/économies/(éco)systèmes ils craignaient que l'IA ne porte préjudice au cours des 18 mois suivants. Les populations marginalisées et le monde du Sud, ainsi que les enfants, les femmes, les jeunes, les créatifs et les personnes occupant les emplois les plus exposés à l'automatisation ont été les plus cités (voir fig. 3).

Encadré 5 : Placer les enfants au centre de la gouvernance de l'IA

Afin que les entreprises et les écoles répondent aux besoins et aux droits des enfants, il faut une approche globale de la gouvernance qui tienne compte de leur situation particulière. Les enfants génèrent un tiers des données et grandiront dans une économie et un monde infusés par l'IA et habitués à l'utiliser. Le présent encadré contient un résumé de certaines des mesures examinées lors de nos travaux de recherche approfondis sur ce sujet.

Donner la priorité aux droits et à la voix des enfants :

La gouvernance de l'IA doit prendre en compte les enfants en tant qu'acteurs prioritaires, en particulier leur droit à se développer à l'abri des effets d'accoutumance de la technologie et à s'en libérer. Contrairement aux approches générales centrées sur l'homme, la gouvernance centrée sur l'enfant doit prendre en compte les effets à long terme sur la perception des enfants, l'image qu'ils ont d'eux-mêmes, leurs choix de vie et les possibilités qui s'offrent à eux. Il est essentiel d'intégrer les enfants dans les processus de conception et de gouvernance afin que les systèmes d'IA soient sûrs et adaptés à leur utilisation.

Travaux de recherche et élaboration de politiques :

Des travaux de recherche approfondis sont nécessaires pour comprendre les effets de l'IA sur le développement social, cognitif et émotionnel des enfants au fil du temps. Ils devraient éclairer les débats de politique générale et orienter les mesures de protection dans les différents pays.

Protection et respect de la vie privée :

Les enfants ne devraient pas être utilisés comme sujets d'expérimentation de l'IA. La protection de la vie privée des enfants est primordiale. Les technologies d'IA doivent intégrer des protocoles rigoureux de protection des données et fournir un contenu adapté à l'âge des utilisateurs.

Études d'impact sur les enfants et conception adaptée aux enfants :

Il faut rendre obligatoires les études d'impact des systèmes d'IA sur les enfants afin de garantir leur adéquation et leur sécurité. Il faut concevoir des systèmes d'IA qui tiennent compte des besoins des enfants et intègrent dès le départ des dispositifs de sécurité et de restriction. Les choix de conception doivent tenir compte de l'avis des enfants.

Inclusion numérique et équité :

L'accès à l'IA devrait permettre aux enfants d'agir, de choisir et de s'exprimer, l'accent étant mis sur des approches globales de l'inclusion numérique. Il s'agit notamment de fournir le contenu de l'IA en plusieurs langues et de veiller à ce qu'il soit culturellement adapté aux enfants qui ne parlent pas l'anglais.

Coopération et normes internationales :

L'interopérabilité mondiale des règles relatives à l'utilisation des technologies de l'IA par les enfants doit permettre de les protéger dans différents environnements d'éducation et de développement. Il sera essentiel d'établir des normes mondiales pour traiter les flux de données transfrontaliers et l'utilisation éthique de l'IA pour les enfants.

Figure 3 : Préoccupations liées à la vulnérabilité recensées dans l'enquête sur les risques liés à l'IA

« Craignez-vous que l'IA porte préjudice à des personnes, groupes ou sociétés/économies/(éco)systèmes au cours des 18 prochains mois ? » [réponse en texte libre] (n=188 réponses utiles à cette question)

Indicatif



Note : les mots-clés ont été marqués pour chaque réponse par le Bureau de l'Envoyé du Secrétaire général pour les technologies. Afficher uniquement les mots-clés recensés dans plus de deux réponses. La taille des caractères est proportionnelle au nombre de réponses mentionnées. À titre d'exemple, le « monde du Sud » a été cité par 46 des 188 répondants qui ont fourni des réponses utiles à cette question ; « populations marginalisées » par 43 répondants sur 188. Source : OSET Poulos des risques de l'IA, 13-25 mai 2024.

30 Les résultats illustrent l'importance d'une représentation inclusive pour parvenir à une vision commune des risques liés à l'IA et à un socle commun sur les programmes stratégiques, conformément aux recommandations 1 et 2. En l'absence d'une telle représentation, les programmes stratégiques en matière de gouvernance de l'IA pourraient être conçus sans tenir compte des préoccupations d'une partie de l'humanité qui sera néanmoins concernée.

F. Difficultés à résoudre

31 Outre les risques et les préjudices à venir, l'évolution de la mise au point, du déploiement et de l'utilisation de l'IA pose également des difficultés dans le contexte des organismes existants, et celles-ci ont également des effets sur les stratégies de gouvernance de l'IA. Le rythme technologique de l'IA avancée – et son caractère polyvalent – met encore plus à l'épreuve la capacité de l'humanité à réagir à temps.

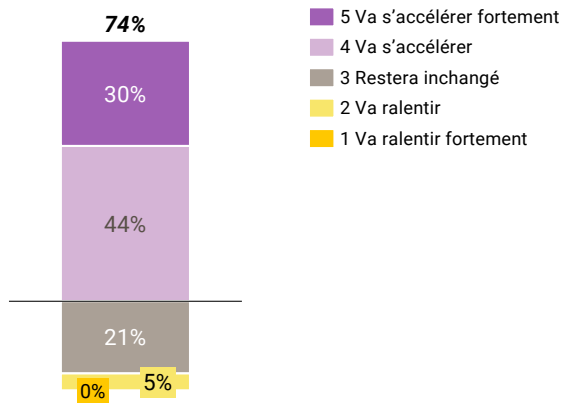
32 La course à la mise au point et au déploiement des systèmes d'IA pose problème pour les systèmes réglementaires et les régimes de gouvernance traditionnels. La plupart des experts interrogés dans le cadre de l'enquête sur les risques liés à l'IA s'attendaient à une accélération de l'IA au cours des 18 prochains mois, tant au niveau de son développement (74 %) que de son adoption et application (89 %) (voir fig. 4).

33 Comme indiqué au paragraphe 23 ci-dessus, certains experts prévoient le déploiement de systèmes agentiques en 2025. En outre, les principaux experts techniques admettent que de nombreux modèles d'IA restent opaques et que leurs résultats ne sont pas entièrement prévisibles ou contrôlables, alors même que les effets négatifs indirects peuvent toucher d'autres personnes à l'échelle mondiale.

Figure 4 : Attentes des experts concernant la mise au point technologique de l'IA

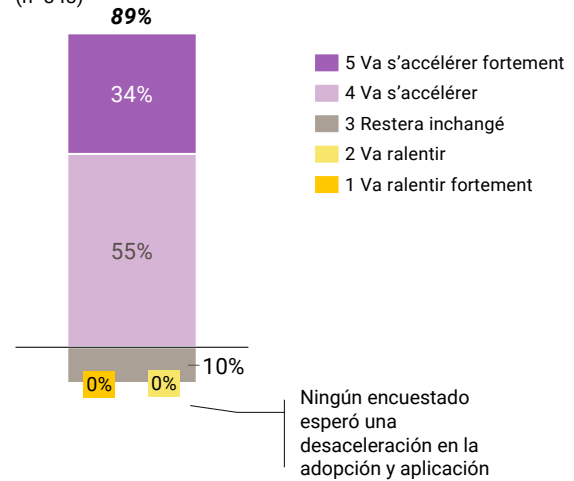
74 % s'attendent à une accélération du rythme des changements techniques (30 % à une forte accélération)

« Au cours des 18 prochains mois, par rapport aux 3 derniers mois, pensez-vous que le rythme des changements techniques dans le domaine de l'IA (par exemple la mise au point/lancement de nouveaux modèles)... » (n=348)



89 % s'attendent à une accélération du rythme d'adoption et d'application (34 % à une forte accélération)

« Au cours des 18 prochains mois, par rapport aux 3 derniers mois, pensez-vous que le rythme d'adoption et d'application de l'IA (par exemple, les nouvelles utilisations de l'IA dans les entreprises/l'administration)... » (n=348)



Note : Aucun répondant ne s'attend à un ralentissement de l'adoption et de l'application. Les chiffres ayant été arrondis, leur somme peut ne pas correspondre à 100 %.
Source : OSET Poulx des risques de l'IA, 13-25 mai 2024.

34 Le recours croissant à la prise de décision automatisée et à la création de contenu par des algorithmes opaques peut nuire à l'équité de traitement et à la sûreté. Si les êtres humains restent souvent juridiquement responsables des décisions d'automatiser des processus qui ont des effets sur autrui, les dispositifs de reddition des comptes n'évoluent pas toujours assez rapidement pour que cette responsabilité soit rapidement et réellement mise en œuvre.

35 Ainsi, il existe un risque sociétal qu'un nombre de plus en plus restreint de personnes soient tenus responsables des préjudices résultant de leurs décisions d'automatiser des processus à l'aide de l'IA, alors même que des systèmes de plus en plus puissants font leur apparition à travers le monde. Cela exige une gouvernance agile afin de garantir que les dispositifs de reddition des comptes suivent l'accélération de l'IA.

36 Si le rythme de mise au point et de déploiement de l'IA pose problème aux organismes existants, il en va de même pour son ampleur. Technologie polyvalente d'envergure mondiale, l'IA avancée peut être déployée dans tous les domaines et avoir de nombreux effets sur les sociétés, avec de vastes implications stratégiques.

37 Les implications et les possibles effets de l'intersection de l'IA avec de nombreux domaines, notamment la finance, les marchés du travail, l'éducation et les systèmes politiques, laissent présager de vastes conséquences qui exigent une approche de l'ensemble de la société (voir les exemples figurant à l'encadré 6). Les organismes existants doivent mettre en place des mesures globales et intersectorielles pour faire face aux vastes effets sociétaux de l'IA.

Encadré 6 : Effets sociétaux liés à l'IA

Dans le cadre de leur collaboration plus large, les membres de l'Organe consultatif ont consulté une série de parties prenantes afin d'examiner des implications de l'IA sur la société. L'encadré présente un résumé des principales préoccupations et possibles initiatives mises en avant dans le cadre des études approfondies sur ce sujet.

Effets sociétaux, psychologiques et communautaires :

À mesure que l'IA devient plus puissante et plus répandue, sa mise au point, son déploiement et son application deviendront plus personnalisés, quitte à développer l'aliénation et la dépendance. Pour certains membres de l'Organe consultatif, l'IA entraînée à partir des données d'une personne et le rôle d'interlocuteur et d'intermédiaire qui en découle peut représenter un tournant pour les êtres humains car elle peut créer de nouveaux défis sociétaux urgents et exacerber les défis existants.

Par exemple, les futurs systèmes d'IA pourraient être en mesure de générer un flux continu de contenu vidéo de haute qualité adapté aux préférences personnelles des personnes, avec pour conséquences le renforcement de l'isolement social, l'aliénation, les problèmes de santé mentale, la perte d'autonomie et les répercussions sur l'intelligence émotionnelle et le développement social, entre autres.

Si peu d'études ont été menées sur ces questions par les décideurs dans le contexte de technologies telles que les appareils intelligents et Internet, on n'en trouve quasiment aucune dans le contexte de l'IA, les cadres de gouvernance actuels privilégiant les risques pour les personnes plutôt que pour la société dans son ensemble.

Lorsque les décideurs réfléchissent aux mesures à adopter en matière d'IA à l'avenir, ils doivent également tenir compte de ces facteurs et élaborer des politiques qui favorisent le bien-être de la société, en particulier celui des jeunes. Les pouvoirs publics pourraient favoriser des environnements qui privilégient les interactions en face à face entre les êtres humains, en facilitant l'accès aux services de santé mentale et en investissant davantage dans les installations sportives, les bibliothèques publiques et les arts.

Toutefois mieux vaut prévenir que guérir : les développeurs devraient concevoir leurs produits sans caractéristiques personnalisées créant une dépendance, veiller à ce qu'ils ne nuisent pas à la santé mentale et promouvoir – plutôt que saper – un sentiment d'appartenance commune à la société. Les entreprises technologiques devraient établir des politiques de gestion des risques sociétaux au même titre que d'autres risques, dans le cadre de leurs efforts d'identification et d'atténuation des risques tout au long du cycle de vie des produits d'IA.

Désinformation et confiance :

Les hypertrucages, les clones vocaux et les campagnes de désinformation automatisées constituent une menace spécifique et grave pour les institutions et les processus démocratiques tels que les élections, ainsi que pour les sociétés démocratiques et la confiance sociale de manière plus générale, notamment par la manipulation de l'information et l'ingérence étrangères. Le développement d'écosystèmes d'information en boucle fermée, renforcés par l'IA et exploitant les données personnelles, peut avoir des effets profonds sur les sociétés, les rendant davantage susceptibles d'accepter l'intolérance et la violence à l'égard d'autrui.

Encadré 6 : Effets sociétaux liés à l'IA (suite)

La protection de l'intégrité des institutions et des processus gouvernementaux représentatifs nécessite des systèmes robustes de vérification et de détection des faux contenus, ainsi que des procédures de notification et de retrait rapides lorsque ces contenus sont susceptibles de tromper d'une manière qui cause des préjudices ou des divisions sociétales, ou qui encourage la propagande de guerre, les conflits et les discours de haine. Il convient de protéger les personnes qui ne sont pas des personnalités publiques face à la création d'hypertrucages de leur image à des fins frauduleuses, diffamatoires ou abusives. Les trucages sexualisés, qui sont particulièrement préoccupants pour les femmes et les jeunes filles, peuvent constituer une forme de violence fondée sur le genre.

Les engagements pris à titre volontaire par les acteurs du secteur privé, qui consistent notamment à signaler les hypertrucages ou à permettre aux utilisateurs de marquer puis de retirer les hypertrucages fabriqués ou diffusés avec des intentions malveillantes, constituent une première étape importante. Toutefois, cela ne permet pas d'atténuer suffisamment les risques sociétaux. Il faut adopter une approche globale et multipartite, ainsi que des engagements contraignants. Des normes communes pour l'authentification des contenus et la provenance numérique permettraient d'utiliser des techniques mondialement reconnues afin de détecter des images, vidéos et sons synthétiques et modifiés par l'IA.

En outre, l'échange de connaissances en temps réel entre les acteurs publics et privés, sur la base de normes internationales, permettrait de mettre en place des capacités de réaction rapide afin de supprimer immédiatement les contenus trompeurs, y compris ceux liés à la manipulation et l'ingérence étrangères, avant qu'ils ne deviennent viraux. Néanmoins, ces processus doivent être assortis de garanties afin d'éviter qu'ils ne fassent l'objet d'une manipulation ou de dérives visant à favoriser la censure.

Ces actions devraient être accompagnées de mesures préventives afin d'accroître la résistance de la société à la désinformation et à la propagande induites par l'IA, telles que des campagnes de sensibilisation du public aux risques que l'IA porte atteinte à l'intégrité de l'information. Les États Membres devraient en outre promouvoir des campagnes d'initiation aux médias et au numérique, soutenir les initiatives de vérification des faits et investir dans le renforcement des capacités de celles et ceux qui luttent contre la manipulation de l'information et l'ingérence étrangères.

- 38** Le rythme, l'ampleur et l'incertitude de la mise au point, du déploiement et de l'utilisation de l'IA montrent qu'il importe d'adopter une approche globale, transversale et agile de l'IA. Au niveau international, une perspective globale doit se traduire par une approche institutionnelle en réseau de la gouvernance de l'IA à travers les secteurs et les frontières, qui mobilise les parties prenantes sans se laisser accaparer par elles.
- 39** S'agissant des changements climatiques, le monde n'a pris conscience que tardivement que l'action collective mondiale devait reposer sur une approche globale. Avec l'IA, il est possible de le faire dès la conception.
- 40** Aux difficultés susmentionnées s'ajoutent la concentration des richesses et la prise de décision aux mains d'une poignée de développeurs et de déployeurs privés en matière d'IA, en particulier des entreprises multinationales. Cela soulève une autre question : comment les parties prenantes peuvent-elles être impliquées dans la gouvernance de l'IA sans nuire à l'intérêt général ?

2. La nécessité d'une gouvernance mondiale

- 41** Il existe aujourd'hui un manque de gouvernance mondiale en matière d'IA. Malgré de nombreux débats sur l'éthique et les principes, la mosaïque de normes, d'organismes et d'initiatives est encore peu développée et présente de grandes lacunes. Le principe de responsabilité et les recours en cas de préjudice brillent souvent par leur absence. Le respect des règles repose sur le volontariat. Il existe un décalage fondamental entre la rhétorique de haut niveau, les systèmes développés, déployés et utilisés, et les conditions requises pour la sûreté et l'inclusion. Comme nous l'avons indiqué dans notre rapport d'étape, la gouvernance de l'IA est cruciale, non seulement pour traiter les difficultés et les risques, mais aussi pour veiller à exploiter ses atouts sans laisser personne de côté¹¹.
- 42** La nécessité de mettre en place une gouvernance mondiale, en particulier, est incontestable. Les matières premières de l'IA, qu'il s'agisse de minéraux essentiels ou de données d'entraînement proviennent du monde entier. L'IA à usage général, déployée au-delà des frontières, génère de nombreuses applications à travers le monde. Le développement accéléré de l'IA concentre le pouvoir et la richesse à l'échelle mondiale, ce qui a des conséquences géopolitiques et géoéconomiques. En outre, personne n'est actuellement en mesure de comprendre tous les rouages de l'IA au point de contrôler pleinement ses résultats ou de prédire son évolution. La responsabilité des décideurs n'est pas non plus engagée dans la mise au point, le déploiement ou l'utilisation de systèmes qu'ils ne maîtrisent pas. Par ailleurs, les répercussions négatives, y compris indirectes, que ces décisions engendrent peuvent également avoir une portée mondiale.
- 43** Malgré la portée mondiale de l'IA, les structures institutionnelles et les réglementations nationales et régionales s'arrêtent aux frontières physiques. Cela réduit la capacité de tout pays à régir les applications en aval de l'IA qui entraînent des préjudices transfrontaliers, ou à traiter les problèmes le long des chaînes d'approvisionnement transfrontalières complexes de l'infrastructure informatique, des flux de données d'entraînement et des sources d'énergie qui sous-tendent la mise au point et l'utilisation de l'IA. Les grandes entreprises en matière d'IA ont souvent une influence plus directe sur les applications en aval (en atténuant les risques en amont) que la plupart des pays agissant seuls.
- 44** La mise au point, le déploiement et l'utilisation de la technologie ne peuvent pas être laissés aux aléas des marchés. Les gouvernements nationaux et les organisations régionales auront un rôle crucial à jouer. Toutefois, outre les considérations d'équité, d'accès, de prévention et de réparation des préjudices, le caractère transfrontière de la structure et de l'application de la technologie appelle une approche multisectorielle mondiale. En l'absence d'un cadre global qui mobilise les parties prenantes, et compte tenu de la dynamique concurrentielle en jeu, les gouvernements et les entreprises pourraient être tentés de rogner sur les coûts afin de donner la priorité à leurs intérêts personnels.
- 45** L'IA présente donc des obstacles à surmonter et des chances à saisir qui nécessitent une approche globale pour traiter de manière transversale des domaines tels que la politique, l'économie, la société, l'éthique, les droits humains, la technique et l'environnement. Cette approche peut transformer

11 Voir <https://un.org/ai-advisory-body>.

une mosaïque d'initiatives en constante évolution en un dispositif cohérent et interopérable, fondé sur le droit international et adaptable à tous les contextes et à toutes les époques.

- 46** Le besoin d'une gouvernance mondiale de l'IA se manifeste au moment où la concurrence géopolitique et géoéconomique pour l'influence et les marchés s'intensifie. Cependant, il faut mener une action concertée au niveau mondial pour faire face aux risques de l'IA et en même temps tirer parti équitablement des possibilités qu'elle offre. L'élargissement de la fracture numérique pourrait limiter les avantages de l'IA à une poignée d'États et de personnes, alors que les risques et les préjudices toucheraient un grand nombre de personnes, en particulier les groupes vulnérables.

A. Principes directeurs et fonctions de la gouvernance internationale de l'IA

- 47** Dans notre rapport d'étape, nous avons exposé cinq principes qui devraient orienter la création de nouveaux organismes internationaux de gouvernance de l'IA.
- **Principe directeur 1** : l'IA doit être gouvernée de manière inclusive, par toutes et par tous et dans l'intérêt commun
 - **Principe directeur 2** : l'IA doit être gouvernée dans l'intérêt général
 - **Principe directeur 3** : la gouvernance de l'IA doit être articulée avec la gouvernance des données et la promotion des biens communs en matière de données
 - **Principe directeur 4** : la gouvernance de l'IA doit être universelle, en réseau et ancrée dans une collaboration multipartite adaptative
 - **Principe directeur 5** : la gouvernance de l'IA devrait prendre ses racines dans la Charte des Nations Unies, le droit international des droits de l'homme et d'autres engagements convenus au niveau international, tels que les objectifs de développement durable.

- 48** L'encadré 7 présente un résumé des commentaires reçus sur ces principes, qui soulignent l'importance des droits humains et le besoin de clarifier davantage la mise en œuvre effective des principes directeurs, notamment en ce qui concerne la gouvernance des données. Il nous a donc fallu chercher à appuyer l'inclusivité par des mesures concrètes et veiller à ce que les groupes marginalisés soient représentés.

- 49** Dans notre rapport d'étape, nous avons également proposé plusieurs fonctions institutionnelles qui pourraient être exercées au niveau international (voir fig. 5). Les commentaires reçus ont largement confirmé qu'il fallait établir ces fonctions au niveau mondial et appeler à d'autres fonctions complémentaires liées à la gouvernance des données et de l'IA afin de mettre en pratique le principe directeur 3 (la gouvernance de l'IA doit être articulée avec la gouvernance des données et la promotion des biens communs en matière de données).

- 50** S'agissant des fonctions institutionnellement plus « rigides » de la gouvernance de l'IA, à savoir le contrôle, la vérification, l'établissement de rapports, la conformité, la stabilisation de la responsabilité, la réaction et l'application, il a été observé qu'il faudrait d'abord appliquer les obligations découlant de traités internationaux avant d'institutionnaliser ces fonctions, et que les arguments en faveur de l'institutionnalisation de ces fonctions dans la gouvernance de l'IA en tant que technologie n'avaient pas encore convaincu.

- 51** Il n'est pas indispensable que l'ONU exerce à elle seule toutes les fonctions. Toutefois, si l'on veut que la mosaïque de normes et d'organismes serve de filet de sécurité pour promouvoir et soutenir l'innovation durable au bénéfice de l'ensemble de l'humanité, il faut une vision commune de la science et un socle commun concernant les règles et les normes qui nous permettent de déterminer si la gouvernance atteint ses objectifs.

- 52** Au cours de nos consultations, nous avons entendu des appels en faveur d'une analyse plus détaillée du paysage des efforts existants et émergents pour gouverner l'IA au niveau international, et des lacunes à combler pour une gouvernance internationale équitable, efficace et efficiente de l'IA.

Encadré 7 : Commentaires sur les principes directeurs

Approche de la gouvernance de l'IA fondée sur les droits humains :

Sur la base des vastes consultations menées par l'Organe consultatif à la suite de la publication de son rapport d'étape, le principe directeur 5 (la gouvernance de l'IA devrait prendre ses racines dans la Charte des Nations Unies, le droit international des droits de l'homme et d'autres engagements convenus au niveau international) a recueilli le soutien le plus large de tous les secteurs de parties prenantes, notamment les gouvernements, la société civile, la communauté technique, le milieu universitaire et le secteur privé. Il s'agit notamment du respect, de la promotion et de la réalisation des droits humains et de la poursuite de leurs violations, ainsi que de la résolution de l'Assemblée générale 78/265 intitulée « Saisir les possibilités offertes par des systèmes d'intelligence artificielle sûrs, sécurisés et dignes de confiance pour le développement durable », adoptée à l'unanimité en mars 2024.

Au cours des débats, il est apparu clairement à l'Organe consultatif qu'afin d'atténuer les risques et les préjudices liés à l'IA, traiter les nouveaux cas d'utilisation et veiller à ce que l'IA puisse réellement profiter à l'ensemble de l'humanité et ne laisser personne de côté, les droits humains en particulier doivent être au centre de la gouvernance de l'IA de manière à garantir une responsabilité fondée sur les droits dans tous les pays. Cet engagement fondamental en faveur des droits humains est transversal et s'applique à toutes les recommandations formulées dans le présent rapport final.

Dispositifs de mise en œuvre spécifiques et clarté sur les lignes directrices :

De nombreuses parties prenantes ont souligné qu'il fallait élaborer des plans d'action détaillés et des lignes directrices claires afin de garantir une mise en œuvre efficace des principes directeurs formulés sur la gouvernance internationale de l'IA par l'Organe consultatif. Des organismes publics ont proposé d'élaborer des recommandations claires pour définir et garantir l'intérêt général, ainsi que des dispositifs de participation et de contrôle du public. Les entités du secteur privé ont souvent insisté sur la nécessité d'adopter des politiques claires et de tirer parti des cadres réglementaires existants pour maintenir des marchés de l'IA compétitifs et novateurs. De nombreuses organisations internationales et organisations de la société civile ont également appelé à la mise en place de systèmes de gouvernance souples, conçus pour répondre en temps utile à l'évolution des technologies. D'autres ont expressément demandé la création d'une nouvelle entité dotée de « muscles et de dents », dont le rôle ne se limiterait pas à une simple coordination.

Dispositifs visant à responsabiliser les acteurs clés :

Une préoccupation commune a été l'application du principe de responsabilité concernant le caractère discriminatoire, partiel et préjudiciable de l'IA, avec des propositions de création de dispositifs visant à assurer la reddition de comptes et la réparation des préjudices, et à traiter la question de la concentration des capacités technologiques et de la puissance de marché. De nombreuses organisations ont souligné qu'il fallait s'attaquer au pouvoir incontrôlé, et garantir les droits des consommateurs et une concurrence loyale. Les institutions universitaires ont souligné les points forts des principes directeurs, à savoir leur caractère universel et inclusif, mais ont proposé d'améliorer la mobilisation des parties prenantes. Les acteurs du secteur privé ont mis l'accent sur l'utilisation responsable de l'IA et sur la suppression des obstacles à l'accès.

Encadré 7 : Commentaires sur les principes directeurs (suite)

Fonctions plus concrètes sur la gouvernance des données d'IA :

L'absence de systèmes de gouvernance des données a été mentionnée dans de nombreuses consultations, les parties prenantes indiquant que l'ONU était une enceinte adaptée pour les dialogues sur la gouvernance des données. Les gouvernements ont souligné qu'il fallait mettre en place des cadres solides de gouvernance des données qui donnaient la priorité à la protection de la vie privée, à la protection des données et à l'utilisation équitable des données, et plaidé en faveur de lignes directrices internationales pour gérer la complexité des données dans le cadre de la mise au point de l'IA. Il a été demandé que les cadres s'inscrivent dans un processus transparent et inclusif, intégrant des considérations éthiques telles que le consentement et la protection de la vie privée.

Le milieu universitaire a souligné que la gouvernance des données devrait être traitée en priorité à court terme. Les entités du secteur privé ont noté que des mesures de gouvernance des données devraient venir compléter la gouvernance de l'IA, notamment des lois complètes sur la protection de la vie privée et une utilisation responsable de l'IA. Les organisations internationales et les organisations de la société civile ont souligné que la gouvernance des données d'entraînement de l'IA devrait protéger les droits des consommateurs et favoriser une concurrence loyale entre les développeurs d'IA grâce à un accès non exclusif aux données d'entraînement, soulignant l'appel à des mesures concrètes et réalisables en matière de gouvernance des données. L'ONU a été considérée comme un lieu idoine pour relever ces défis de gouvernance et combler les disparités en matière de ressources.

Figure 5 : Fonctions de gouvernance de l'IA proposées au niveau international

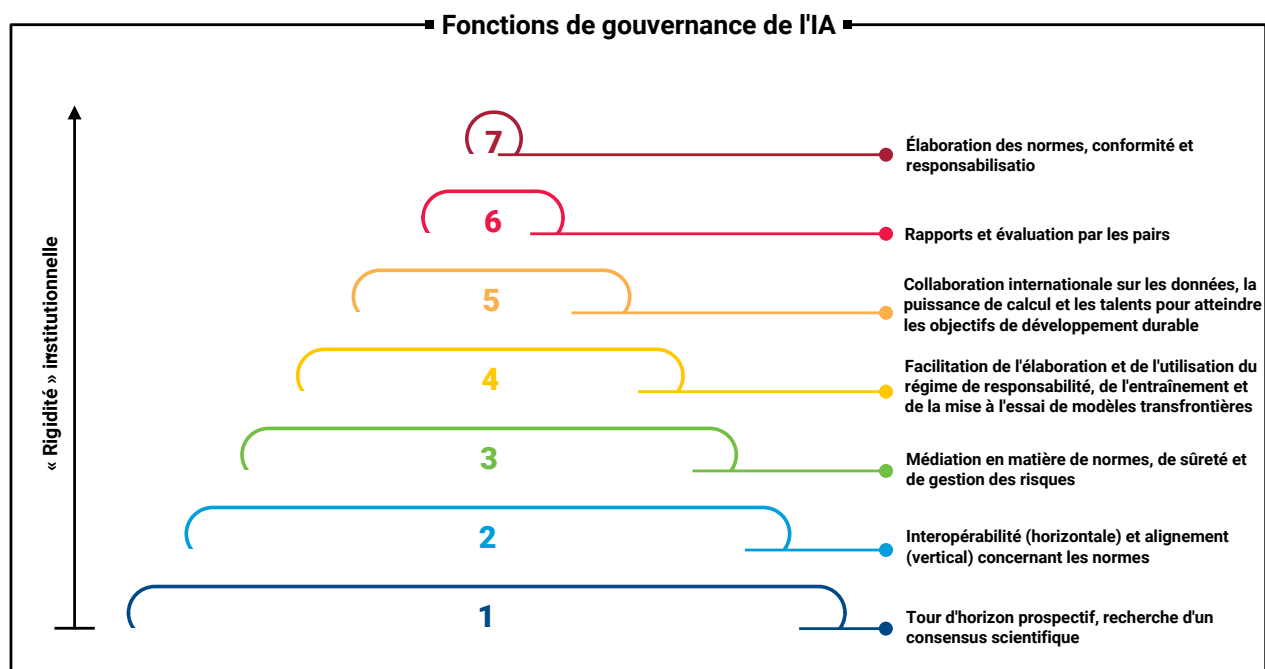
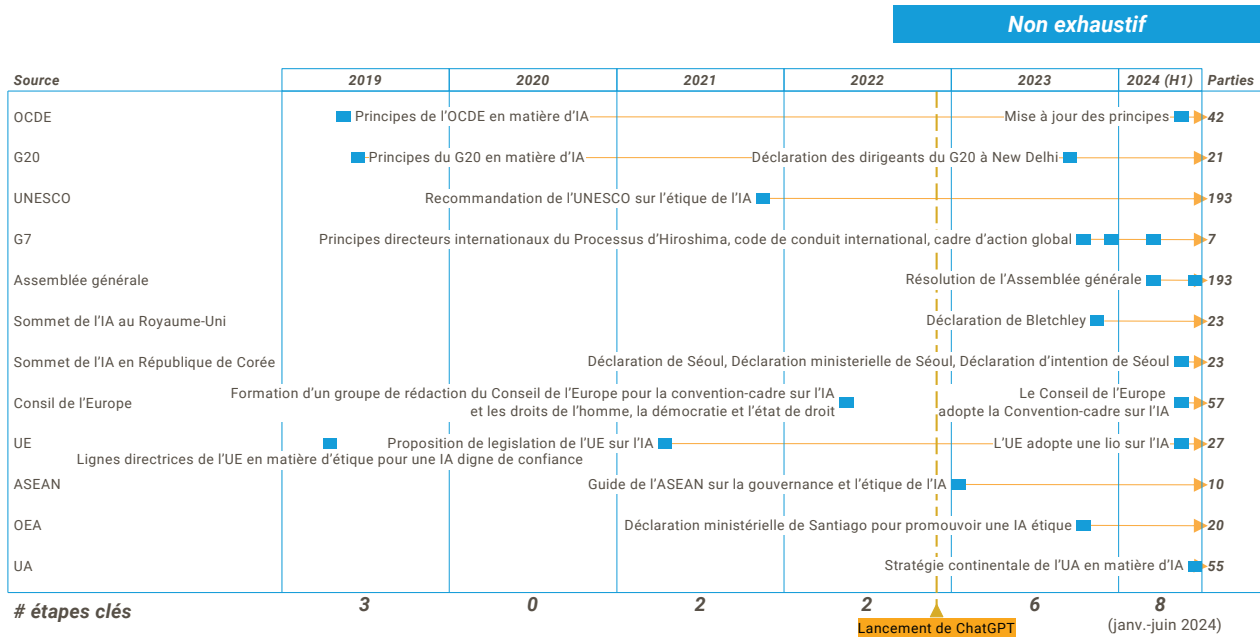


Figure 6 : Initiatives interrégionales et régionales de gouvernance de l'IA, étapes clés 2019-2024



Abbreviations : ASEAN, Association des nations de l'Asie du Sud-Est ; UA, Union africaine ; Conseil de l'Europe ; UE, Union européenne ; G20, Groupe des 20 ; G7, Groupe des Sept ; PMIA, Partenariat mondial sur l'intelligence artificielle ; OEA, Organisation des États américains ; OCDE, Organisation de coopération et de développement économiques ; UNESCO, Organisation des Nations Unies pour l'éducation, la science et la culture.

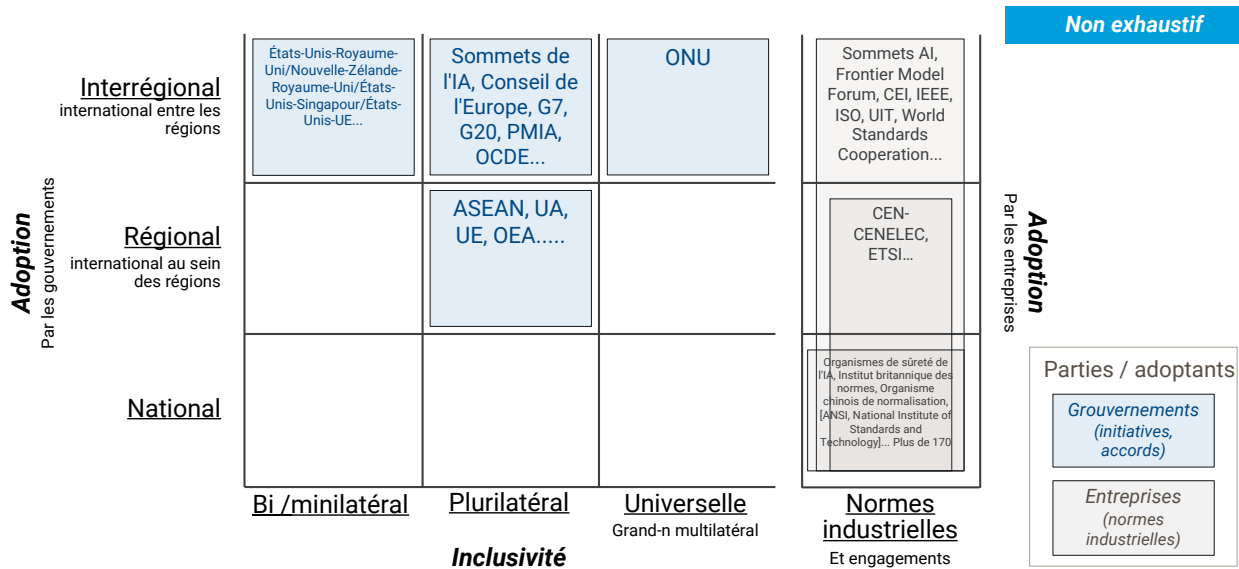
B. Paysage international émergent de la gouvernance de l'IA

53 Il est certain que les documents et débats consacrés actuellement à la gouvernance de l'IA ne manquent pas. Des entreprises et des consortiums ainsi que des organisations régionales et internationales ont adopté des centaines de guides, de cadres et de principes. Des dizaines d'enceintes réunissent des acteurs divers, qu'il s'agisse de processus intergouvernementaux établis, d'organes d'experts ou d'initiatives multipartites ad hoc. Il existe en outre une réglementation aux niveaux national et régional, qui continue d'évoluer.

54 Les initiatives internationales des gouvernements se multiplient (voir fig. 6). Ces nouvelles initiatives adoptent de plus en plus une approche transversale de la gouvernance de l'IA au niveau international, qui comprend des principes, des déclarations et d'autres textes qui traitent de l'IA de manière globale, plutôt que dans des domaines particuliers. Elles se sont fortement accélérées depuis 2023, stimulées par la publication de plusieurs grands modèles de langage d'IA à usage général, après le lancement de ChatGPT en novembre 2022.

55 Parallèlement, des normes industrielles sur l'IA ont été élaborées et publiées en vue de leur adoption au niveau international. D'autres initiatives multipartites ont également cherché à combler le fossé entre le secteur public et le secteur privé, y compris dans des enceintes comme le Forum sur la gouvernance d'Internet.

Figure 7 : Sources des initiatives de gouvernance consacrées à l'IA



Abbréviations : ANSI, Institut national américain de normalisation ; ASEAN, Association des nations de l'Asie du Sud-Est ; UA, Union africaine ; BSI, Institut britannique des normes ; CEN, Comité européen de normalisation ; CENELEC, Comité européen de normalisation électrotechnique ; Conseil de l'Europe ; ETSI, Institut européen de normalisation des télécommunications ; UE, Union européenne ; Frontier Model Forum ; G20, Groupe des 20 ; G7, Groupe des Sept ; PMIA, Partenariat mondial sur l'intelligence artificielle ; CEI, Commission électrotechnique internationale ; IEEE, Institute of Electrical and Electronic Engineers ; ISO, Organisation internationale de normalisation ; UIT, Union internationale des télécommunications ; National Institute of Standards and Technology ; OEA, Organisation des États américains ; OCDE, Organisation de coopération et de développement économiques ; Organisme chinois de normalisation ; World Standards Cooperation.

56 On trouvera à la figure 7 un aperçu de certaines sources d'initiatives de gouvernance de l'IA et de normes sectorielles, classées en fonction de la portée géographique et du caractère inclusif (en dressant la liste de ces travaux récents, nous saluons les efforts déployés depuis de nombreuses années par le milieu universitaire, la société civile et les organismes professionnels).

57 Parmi les initiatives plurilatérales régionales et interrégionales pertinentes, on peut citer celles menées par l'Union africaine, divers hôtes de sommets sur l'IA, l'Association des nations de l'Asie du Sud-Est, le Conseil de l'Europe, l'Union européenne, le Groupe des Sept (G7), le Groupe des 20 (G20), le PMIA, l'Organisation des États américains et l'OCDE, entre autres.

58 Notre analyse des dispositions actuelles en matière de gouvernance sera probablement dépassée en l'espace de quelques mois. Toutefois, elle peut aider à illustrer la manière dont les initiatives internationales actuelles et nouvelles en matière de gouvernance de l'IA sont liées à nos principes directeurs pour la formation de nouveaux organismes mondiaux de gouvernance de l'IA, notamment le principe directeur 1 (l'IA doit être gouvernée de manière inclusive, par toutes et par tous et dans l'intérêt commun).

3. Lacunes dans la gouvernance mondiale de l'IA

- 59** Les multiples initiatives, notamment nationales, régionales et multipartites, susmentionnées ont permis d'obtenir des résultats significatifs et d'éclairer notre travail ; nombre de leurs représentants ont contribué à nos délibérations par écrit ou participé à nos consultations.
- 60** Pourtant, hormis quelques initiatives de l'ONU¹², aucune n'a une portée véritablement mondiale. Ces lacunes dans la représentation de la gouvernance de l'IA au niveau international constituent un problème car la technologie est mondiale et aura des effets à l'échelle mondiale.
- 61** Le manque de coordination entre les initiatives et les organismes risque de diviser le monde en régimes de gouvernance de l'IA déconnectés et incompatibles.
- 62** En outre, les lacunes en matière de mise en œuvre et de responsabilité réduisent la capacité des États, du secteur privé, de la société civile, du milieu universitaire et de la communauté technique à traduire les engagements, aussi représentatifs soient-ils, en résultats concrets.
- 64** Beaucoup excluent des régions entières du monde. Comme le montre la figure 8, si l'on considère 7 initiatives d'IA plurilatérales et interrégionales ne relevant pas de l'ONU et dont les membres se chevauchent, 7 pays sont parties à toutes ces initiatives, tandis que 118 pays ne sont parties à aucune (essentiellement dans le monde du Sud, avec une représentation inégale, même parmi les pays à la pointe de l'IA (voir fig. 12).
- 65** La sélectivité est compréhensible à un stade précoce de la gouvernance, lorsqu'il y a un certain degré d'expérimentation et de concurrence autour des normes et divers niveaux d'aisance avec les nouvelles technologies. Toutefois, à mesure que la gouvernance internationale de l'IA gagne en maturité, la représentation mondiale devient plus importante en termes d'équité et d'efficacité.
- 66** Outre le caractère non inclusif des efforts existants, il existe également des lacunes en matière de représentation dans les initiatives nationales et régionales visant à parvenir à une vision scientifique commune de l'IA. Ces lacunes en matière de représentation peuvent se manifester dans les processus de prise de décision concernant l'étendue, les ressources et la réalisation des évaluations.

A. Lacunes en matière de représentation

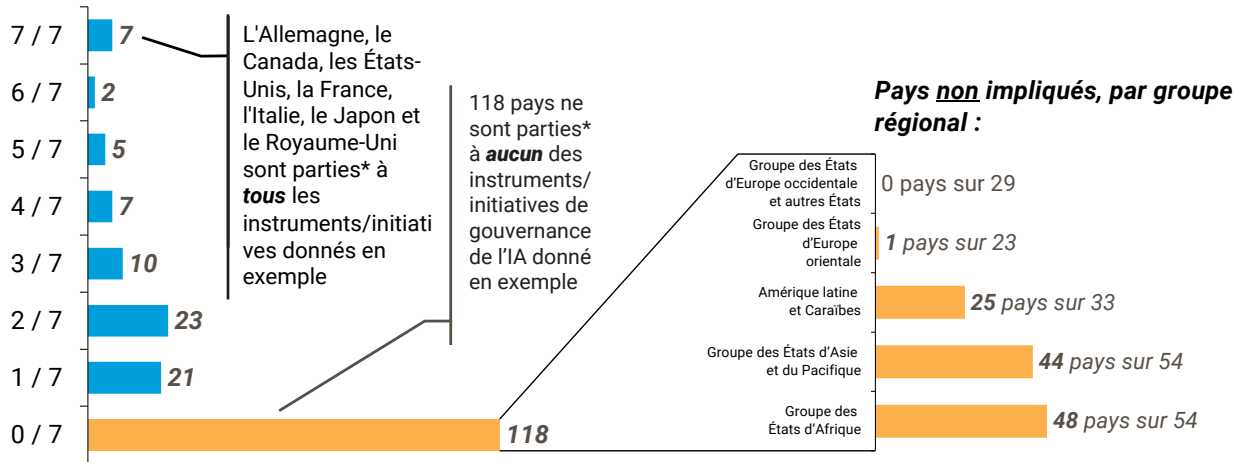
- 63** Notre analyse des diverses initiatives de gouvernance de l'IA qui ne relèvent pas de l'ONU et qui s'étendent sur plusieurs régions montre que la plupart d'entre elles ne sont pas pleinement représentatives dans leurs dimensions intergouvernementales.
- 67** L'équité exige qu'un plus grand nombre de voix jouent un rôle majeur dans les décisions concernant la manière de gouverner les technologies qui nous concernent tous, et qu'on reconnaisse que de nombreuses populations ont été exclues de ces débats au cours de l'histoire. L'absence relative

12 Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle (2021) et deux résolutions de l'Assemblée générale sur l'IA.

Figure 8 : Représentation dans sept initiatives internationales de gouvernance de l'IA ne relevant pas de l'ONU

Exemple : Principes de l'OCDE en matière d'IA (2019), principes du G20 en matière d'IA (2019), groupe de rédaction de la convention du Conseil de l'Europe sur l'IA (2022-2024), Déclaration ministérielle du PMIA (2022), Déclaration des Ministres du G7 (2023), Déclaration de Bletchley (2023) et Déclaration ministérielle de Séoul (2024).

**Interrégional
uniquement, excluant le
niveau régional**



*Par approbation des textes intergouvernementaux pertinents. Les pays ne sont pas considérés comme participant à une initiative plurilatérale du seul fait de leur appartenance à l'Union européenne ou à l'Union africaine. *Abréviations* : Groupe des États d'Afrique ; Groupe des États d'Asie et du Pacifique ; Groupe des États d'Europe orientale ; G20, Groupe des 20 ; G7, Groupe des Sept ; PMIA, Partenariat mondial sur l'intelligence artificielle ; Amérique latine et Caraïbes ; OCDE, Organisation de coopération et de développement économiques ; Groupe des États d'Europe occidentale et autres États.

de thèmes à l'ordre du jour des grandes initiatives prioritaires pour certaines régions est le signe d'un déséquilibre dû à une sous-représentation¹³.

68 Les régimes de gouvernance de l'IA doivent s'étendre à l'ensemble du globe pour être efficaces, c'est-à-dire pour instaurer la confiance, éviter les « courses aux armements de l'IA » ou le nivellement par le bas en matière de sûreté et de droits, traiter efficacement les difficultés liées au caractère transfrontalier de l'IA, stimuler l'apprentissage, encourager l'interopérabilité et mettre en commun les avantages de l'IA¹⁴. En outre, le fait d'inclure divers points de vue, y compris de sensibilités différentes, permet d'anticiper les menaces et de calibrer des mesures créatives et adaptables.

69 En limitant l'éventail des pays inclus dans les processus clés d'élaboration du programme, d'établissement de relations et d'échange d'informations, le plurilatéralisme sélectif peut limiter la réalisation de ses propres objectifs. Il s'agit notamment de la compatibilité des approches émergentes en matière de gouvernance de l'IA, de la sûreté mondiale de l'IA et de la vision commune de la science de l'IA au niveau mondial (voir les recommandations 1, 2 et 3 sur ce qui rend une approche mondiale particulièrement efficace).

70 Avec l'adoption des deux résolutions sur l'IA par l'Assemblée en 2024¹⁵, les nations à la pointe de l'IA reconnaissent qu'il faut combler les lacunes en matière de représentation concernant la

13 Par exemple la gouvernance des ensembles de données d'entraînement de l'IA, l'accès à la puissance de calcul, le développement des capacités d'IA, les risques liés à l'IA concernant la discrimination des groupes marginalisés et l'utilisation de l'IA dans les conflits armés (voir à l'annexe E les résultats de l'enquête sur les risques liés à l'IA qui montre une perception différente des risques parmi les répondants du Groupe des États d'Europe occidentale et autres États d'une part et d'autres pays d'autre part). De nombreux États et groupes marginalisés ont également été exclus des avantages de l'IA ou risquent d'en subir les effets néfastes de manière disproportionnée. L'équité exige une approche diversifiée et inclusive qui tienne compte des points de vue de toutes les régions et répartisse les possibilités offertes de manière égale tout en atténuant les risques.

14 Si et quand des lignes rouges sont établies - par analogie, peut-être, avec l'interdiction du clonage humain - elles ne seront applicables que si la norme est acceptée au niveau mondial et si son respect est contrôlé. Cela reste vrai même si, paradoxalement, dans le paradigme actuel, alors que les coûts d'un système d'IA donné diminuent, les coûts des systèmes d'IA avancés (sans doute les plus importants à contrôler) augmentent.

15 Résolutions 78/265 (Saisir les possibilités offertes par des systèmes d'intelligence artificielle sûrs, sécurisés et dignes de confiance pour le développement durable) et 78/311 (Intensifier la coopération internationale en matière de renforcement des capacités dans le domaine de l'intelligence artificielle).

gouvernance internationale de l'IA, et l'ONU pourrait être l'enceinte permettant de rassembler le monde à cet égard.

71 Le Pacte numérique mondial et le Forum mondial sur la société de l'information, qui se tiendront en septembre 2024 et en 2025, fournissent deux enceintes stratégiques supplémentaires où un ensemble de processus de gouvernance de l'IA représentatif au niveau mondial pourrait être institutionnalisé afin de combler les lacunes en matière de représentation¹⁶.

B. Lacunes en matière de coordination

72 L'émergence et l'évolution constantes des initiatives de gouvernance de l'IA ne garantissent pas une collaboration efficace au bénéfice de l'humanité. Au demeurant, des lacunes en matière de coordination sont apparues. Une concertation efficace entre les initiatives plurilatérales sélectives (voir fig. 8) et d'autres initiatives régionales n'est pas assurée, ce qui risque d'entraîner une incompatibilité entre les régions.

73 À l'échelle mondiale, il n'existe pas non plus de dispositifs permettant à toutes les organisations internationales de normalisation (voir fig. 7), aux initiatives internationales de recherche scientifique ou aux initiatives de renforcement des capacités en matière d'IA de se coordonner, ce qui compromet l'interopérabilité des approches et entraîne une

fragmentation. Le manque de coordination qui en résulte entre les différentes initiatives sous-mondiales est dans certains cas mieux traité au niveau mondial.

74 On observe également un manque de coordination au sein du système des Nations Unies, comme en témoigne le large éventail de documents et d'initiatives menées en matière d'IA. La figure 9 présente 27 instruments de l'ONU dans des domaines spécifiques susceptibles de s'appliquer à l'IA, dont 23 sont contraignants et nécessiteront une interprétation dans la mesure où ils concernent l'IA. En outre, 29 documents des Nations Unies et d'organisations apparentées portent sur l'IA, mais aucun d'entre eux n'est contraignant¹⁷. Dans certains cas, ils peuvent traiter les risques liés à l'IA et exploiter les avantages de l'IA dans des domaines spécifiques.

75 Le niveau d'activité montre l'importance de l'IA pour les programmes des Nations Unies. À mesure que l'IA se développe et touche des segments de plus en plus larges de la société, des appels toujours plus nombreux seront lancés à diverses parties du système des Nations Unies pour qu'elles agissent, y compris au moyen de normes contraignantes. Cela montre également le caractère ad hoc des réponses, qui se sont largement développées de manière organique dans des domaines précis et sans stratégie globale. Le manque de coordination qui en résulte favorise les chevauchements et entrave l'interopérabilité et les résultats.

16 Diverses initiatives plurilatérales, notamment les principes de l'OCDE en matière d'IA, le processus d'Hiroshima du G7 sur l'IA et la Convention-cadre du Conseil de l'Europe sur l'IA, sont ouvertes à des soutiens ou à des adhérents autres que les pays qui en sont à l'origine. Une telle ouverture pourrait toutefois ne pas assurer la représentation et la légitimité à la vitesse et à l'échelle requises pour suivre le rythme accéléré de la prolifération de l'IA à l'échelle mondiale. Dans le même temps, les processus de gouvernance internationale de l'IA présentent toujours des lacunes en matière de représentation, la prise de décision étant concentrée entre les mains de quelques pays et entreprises.

17 Une enquête menée par le Conseil des chefs de secrétariat des organismes des Nations Unies en février 2024 auprès de 57 entités des Nations Unies a fait état de 50 documents concernant la gouvernance de l'IA ; 44 des 57 entités ont répondu, dont la Commission économique pour l'Amérique latine et les Caraïbes ; la Commission économique et sociale pour l'Asie et le Pacifique ; la Commission économique et sociale pour l'Asie occidentale ; la FAO ; l'Agence internationale de l'énergie atomique (AIEA) ; l'Organisation de l'aviation civile internationale (OACI) ; le Fonds international de développement agricole ; l'OIT ; le FMI ; l'Organisation internationale pour les migrations ; le Centre du commerce international ; l'UIT ; l'Entité des Nations Unies pour l'égalité des sexes et l'autonomisation des femmes ; le Programme commun des Nations Unies sur le VIH/sida ; la CNUCED ; le Département des affaires économiques et sociales ; le Département de la communication globale ; le Cabinet du Secrétaire général ; le Bureau de la coordination des affaires humanitaires ; le Haut-Commissariat des Nations Unies aux droits de l'homme ; le Bureau de lutte contre le terrorisme ; le Bureau des affaires de désarmement ; le Bureau de l'informatique et des communications ; le Bureau de l'Envoyé du Secrétaire général pour les technologies ; le PNUD ; le Bureau des Nations Unies pour la prévention des catastrophes ; le Programme des Nations Unies pour l'environnement ; l'UNESCO ; la Convention-cadre des Nations Unies sur les changements climatiques ; le Fonds des Nations Unies pour la population ; le Haut-Commissariat des Nations Unies pour les réfugiés ; le UNICEF ; l'Institut interrégional de recherche des Nations Unies sur la criminalité et la justice ; l'Organisation des Nations Unies pour le développement industriel (ONUDI) ; l'Office des Nations Unies contre la drogue et le crime/Office des Nations Unies à Vienne ; le Bureau des Nations Unies pour les services d'appui aux projets ; l'Office de secours et de travaux des Nations Unies pour les réfugiés de Palestine dans le Proche-Orient ; l'ONU ; le Programme des Volontaires des Nations Unies ; l'Organisation mondiale du commerce (OMC) ; l'Union postale universelle ; le Groupe de la Banque Mondiale ; le PAM ; l'OMS ; l'OMPI. Voir le livre blanc du système des Nations Unies sur la gouvernance de l'IA : une analyse des modèles institutionnels et des fonctions du système des Nations Unies, ainsi que des cadres normatifs internationaux existants applicables à la gouvernance de l'IA (disponible à l'adresse suivante : <https://unsceb.org/united-nations-system-white-paper-ai-governance>).

Figure 9 : Documents relatifs à la gouvernance de l'IA émanant de l'ONU et d'organisations apparentées

Non exhaustif					
<p>Éthique et stratégie</p> <p>UNESCO</p> <ul style="list-style-type: none"> Recommandation sur l'éthique de l'intelligence artificielle <p>OMS</p> <ul style="list-style-type: none"> Guidance on Ethics & Governance of Artificial Intelligence for Health <p>Fonds des Nations Unies pour l'enfance (UNICEF)</p> <ul style="list-style-type: none"> Orientations stratégiques sur l'IA destinée aux enfants The Case for Better Governance of Children's Data: A Manifesto Responsible Data for Children (rd4c.org) <p>Programme des Nations Unies pour les établissements humains (ONU-HABITAT)</p> <ul style="list-style-type: none"> Guide to mainstream human rights in the digital transformation of cities Policy framework for centring people, inclusion, and human rights in smart city development <p>ONU-Femmes</p> <ul style="list-style-type: none"> Soixante-septième session de la Commission de la condition de la femme (« conclusions concertées ») <p>Fonds des Nations Unies pour la population</p> <ul style="list-style-type: none"> Programme of action for the international Conference on Population and Development: A Population-focused human rights-based framework <p>ESSENTIEL: S'applique à l'IA Peut s'appliquer à l'IA * Force obligatoire</p>	<p>Droits humains</p> <p>ACNUDH</p> <ul style="list-style-type: none"> Haut-Commissariat des Nations Unies aux droits de l'homme Convention internationale sur l'élimination de toutes les formes de discrimination raciale* Pacte international relatif aux droits civils et politiques* Pacte international relatif aux droits économiques, sociaux et culturels* Convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes* Convention contre la torture et autres peines ou traitements cruels, inhumains ou dégradants* <p>OIT</p> <ul style="list-style-type: none"> Convention (no 182) sur les pires formes de travail des enfants, 1999* Convention (no 155) sur la sécurité et la santé des travailleurs, 1981 Convention (n° 187) sur le cadre promotionnel pour la sécurité et la santé au travail, 2006* Convention (no 111) concernant la discrimination (emploi et profession), 1958* Convention (n° 135) concernant les représentants des travailleurs, 1971* Convention (no 122) sur la politique de l'emploi, 1964* Recueil de directives pratiques du BIT sur la protection des données personnelles des travailleurs <p>UNICEF</p> <ul style="list-style-type: none"> Convention relative aux droits de l'enfant* 	<p>Normes techniques</p> <p>UIT</p> <ul style="list-style-type: none"> LTA dans l'exploitation et la gestion des télécommunications LTA dans les systèmes et villes intelligents LTA dans la gestion des réseaux et les services, l'IA dans des technologies ou applications ciblées <p>PNUD</p> <ul style="list-style-type: none"> La norme sur les biens publics numériques pour les systèmes d'intelligence artificielle (élaborée en collaboration avec Digital Public Goods Alliance) <p>OACI</p> <ul style="list-style-type: none"> Annexes de la Convention de Chicago* <p>Paix et sécurité</p> <p>Bureau des affaires de désarmement</p> <ul style="list-style-type: none"> Article 36 du Protocole additionnel I aux Conventions de Genève* Convention sur les armes biologiques* Convention sur les armes chimiques* <p>Santé</p> <p>OMS</p> <ul style="list-style-type: none"> Regulatory considerations on artificial intelligence for health Generating Evidence for Artificial Intelligence Based Medical Devices: A Framework for Training Validation and Evaluation Guidance on Ethics & Governance of Artificial Intelligence for Health 	<p>Communications</p> <p>Département de la communication globale</p> <ul style="list-style-type: none"> Elaboration de principes sur l'intégrité de l'information <p>UNESCO</p> <ul style="list-style-type: none"> Principes pour la gouvernance des plateformes numériques <p>Commerce</p> <p>OMC</p> <ul style="list-style-type: none"> Accord général sur le commerce des services* Obstacles techniques au commerce* Accord sur les technologies de l'information* Aspects des droits de propriété intellectuelle qui touchent au commerce* Accord sur la facilitation des échanges <p>CNUDCI</p> <ul style="list-style-type: none"> Projet de dispositions sur les contrats automatisés <p>Propriété intellectuelle</p> <p>OMPI</p> <ul style="list-style-type: none"> Convention de Rome sur la protection des artistes interprètes ou exécutants, des producteurs de phonogrammes et des organismes de radiodiffusion* Convention de Berne pour la protection des œuvres littéraires et artistiques* Traité de Beijing sur les interprétations et exécutions audiovisuelles* Traité de coopération en matière de brevets* 	<p>Drogues et criminalité</p> <p>Office des Nations Unies contre la drogue et le crime</p> <ul style="list-style-type: none"> Déclaration de Kyoto <p>UNICRI</p> <ul style="list-style-type: none"> Policy Framework for Responsible Limits on Facial Recognition. Use Case: Law Enforcement Investigations Toolkit for Responsible Artificial Intelligence Innovation in Law Enforcement <p>Bureau de lutte contre le terrorisme</p> <ul style="list-style-type: none"> Huitième examen de la Stratégie antiterroriste mondiale des Nations Unies (A/RES/77/298) <p>Éducation</p> <p>UNESCO</p> <ul style="list-style-type: none"> Orientations pour l'IA générative dans l'éducation et la recherche Référentiels de compétences en IA pour les élèves et les enseignants Compétences en IA et transformation numérique pour les fonctionnaires <p>Autres</p> <p>ONU-Habitat</p> <ul style="list-style-type: none"> AI Risk Assessment Framework Directives internationales pour des villes intelligentes centrées sur les personnes <p>Organisation des Nations Unies pour le développement industriel</p> <ul style="list-style-type: none"> Déclaration d'Abou Dhabi, Dix-huitième Conférence générale de l'ONUDI <p>Bureau des Nations Unies pour la prévention des catastrophes</p> <ul style="list-style-type: none"> Cadre de Sendai pour la réduction des risques de catastrophe* 	

Source : livre blanc du système des Nations Unies sur la gouvernance de l'IA : une analyse des modèles institutionnels et des fonctions du système des Nations Unies, ainsi que des cadres normatifs internationaux existants applicables à la gouvernance de l'IA, 28 février 2024.

76 Le nombre et la diversité des approches montrent que le système des Nations Unies répond à une question émergente. Avec une stratégie adaptée et des processus intégrés, ces travaux peuvent ouvrir une voie efficace et durable vers une gouvernance internationale inclusive de l'IA dans des domaines précis. Cela permettrait d'avoir des effets majeurs, harmonisés et coordonnés dans des domaines tels que la santé, l'éducation, les normes techniques et l'éthique, au lieu de simplement contribuer à multiplier les initiatives et les organismes dans ce domaine en pleine expansion. Le droit international, y compris le droit international des droits de l'homme, constitue une base normative commune pour tous les travaux liés à l'IA, facilitant ainsi la coordination et la cohérence.

77 Bien que de nombreuses entités des Nations Unies abordent la gouvernance de l'IA, leurs mandats spécialisés ne leur permettent pas de la traiter dans sa globalité, pas plus que les points focaux gouvernementaux désignés, qui sont tout aussi spécialisés¹⁸. Cela limite la capacité des entités existantes à traiter seules les implications pluridimensionnelles de l'IA à l'échelle mondiale. Aux niveaux national et régional, de nouvelles institutions¹⁹, telles que les organismes de sûreté de l'IA ou les bureaux de l'IA, s'emploient à combler ces lacunes dans le cadre d'une approche transversale adaptée.

18 Par exemple les ministères de l'éducation, de la science et de la culture (UNESCO) ; des télécommunications ou des TIC (UIT) ; de l'industrie (ONUDI) ; et du travail (OIT).

19 Notamment ceux mis en place par le Canada, les États-Unis, le Japon, la République de Corée, le Royaume-Uni, Singapour et l'Union européenne.

C. Lacunes en matière d'exécution

78 Toutefois, la représentation et la coordination ne suffisent pas. Il faut établir des processus d'action et de suivi afin de garantir que les engagements en faveur de la bonne gouvernance se traduisent par des résultats concrets dans la pratique.

Il faut faire plus pour garantir l'application du principe de responsabilité. L'influence des pairs et l'apprentissage entre pairs sont deux éléments qui peuvent encourager l'application de ce principe.

79 La collaboration avec le secteur privé sera tout aussi importante pour garantir une responsabilité effective et réparer les préjudices. L'ONU en a fait l'expérience dans ses Principes directeurs relatifs aux entreprises et aux droits de l'homme. De même, nous aurions besoin d'une participation forte de la société civile et des experts scientifiques pour que les gouvernements et les entreprises privées restent honnêtes quant à leurs engagements et à leurs affirmations.

80 L'absence d'outils permettant d'exploiter les avantages de l'IA dans l'intérêt général au sein des pays et entre les pays constitue une lacune importante dans la mise en œuvre. De nombreux pays ont mis en place des stratégies nationales pour stimuler les infrastructures et les talents liés à l'IA et quelques initiatives d'assistance internationale voient le jour²⁰. Toutefois, ces derniers ne disposent pas de réseaux ni de ressources suffisantes.

81 Au niveau mondial, la coordination des initiatives nationales et régionales de développement des capacités et la mutualisation des ressources en vue de soutenir les pays exclus de ces efforts peuvent contribuer à faire en sorte qu'aucun pays ne soit laissé pour compte dans l'accès aux possibilités offertes par l'IA. Une autre lacune importante dans la mise en œuvre est l'absence d'un fonds dédié au renforcement des capacités en matière d'IA, malgré l'existence de certains dispositifs de financement pour les capacités numériques (encadré 8).

²⁰ Les initiatives nationales pourraient continuer à utiliser des outils de diagnostic tels que la méthode d'évaluation de l'état de préparation à l'IA élaborée par l'UNESCO afin de recenser les lacunes au niveau des pays, le réseau international aidant à les combler.

Encadré 8 : Insuffisance du financement mondial des capacités d'IA

L'Organe consultatif estime qu'il n'existe pas de fonds mondial pour le renforcement des capacités en matière d'IA doté de la portée et du mandat nécessaires pour financer les investissements importants requis pour combler le fossé en matière d'IA.

Les estimations indicatives situent le montant nécessaire dans une fourchette de 350 millions à 1 milliard de dollars par an^a, ce qui comprend les contributions en nature du secteur privé, avec pour mandat de cibler les capacités d'IA sur l'ensemble des facilitateurs de l'IA, notamment les talents, la puissance de calcul, les données d'entraînement, la mise au point de modèles et la collaboration interdisciplinaire pour les applications.

Les dispositifs de financement multilatéraux comprennent les exemples suivants :

a) Fonds commun pour les objectifs de développement durable :

Grâce à sa large portée, le Fonds englobe tous les objectifs de développement durable ainsi que l'intervention d'urgence. Il soutient les initiatives au niveau des pays en faveur d'une politique intégrée de l'ONU et d'un appui financier stratégique aux pays afin de faire progresser la réalisation des objectifs de développement durable. Il aide l'ONU à fournir et à catalyser le financement et la programmation des objectifs de développement durable. Depuis 2017, 30 entités des Nations Unies participantes ont reçu un montant total de 223 millions de dollars. Le Fonds ne finance pas directement les gouvernements nationaux, les populations ou les entités, ni ne finance les initiatives transfrontières.

En 2023, le Fonds comptait environ 16 donateurs pour un total de 57,7 millions de dollars et un montant estimé à 58,8 millions de dollars en 2024. Le secteur privé a fourni un montant de 83 155 de dollars depuis 2017, et aucune contribution en 2023 ou 2024 à ce jour.

La plupart du Fonds (60 %) est consacrée à des initiatives relevant des cinq objectifs de développement durable suivants : les objectifs 2 (faim zéro), 5 (égalité des genres), 7 (énergie abordable et propre), 9 (industrie, innovation et infrastructure) et 17 (partenariats).

L'UIT et le PNUD ont financé à parts égales un projet d'un montant de 250 000 dollars au titre du volet Transformation numérique des politiques du Fonds [lancé en 2023]. À la fin de l'exercice 2023, son taux d'exécution était de 2,27 %. Les activités de transformation numérique ne représentent qu'une petite partie des activités du Fonds et sont généralement liées à d'autres objectifs de développement durable (par exemple la connectivité et l'infrastructure numérique pour soutenir la prestation de services, tel que dans les petits États insulaires en développement).

b) Partenariat numérique pour le développement de la Banque mondiale

Ce fonds aide les pays à élaborer et à mettre en œuvre la transformation numérique en mettant l'accent sur l'infrastructure à haut débit, l'accès et l'utilisation, l'infrastructure publique numérique et la production, l'accessibilité et l'utilisation des données. Fin 2022, il avait investi un montant de 10,7 milliards de dollars dans plus de 80 pays.

Le partenariat comprend un fonds d'affectation spéciale multidonateurs associé à la cybersécurité (Estonie, Allemagne, Japon et le Royaume des Pays-Bas) qui vise à soutenir le développement des capacités nationales en la matière.

a Soit moins de 1 % des investissements annuels estimés du secteur privé dans l'IA en 2023.

4. Renforcement de la coopération mondiale

82 Après avoir exposé le manque de gouvernance mondiale, nous examinons maintenant les recommandations qui visent à combler les lacunes prioritaires à court terme.

83 Nos recommandations mettent en avant une vision globale qui favorise une gouvernance de l'IA souple, agile et interconnectée à l'échelle mondiale au bénéfice de l'humanité et qui repose sur une compréhension commune, un socle commun et des avantages communs afin d'améliorer la représentation, de permettre la coordination et de renforcer la mise en œuvre (voir fig. 10). Seule une approche globale et inclusive de la gouvernance de l'IA peut permettre de traiter les défis et les

avantages multiples et évolutifs que l'IA présente à l'échelle mondiale et de promouvoir ainsi la stabilité internationale et le développement équitable.

84 Guidées par les principes énoncés dans notre rapport d'étape (voir par. 47), nos propositions visent à combler les lacunes et à apporter de la cohérence à l'écosystème de mesure et d'initiatives internationales qui progresse rapidement en matière de gouvernance de l'IA, afin d'éviter la fragmentation et les occasions manquées. Afin de soutenir efficacement ces mesures et d'établir des partenariats utiles avec d'autres organismes, nous proposons une structure légère et agile dans le cadre d'une stratégie cohérente : un bureau de

Figure 10 : Vue d'ensemble des recommandations et de la manière dont elles comblent le manque de gouvernance de l'IA au niveau mondial

Objectif	Améliorer la représentation	Permettre la coordination	Renforcer la mise en œuvre
Vision commune Groupe scientifique international sur l'IA	✓	✓	
Socle commun Concertation sur les mesures à prendre en matière de gouvernance de l'IA Échange de normes sur l'IA	✓	✓	(✓)
Avantages communs Réseau de développement des capacités Fonds mondial pour l'IA Cadre mondial de données sur l'IA	✓	✓	✓
Stratégie cohérente Bureau de l'IA au Secrétariat	Conseiller le Secrétaire général sur les questions liées à l'IA, assurer la cohérence du système des Nations Unies, mobiliser les États et les parties prenantes, établir des partenariats et des interfaces avec d'autres processus et organismes, et soutenir d'autres propositions le cas échéant.		

l'IA au Secrétariat de l'ONU, proche du Secrétaire général, qui servirait de « ciment » pour fédérer les initiatives proposées ici de manière efficace et durable.

85 L'ONU est loin d'être parfaite. Toutefois, la légitimité découlant de son caractère unique et inclusif, associée aux fondements normatifs contraignants du droit international, notamment du droit international des droits de l'homme, permet d'espérer une gouvernance de l'IA qui permette de préserver les intérêts et la vie de l'humanité, d'une manière équitable, efficace et efficiente ²¹.

A. Vision commune

86 Une approche globale de la gouvernance de l'IA commence par une vision commune de ses capacités, possibilités, risques et incertitudes.

87 Le domaine de l'IA évolue rapidement, ce qui produit une quantité considérable d'informations et entrave le décryptage entre le battage médiatique et la réalité. Cela peut prêter à confusion, empêcher l'établissement d'une vision commune et avantager les grandes entreprises d'IA au détriment des décideurs, de la société civile et du public.

88 En outre, le manque de collaboration scientifique internationale et d'échange d'informations peut engendrer des perceptions erronées à l'échelle mondiale et saper la confiance internationale.

89 Il convient de disposer en temps utile de connaissances et d'informations scientifiques impartiales et fiables sur l'IA afin que les États Membres puissent élaborer une vision fondamentale commune à l'échelle mondiale, et d'équilibrer les asymétries d'information entre les entreprises hébergeant des laboratoires d'IA coûteux et le reste du monde, notamment par l'échange d'informations entre les entreprises d'IA et l'ensemble de la communauté de l'IA.

90 Cela est particulièrement efficace au niveau mondial pour investir conjointement dans un bien public mondial et collaborer, dans l'intérêt général, à des efforts qui, autrement, seraient fragmentés et feraient double emploi.

Groupe scientifique international sur l'IA

Recommandation 1 : groupe scientifique international sur l'intelligence artificielle

Nous recommandons la création d'un groupe scientifique international indépendant sur l'IA, composé de divers experts multidisciplinaires du domaine, participant à titre personnel et sur la base du volontariat. Soutenu par le bureau de l'IA dont la création est proposée au Secrétariat et par d'autres organismes des Nations Unies concernés, en partenariat avec d'autres organisations internationales compétentes, le mandat du groupe comprendrait les éléments suivants :

- a. La publication d'un rapport annuel examinant les capacités, les possibilités, les risques et les incertitudes liés à l'IA ainsi que les domaines de consensus scientifique sur les tendances technologiques et les domaines dans lesquels des travaux de recherche supplémentaires doivent être menés ;
- b. La production de dossiers de recherche thématiques trimestriels sur les domaines dans lesquels l'IA pourrait contribuer à la réalisation des objectifs de développement durable, qui seraient axés sur les domaines d'intérêt général susceptibles d'être négligés ;
- c. La publication de rapports ad hoc sur de nouvelles questions, en particulier l'apparition de nouveaux risques ou de lacunes majeures dans le paysage de la gouvernance.

91 Il existe des précédents pour ce type d'organisme, par exemple l'UNSCEAR, la Plateforme intergouvernementale scientifique et politique sur la biodiversité et les services écosystémiques, le Comité scientifique pour les recherches antarctiques et le GIEC.

²¹ Elle devrait également être inclusive, renforcer la cohésion et consolider la paix et la sécurité dans le monde.

- 92** Ces modèles sont connus pour leurs approches systématiques de questions complexes et omniprésentes touchant divers secteurs et populations mondiales. Cependant, bien qu'ils puissent servir d'exemples à suivre, aucun n'est parfaitement adapté à l'évaluation de la technologie de l'IA et ne devrait être reproduit directement. Au contraire, il faut établir une approche sur mesure.
- 93** En s'inspirant de ces précédents, un groupe scientifique indépendant, international et multidisciplinaire sur l'IA pourrait rassembler et catalyser des travaux de recherche de pointe au service de ceux qui recherchent des éclairages scientifiques sur la technologie de l'IA ou ses applications à partir d'une source impartiale et crédible. Le débat en cours sur les systèmes d'IA ouverts ou fermés, traité à l'encadré 9, est un exemple de question à laquelle le groupe pourrait contribuer.
- 94** Un groupe scientifique placé sous les auspices de l'ONU aurait un champ d'action large afin de couvrir de manière globale un éventail complet de priorités. Il peut s'agir notamment d'apporter un éclairage technique sur les possibilités liées à l'IA, de faciliter les études approfondies dans les domaines appliqués des objectifs de développement durable, tels que les soins de santé, l'énergie, l'éducation, la finance, l'agriculture, le climat, le commerce et l'emploi.
- 95** Les estimations des risques pourraient également s'appuyer sur les travaux d'autres initiatives de recherche sur l'IA, l'ONU fournissant un espace de confiance unique pour que les chercheurs puissent échanger des idées sur les dernières avancées en la matière. Le droit international, notamment les droits humains, servirait de boussole pour définir les risques pertinents. En mettant en commun les connaissances au-delà des silos au sein de pays ou d'entreprises qui pourraient autrement ne pas participer ou être mobilisés, un groupe créé sous l'égide de l'ONU peut contribuer à rectifier les perceptions erronées et à renforcer la confiance à l'échelle mondiale.
- 96** Un groupe scientifique de ce type ne mènerait pas nécessairement ses propres travaux de recherche mais servirait de catalyseur pour une action en réseau²². Il pourrait compiler, distiller et traduire les avancées en matière d'IA à l'intention de son public, en mettant en évidence les cas d'utilisation possibles. Il permettrait de réduire l'asymétrie de l'information, d'éviter les investissements mal orientés et d'assurer la circulation de l'information au sein d'un réseau mondial d'experts.
- 97** Le groupe s'adresserait principalement à trois publics :
- Le premier est la communauté scientifique mondiale²³. Le transfert de la recherche fondamentale sur l'IA vers des sociétés privées, motivé en partie par le coût de la puissance de calcul, a fait craindre que ces travaux de recherche puissent servir des intérêts financiers. Un groupe scientifique pourrait encourager les organismes publics du monde entier à axer davantage les travaux de recherches sur le bien public.
 - Deuxièmement, des évaluations indépendantes régulières seraient mises au service des États Membres, des décideurs et d'autres dispositifs recommandés dans le présent rapport. Une enquête annuelle sur les risques, menée par des experts mondiaux, contribuerait à définir la teneur des concertations sur la gouvernance de l'IA proposées dans la recommandation 2. Le rapport sur l'état des connaissances servirait de base à l'élaboration des normes proposées dans la recommandation 3, ainsi qu'à la mise en place du réseau de développement des capacités proposé dans la recommandation 4.
 - Troisièmement, grâce à ses rapports publics, le groupe pourrait servir de source impartiale d'informations de qualité pour le public.

²² Il pourrait notamment s'appuyer sur les groupes sectoriels ou régionaux existants.

²³ Il pourrait également mener des activités de sensibilisation auprès de publics plus larges, notamment la société civile et le grand public.

- 98** La portée mondiale des réseaux accessibles uniquement par l'intermédiaire de l'ONU permettrait une vision commune sur la base la plus large possible, et mettrait à disposition les résultats sous des formes adaptées aux différents contextes socioéconomiques et géographiques. Le groupe peut ainsi faire de l'ONU une plateforme fiable pour une vision multidisciplinaire interconnectée de toutes les parties prenantes.
- 99** Le groupe pourrait être créé pour une période initiale de 3 à 5 ans (avec une prorogation soumise à l'examen du Secrétaire général) et pourrait fonctionner sur la base suivante :
- a. Initialement, le groupe pourrait être constitué de 30 à 50 membres nommés par les États Membres et par eux-mêmes, comme pour la création de l'Organe consultatif. Il devrait se concentrer sur les connaissances scientifiques dans toutes les disciplines et devrait assurer une représentation diversifiée par région et par genre, tout en tenant compte du caractère interdisciplinaire de l'IA. Les membres pourraient être renouvelés à intervalles réguliers dans le cadre du mandat global de 3 à 5 ans.
 - b. Le groupe se réunirait en ligne (et en présentiel en séance plénière, peut-être à raison de deux fois par an par exemple). Les réunions pourraient se dérouler tour à tour dans les villes accueillant les entités des Nations Unies concernées, notamment dans le monde du Sud. Il conviendrait d'encourager la formation de groupes de travail thématiques, notamment d'ajouter des membres supplémentaires si nécessaire et de mobiliser des réseaux de partenaires universitaires. La participation d'entités compétentes des Nations Unies à ces groupes de travail pourrait être envisagée²⁴.
 - c. Le groupe fonctionnerait de manière indépendante, notamment en ce qui concerne ses constatations et ses conclusions, avec le soutien d'une équipe du système des Nations Unies composée de membres du bureau de l'IA proposé et des organismes compétents des Nations Unies, tels que l'UIT et l'UNESCO.
 - d. Le groupe devrait s'associer aux travaux de recherche menés par d'autres organismes internationaux tels que l'OCDE et le PMIA, ainsi qu'à d'autres travaux pertinents tels que le récent rapport scientifique sur les risques de l'IA avancée réalisé à la demande du Gouvernement du Royaume-Uni²⁵, et aux organisations régionales concernées, et s'appuyer sur ces travaux.
 - e. Un comité directeur élaborerait un programme de recherche qui tiendrait compte des différents points de vue et des considérations éthiques, superviserait l'allocation des ressources, encouragerait la collaboration avec un réseau d'institutions universitaires et d'autres parties prenantes, et passerait en revue les activités et les résultats du groupe.
- 100** En s'appuyant sur le pouvoir de rassemblement unique de l'ONU et la portée mondiale inclusive des groupes de parties prenantes, un groupe scientifique international peut fournir des processus et des résultats de collaboration scientifique fiables et corriger les asymétries d'information afin de combler les lacunes en matière de représentation et de coordination recensées aux paragraphes 66 et 73, promouvant ainsi une gouvernance internationale de l'IA équitable et efficace.

²⁴ Voir à la figure 9 la liste des entités des Nations Unies actives dans ce domaine.

²⁵ Rapport scientifique international sur la sécurité de l'IA avancée : Rapport d'étape, <https://gov.uk/government/publications/international-scientific-report-on-the-safety-of-advanced-ai>.

Encadré 9 : Systèmes d'IA ouverts ou fermés

Parmi les sujets traités lors de nos consultations figure le débat actuel sur les systèmes d'IA ouverts ou fermés. Les systèmes d'IA qui sont ouverts à des degrés divers sont souvent appelés « IA à code source ouvert », mais il s'agit d'une appellation quelque peu erronée si on la compare aux logiciels (codes) à code source ouvert. Il importe de souligner que l'ouverture des systèmes d'IA est davantage un spectre qu'un attribut unique.

Selon un article, « un système d'IA entièrement fermé n'est accessible qu'à un groupe particulier. Il peut s'agir d'une entreprise de développement de l'IA ou d'un groupe spécifique au sein de cette entreprise à des fins de recherche et de développement internes. En revanche, les systèmes plus ouverts peuvent autoriser un accès public ou mettre à disposition certaines parties, telles que les données, le code ou les caractéristiques du modèle afin de faciliter le développement externe de l'IA »^a.

Les systèmes d'IA à source ouverte dans le domaine de l'IA générative présentent à la fois des risques et des opportunités. Les entreprises invoquent souvent la « sûreté de l'IA » pour ne pas divulguer les spécifications des systèmes, ce qui traduit la tension permanente entre les approches ouvertes et fermées dans le secteur. Les débats tournent généralement autour de deux extrêmes : l'ouverture totale, qui implique la mise à disposition de tous les composants du modèle et des ensembles de données, et l'ouverture partielle, qui implique de ne divulguer que les poids du modèle.

Les systèmes d'IA à code source ouvert encouragent l'innovation et sont souvent une condition préalable à un financement public. À l'extrémité ouverte du spectre, lorsque le code sous-jacent est en accès libre, les développeurs du monde entier peuvent expérimenter, améliorer et créer de nouvelles applications. Cela favorise un environnement de collaboration où les idées et les connaissances sont facilement mises à disposition. Certaines grandes entreprises affirment que cette ouverture est essentielle à l'innovation et à la croissance économique.

Toutefois, dans la plupart des cas, les modèles d'IA en libre accès sont disponibles sous forme d'interfaces de programmation d'applications. Dans ce cas, le code original n'est pas mis à disposition, les poids originaux ne sont jamais modifiés et les mises à jour de modèles deviennent de nouveaux modèles.

En outre, les modèles en libre accès ont tendance à être plus petits et plus transparents. Cette transparence peut instaurer la confiance, permettre de traiter en amont les considérations éthiques et favoriser la validation et la réplication car les utilisateurs peuvent examiner le fonctionnement interne du système d'IA, comprendre son processus de prise de décision et recenser les biais potentiels.

Les systèmes d'IA fermés offrent un meilleur contrôle à leurs développeurs. En outre, les systèmes à source fermée peuvent être plus rationalisés et plus efficaces car la base de code n'évolue pas constamment grâce aux contributions du public. De nombreuses entreprises considèrent que l'ouverture totale n'est pas réalisable et que l'ouverture partielle est la seule solution possible. Toutefois, ce point de vue ne tient pas compte des possibilités offertes par une approche équilibrée permettant de réaliser une « ouverture significative »^b.

a Angela Luna, « The open or closed AI dilemma », 2 mai 2024. Disponible à l'adresse suivante : <https://bipartisanpolicy.org/blog/the-open-or-closed-ai-dilemma>.

b Inspiré par Andreas Liesenfeld et Mark Dingemans, « Rethinking open source generative AI : open-washing and the EU AI Act », *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)* (juin 2024).

Encadré 9 : Systèmes d'IA ouverts ou fermés (suite)

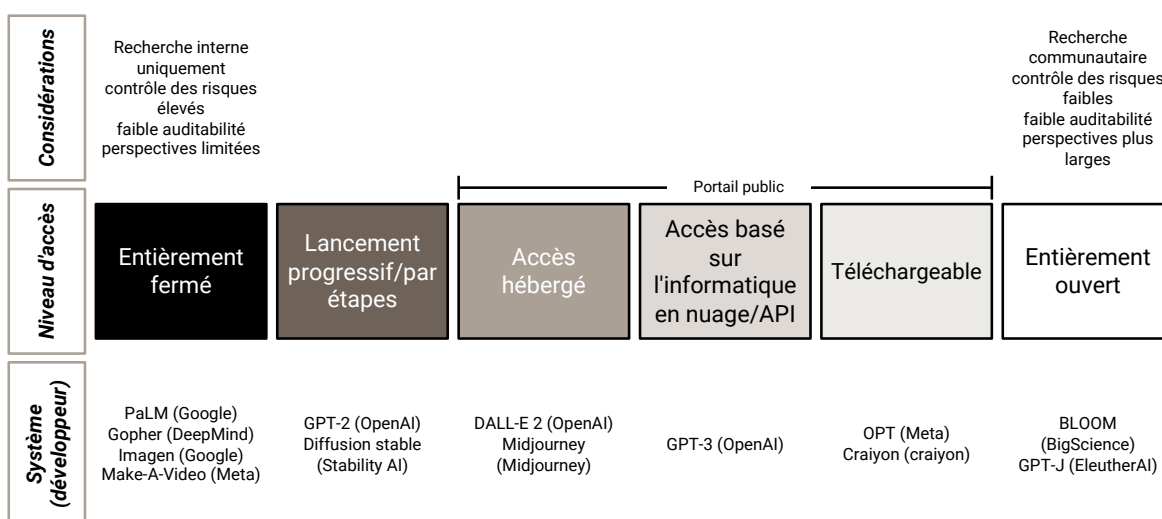
Une ouverture significative existe entre les deux extrêmes du spectre et peut être adaptée à différents cas d'utilisation. Cette méthode équilibrée favorise le développement d'une IA sûre, novatrice et inclusive qui permette un contrôle public et un audit indépendant des données divulguées en matière de formation et de mise au point. L'ouverture, qui ne se limite pas au simple partage des poids des modèles, peut favoriser l'innovation et l'inclusion car elle contribue aux applications dans la recherche et l'éducation.

La définition de l' « IA en accès libre » évolue^c et est souvent influencée par les intérêts des entreprises, comme illustré à la figure 11. Afin d'y remédier, nous recommandons de lancer un processus, coordonné par le groupe scientifique international proposé ci-dessus, afin d'élaborer une approche équilibrée et graduelle de l'ouverture. Cela permettrait d'adopter des approches significatives et fondées sur des données probantes en matière d'ouverture, afin d'aider les utilisateurs et les décideurs à faire des choix éclairés sur des modèles et des architectures d'IA.

La divulgation des données – même limitée à des éléments clés – est essentielle pour comprendre la performance des modèles, garantir la reproductibilité et évaluer les risques juridiques. La clarification des degrés d'ouverture peut contribuer à contrer le « open-washing » des entreprises et à favoriser un écosystème technologique transparent.

Il est également important, à mesure que la technologie mûrit, d'envisager les régimes de gouvernance pour l'application des systèmes d'IA ouverts et fermés. Nous devons élaborer des lignes directrices sur l'IA responsable, des normes contraignantes et des critères mesurables pour les développeurs et les concepteurs de produits et de services qui intègrent des technologies d'IA, ainsi que pour leurs utilisateurs et tous les acteurs impliqués tout au long de leur cycle de vie.

Figure 11 : Intérêts et ouverture des entreprises



Source : Irene Solaiman, « The gradient of generative AI release : methods and considerations », *Proceedings of the 2023 Association for Computing Machinery (ACM) Conference on Fairness, Accountability, and Transparency* (juin 2023), p. 111 à 122.

c The Open Source AI Definition – draft v. 0.0.3. Disponible à l'adresse suivante : <https://opensource.org/deepdive/drafts/the-open-source-ai-definition-draft-v-0-0-3>.

B. Socle commun

101 Outre la vision commune de l'IA, il convient de définir un socle commun afin d'établir des approches de gouvernance interopérables entre les pays et fondées sur des normes internationales, telles que la Déclaration universelle des droits de l'homme (voir principe 5 ci-dessus).

102 Cela est nécessaire au niveau mondial, non seulement pour assurer une représentation équitable, mais aussi pour éviter un nivellement par le bas de la réglementation et réduire les frictions réglementaires entre les pays, maximiser l'interopérabilité technique et ontologique, et enfin détecter les problèmes découlant des décisions prises tout au long du cycle de vie de l'IA et qui relèvent de plusieurs pays, et y répondre.

Concertation sur les mesures à prendre en matière de gouvernance de l'IA

Recommandation 2 : concertation sur les mesures à prendre en matière de gouvernance de l'IA

Nous recommandons le lancement d'une concertation intergouvernementale et multipartite sur la gouvernance de l'IA, deux fois par an, en marge des réunions existantes à l'ONU. Ses objectifs seraient les suivants :

- a. Échanger les meilleures pratiques en matière de gouvernance de l'IA qui encouragent le développement en même temps qu'elles favorisent le respect, la protection et la concrétisation des droits humains, notamment l'exploitation des possibilités et la gestion des risques ;
- b. Promouvoir une vision commune de la mise en œuvre des mesures de gouvernance de l'IA par les développeurs et les utilisateurs des secteurs privé et public afin d'améliorer l'interopérabilité internationale de la gouvernance de l'IA ;

- c. Faire part à titre volontaire des problèmes importants liés à l'IA qui ont mis à rude épreuve ou dépassé la capacité de réaction des organismes publiques ;
- d. Examiner les rapports du groupe scientifique international sur l'IA, selon qu'il convient.

103 La gouvernance internationale de l'IA est aujourd'hui une mosaïque fragmentée, dans le meilleur des cas. Actuellement, 118 pays ne participent à aucune des sept initiatives récentes et importantes de gouvernance de l'IA non liées à l'ONU et dotées d'un volet intergouvernemental²⁶ (voir fig. 8). Des lacunes liées à la représentation existent même parmi les 60 pays les plus performants en matière d'IA, ce qui met en évidence le caractère sélectif de la gouvernance internationale de l'IA aujourd'hui (voir fig. 12).

104 Il convient d'établir une instance stratégique inclusive afin que tous les États Membres, en s'appuyant sur le savoir-faire des parties prenantes, puissent mettre en commun les meilleures pratiques qui encouragent le développement en même temps qu'elles favorisent le respect, la protection et la concrétisation des droits humains, promeuvent des approches de gouvernance interopérables et surveillent les risques communs qui justifient de nouvelles mesures stratégiques.

105 Cela ne signifie pas une gouvernance mondiale de tous les aspects de l'IA (ce qui n'est ni possible ni souhaitable étant donné les intérêts et les priorités divergents des États). Pourtant, l'échange de vues sur les avancées de l'IA et les réponses stratégiques peuvent fixer le cadre de la coopération internationale.

²⁶ Ces initiatives ne sont pas toujours directement comparables. Certaines reflètent le travail d'organisations internationales ou régionales existantes, tandis que d'autres sont basées sur des invitations ad hoc de pays de même sensibilité.

Figure 12 : Les 60 premiers pays en matière d'IA (indice Tortoise 2023) dans l'échantillon des principales initiatives plurilatérales de gouvernance de l'IA avec des voies intergouvernementales

partie a

Pays/Partie*	Classement de l'indice mondial Tortoise sur l'IA (2023)	Nombre d'initiatives en tant que partie	Ordre chronologique						
			Principes de l'OCDE en matière d'IA (2019)	Principes du G20 en matière d'IA (2019)	Rédacteurs du Conseil de l'Europe (2022)	Déclaration ministérielle du PMIA (2022)	Déclaration de Bletchley (2023)	Déclaration ministérielle du G7 sur le processus d'Hiroshima en matière d'IA (2023)	Déclaration ministérielle de Séoul (2024)
États-Unis d'Amérique	1	7							
Chine	2	2							
Singapour	3	4							
Royaume-Uni de Grande-Bretagne et d'Irlande du Nord	4	7							
Canada	5	7							
République de Corée	6	5							
Israël	7	5							
Allemagne	8	7							
Suisse	9	4							
Finlande	10	2							
Royaume des Pays-Bas	11	5							
Japon	12	7							
France	13	7							
Inde	14	4							
Australie	15	6							
Danemark	16	3							
Suède	17	3							
Luxembourg	18	2							
Irlande	19	4							
Autriche	20	2							
Espagne	21	5							
Belgique	22	3							
Italie	23	7							
Norvège	24	2							
Estonie	25	2							
Émirats arabes unis	27	2							
Portugal	28	2							
Fédération de Russie	29	1							
Arabie saoudite	30	3							
Malte	31	2							
Brésil	33	4							
Nouvelle-Zélande	34	3							
Slovénie	35	3							
Hongrie	36	2							
Türkiye	37	6							
Islande	38	2							
Chili	39	3							
Qatar	40	0							
Lituanie	41	2							
Malaisie	42	0							
Grèce	43	2							
Indonésie	44	3							
Viet Nam	45	0							
Colombie	46	1							
Argentine	47	4							
Slovaquie	48	2							
Mexique	49	5							
Égypte	50	1							
Uruguay	51	1							
Arménie	52	1							
Afrique du Sud	53	1							
Tunisie	54	0							
Maroc	55	0							
Bahreïn	56	0							
Pakistan	57	0							
Sri Lanka	58	0							
Nigéria	59	2							
Kenya	60	2							
Union européenne	s.o.	5							
Total (y compris ceux qui ne sont pas indiqués)			47	20	58	29	29	7	28

*Including jurisdictions such as the Holy See and the European Union.

- Sources:
- OECD, Recommendation of the Council on Artificial Intelligence (adopted 21 May 2019), available at <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.
 - G20, AI Principles (June 2019), available at https://www.mofa.go.jp/policy/economy/g20_summit/osaka19/pdf/documents/en/annex_08.pdf.
 - GPAI, 2022 ministerial declaration (22 November 2022), available at [https://one.oecd.org/document/GPAI/G\(2022\)7/FINAL/en.pdf](https://one.oecd.org/document/GPAI/G(2022)7/FINAL/en.pdf).
 - Bletchley Declaration (1 Nov 2023), available at <https://gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>.
 - G7, Hiroshima AI Process G7 Digital & Tech Ministers' Statement (1 Dec 2023), available at https://www.soumu.go.jp/hiroshimaai/process/pdf/document02_en.pdf.
 - Council of Europe, Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (adopted 17 May 2024), available at <https://coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.
 - Seoul Ministerial Statement for advancing AI safety, innovation and inclusivity, AI Seoul Summit (22 May 2024).
 - Tortoise Media, Global AI Index (2023), available at <https://tortoisemedia.com/intelligence/global-ai/#rankings>.

106 L'ONU est particulièrement bien placée pour faciliter ces concertations de manière inclusive, afin d'aider les États Membres à collaborer efficacement. L'ensemble des normes existantes et nouvelles du système des Nations Unies peut fournir des fondements normatifs solides pour une action concertée, ancrés dans la Charte des Nations Unies, les droits humains et le droit international, notamment le droit de l'environnement et le droit international humanitaire – ainsi que les objectifs de développement durable et d'autres engagements internationaux²⁷.

107 Associé aux compétences du groupe scientifique international et au développement des capacités (voir les recommandations 1, 4 et 5), le dialogue inclusif au sein de l'ONU peut aider les États et les entreprises à actualiser leurs approches et méthodes réglementaires afin de suivre le rythme de l'accélération de l'IA d'une manière interoperable qui favorise un socle commun. Certaines des caractéristiques propres à l'ONU peuvent être utiles à cet égard :

- a. L'ancrage du dialogue inclusif dans l'ensemble des normes de l'ONU, notamment la Charte des Nations Unies, les droits humains et le droit international, peut promouvoir un nivellement par le haut des approches de gouvernance. À l'inverse, en l'absence d'une représentation mondiale comme à l'ONU, l'action collective internationale est soumise à une pression plus forte pour succomber à un nivellement par le bas réglementaire entre les pays sur la sûreté de l'IA et son champ d'application.
- b. L'Organisation des Nations Unies réunissant tous les pays du monde, elle peut également permettre de coordonner les initiatives régionales existantes afin de les rendre plus compatibles entre elles. De nombreux

participants à nos consultations ont demandé que l'ONU soit un lieu privilégié pour assurer une coordination souple entre les initiatives régionales et plurilatérales existantes, compte tenu des diverses valeurs véhiculées par les cultures, les langues et les régions.

- c. Les procédures prévisibles, transparentes, fondées sur des règles et justifiables de l'ONU peuvent permettre un engagement politique continu pour rapprocher les pays de sensibilités différentes et modérer les contestations dangereuses. Outre l'instauration d'un climat de confiance, de relations et de lignes de communication en temps de crise, des dialogues fiables et inclusifs peuvent favoriser l'apparition de nouvelles normes, d'un droit coutumier et d'accords qui renforcent la coopération entre les États.

108 Sur le plan opérationnel :

- a. Une concertation pourrait s'engager en marge des réunions qui se tiennent à New York, notamment à l'Assemblée, à Genève²⁸, et dans le monde du Sud.
- b. Une partie de chaque session de dialogue pourrait se concentrer sur les stratégies nationales adoptées par les États Membres, tandis qu'une seconde partie serait consacrée au savoir-faire et aux contributions des principales parties prenantes.
- c. La participation des gouvernements pourrait être ouverte à tous les États Membres, ou à un groupe équilibré sur le plan régional (pour un débat plus ciblé entre un sous-ensemble représentatif et tournant), ou une combinaison des deux, calibrée de manière appropriée pour différents points ou segments de l'ordre du jour au fil du temps, à mesure que la technologie évolue et que les préoccupations mondiales

²⁷ Voir, par exemple, la Charte des Nations Unies (préambule, buts et principes, et art. 13, 55, 58 et 59). Voir également des instruments internationaux fondamentaux sur les droits humains (Déclaration universelle des droits de l'homme ; Pacte international relatif aux droits civils et politiques ; Pacte international relatif aux droits économiques, sociaux et culturels ; Convention internationale sur l'élimination de toutes les formes de discrimination raciale ; Convention relative aux droits de l'enfant ; Convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes ; Convention contre la torture ; Convention relative aux droits des personnes handicapées ; Convention sur les droits des migrants ; Convention internationale pour la protection de toutes les personnes contre les disparitions forcées) ; des instruments internationaux relatifs aux droits humains (conventions de Genève, Convention sur certaines armes classiques ; Convention sur le génocide ; Convention de La Haye) ; des instruments sur des principes connexes tels que la distinction, la proportionnalité et la précaution, ainsi que les 11 principes relatifs aux systèmes d'armes létaux autonomes adoptés dans le cadre de la Convention sur certaines armes classiques) ; des instruments sur le désarmement et la maîtrise des armements en termes d'interdiction des armes de destruction massive (Traité sur la non-prolifération des armes nucléaires ; Convention sur les armes chimiques ; Convention sur les armes biologiques) ; des instruments du droit de l'environnement (Convention-cadre des Nations Unies sur les changements climatiques ; Convention sur l'interdiction d'utiliser des techniques de modification de l'environnement à des fins militaires ou toutes autres fins hostiles) ; l'accord de Paris et les principes connexes tels que le principe de précaution, le principe d'intégration et la participation du public ; et des engagements non contraignants sur le Programme 2030, le genre et l'éthique, comme la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle.

²⁸ Analogue au forum politique de haut niveau pour le développement durable dans le contexte des objectifs de développement durable qui se tient sous les auspices du Conseil économique et social.

se font jour ou prennent de l'importance. Une géométrie fixe pourrait ne pas être utile compte tenu du caractère dynamique de la technologie et du contexte stratégique.

- d. Outre les sessions de dialogue formelles, la collaboration multipartite concernant la politique en matière d'IA pourrait également s'appuyer sur d'autres dispositifs existants tels que la réunion AI for Good de l'UIT, la réunion annuelle du Forum sur la gouvernance d'Internet, le Forum mondial de l'UNESCO sur l'éthique de l'IA et la eWeek de la CNUCED – ouverte à la participation de représentants de tous les États Membres sur la base du volontariat.
- e. Conformément au caractère inclusif du dialogue, les ordres du jour des débats pourraient être larges de manière à englober des perspectives et des préoccupations diverses. Par exemple, une première réunion semestrielle pourrait se concentrer sur les possibilités fournies par les différents secteurs et la seconde sur les risques²⁹. Ces réunions pourraient porter notamment sur l'utilisation de l'IA pour atteindre les objectifs de développement durable, la protection des enfants et la réduction au minimum des effets sur le climat et être l'occasion un échange sur les approches de gestion des risques. Elles pourraient également comprendre un débat sur les définitions des termes utilisés dans la gouvernance de l'IA et les normes techniques de l'IA, ainsi que sur les rapports du groupe scientifique international, le cas échéant.
- f. En outre, diverses parties prenantes, en particulier les entreprises technologiques et les représentants de la société civile, pourraient être invitées à se mobiliser par l'intermédiaire des organismes existants décrits ci-après ainsi que lors d'ateliers stratégiques sur tel ou tel

aspect de la gouvernance de l'IA, notamment les limites (le cas échéant) des approches à source ouverte pour les formes les plus avancées d'IA, les seuils de suivi et le signalement des problèmes liés à l'IA, l'application de la législation sur les droits humains aux nouveaux cas d'utilisation, ou l'application du droit de la concurrence pour traiter les concentrations de pouvoir parmi les entreprises technologiques³⁰.

- g. Le bureau de l'IA proposé pourrait également gérer un référentiel d'exemples de gouvernance de l'IA, comprenant la législation, les politiques et les organismes du monde entier, en vue de les examiner dans le cadre de la concertation, en collaboration avec les travaux existants, par exemple ceux menés par l'OCDE.

109 Malgré deux résolutions sur l'IA adoptées par l'Assemblée en 2024, il n'existe actuellement aucun dispositif institutionnalisé de dialogue sur la gouvernance de l'IA à l'ONU correspondant à la vision inclusive fiable de cette recommandation. Des processus semblables existent au niveau international, mais principalement dans le cadre de constellations régionales ou plurilatérales (par. 57), qui ne sont pas toujours inclusives et globales.

110 En complément d'un processus fluide de sommets plurilatéraux et régionaux sur l'IA³¹, l'ONU peut fournir un lieu stable de concertation sur la gouvernance de l'IA. L'inclusion dès la conception, qui est une condition essentielle pour jouer un rôle stabilisateur en des temps géopolitiquement délicats, peut également combler les lacunes en matière de représentation et de coordination recensées aux paragraphes 64 et 72, au moyen d'une action collective plus efficace sur la gouvernance de l'IA dans l'intérêt commun de tous les pays.

29 Certaines parties du système des Nations Unies pourraient être mobilisées pour mettre en évidence les possibilités et les risques, notamment l'UIT pour les normes en matière d'intelligence artificielle ; l'UIT, la CNUCED, le PNUD et le Bureau de la coordination des activités de développement sur les applications de l'IA pour les objectifs de développement durable ; l'UNESCO sur l'éthique et la capacité de gouvernance ; le Haut-Commissariat des Nations Unies aux droits de l'homme sur la responsabilité en matière de droits de l'homme sur la base des normes et dispositifs existants ; le Bureau des affaires de désarmement sur la réglementation de l'IA dans les systèmes militaires ; le PNUD sur le soutien aux capacités nationales de développement ; le Forum sur la gouvernance d'Internet pour la collaboration et le dialogue entre les différentes parties prenantes ; l'OMPI, l'OIT, l'OMS, la FAO, le PAM, le Haut-Commissariat des Nations Unies pour les réfugiés, l'UNESCO, l'UNICEF, l'OMM et d'autres organismes sur les applications sectorielles et la gouvernance.

30 Ce rassemblement pourrait également être l'occasion d'un débat multipartite sur un éventuel durcissement de la gouvernance mondiale de l'IA. Il peut s'agir, par exemple, d'interdire la mise au point de systèmes d'IA incontrôlables ou incontrôlés, ou d'exiger que tous les systèmes d'IA soient suffisamment transparents pour que leurs conséquences puissent être retracées jusqu'à un acteur juridique susceptible d'en assumer la responsabilité.

31 Bien que les nombreux sommets sur l'IA aient aidé un sous-ensemble de 20 à 30 pays à s'aligner sur les questions de sûreté liées à l'IA, la participation a été irrégulière : le Brésil, la Chine et l'Irlande ont approuvé la Déclaration de Bletchley en novembre 2023, mais pas la Déclaration ministérielle de Séoul six mois plus tard (voir fig. 12), tandis que le Mexique et la Nouvelle-Zélande ont approuvé la Déclaration de Séoul mais pas la Déclaration de Bletchley.

Centre d'échange de normes en matière d'IA

Recommandation 3 : centre d'échange de normes en matière d'IA

Nous recommandons la création d'un centre d'échange de normes en matière d'IA, réunissant des représentants d'organismes nationaux et internationaux de normalisation, d'entreprises technologiques et de la société civile ainsi que des représentants du groupe scientifique international.

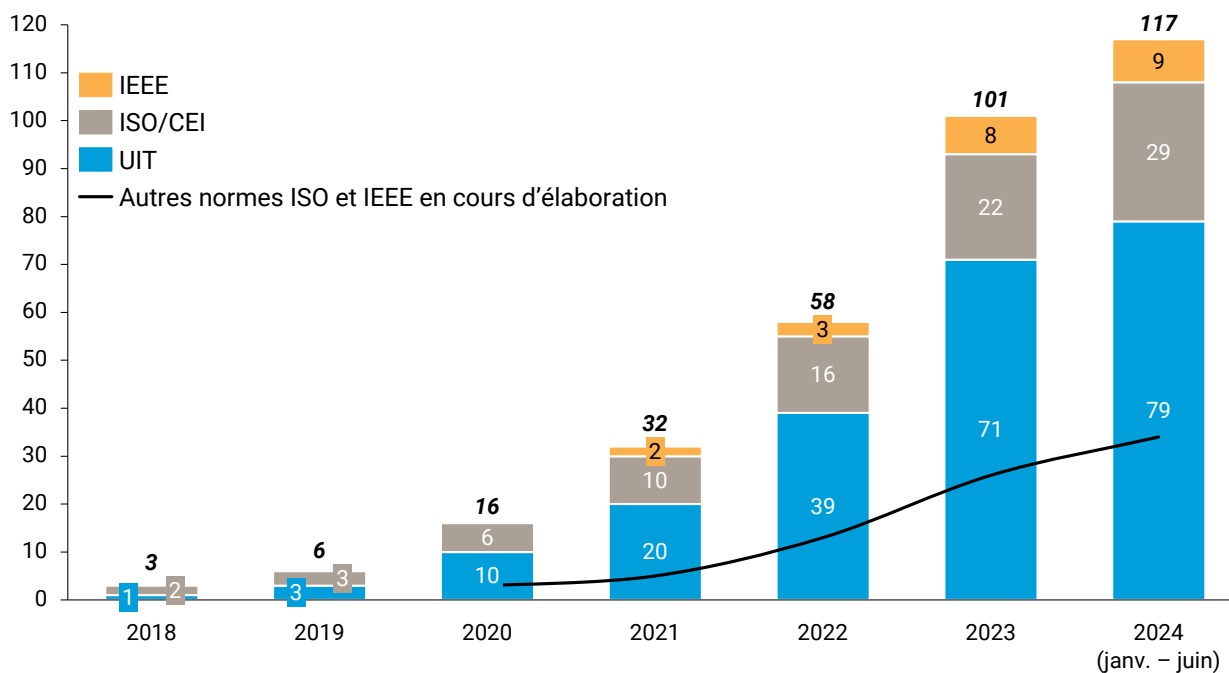
Ce centre aurait les fonctions suivantes :

- Élaborer et tenir à jour un registre des définitions et des normes applicables à la mesure et à l'évaluation des systèmes d'IA ;
- Examiner et évaluer les normes et les processus d'élaboration ;
- Recenser les lacunes qui doivent être comblées par de nouvelles normes.

111 Lorsque les systèmes d'IA ont été étudiés pour la première fois, il existait peu de normes pour aider à naviguer ou à mesurer cette nouvelle frontière. Le test de Turing, qui consiste à déterminer si une machine peut avoir un comportement équivalent (ou impossible à différencier) de celui d'un être humain, a captivé l'imagination populaire, mais sa signification est plus culturelle que scientifique. Il est d'ailleurs révélateur que certaines des plus grandes avancées informatiques aient été mesurées à l'aune de leur succès dans des jeux, comme lorsqu'un ordinateur a pu battre des humains aux échecs, au jeu de Go, au poker ou à Jeopardy. Ces mesures sont facilement compréhensibles par des non-spécialistes, mais ne sont ni rigoureuses ni particulièrement scientifiques.

112 Plus récemment, on a assisté à une multiplication des normes. La figure 13 illustre l'augmentation du nombre de normes pertinentes adoptées par l'UIT, l'ISO, la CEI et l'IEEE³².

Figure 13 : Nombre de normes relatives à l'IA



Sources : IEEE, ISO/CEI, UIT, Coopération mondiale de la normalisation (sur la base de la cartographie de juin 2023, élargie par l'inclusion des normes liées à l'IA).

32 De nombreuses nouvelles normes apparaissent également au niveau national et multinational, telles que les engagements volontaires de la Maison Blanche en matière d'IA aux États-Unis et les codes de pratique de la loi sur l'IA de l'Union européenne.

- 113** Deux tendances se dégagent. Premièrement, ces normes ont été élaborées principalement pour répondre à des questions précises. Il n'existe pas de langage commun et de nombreux termes couramment utilisés en rapport avec l'IA (équité, sûreté, transparence) ne font pas l'objet de définitions convenues ou ne sont pas mesurables (malgré les travaux récents de l'OCDE et du National Institute of Standards and Technology, qui ont adopté une nouvelle approche pour les systèmes dynamiques tels que l'IA).
- 114** Deuxièmement, il existe des divergences entre les normes qui ont été adoptées à des fins techniques étroites ou de validation interne, et celles qui visent à intégrer des principes éthiques plus larges. Les informaticiens et les experts en sciences sociales donnent souvent des interprétations différentes d'une même notion et un paradigme commun de normes sociotechniques est prometteur, mais reste encore à l'état d'hypothèse (voir encadré 10).
- 115** Il en résulte un nouvel ensemble de normes qui ne font pas l'objet d'une définition commune ou qui sont dissociées des valeurs qu'elles étaient censées défendre. Surtout, il existe peu de normes reconnues en matière de consommation d'énergie et d'intelligence artificielle. Il convient également de remédier au manque d'intégration des considérations relatives aux droits humains dans les processus de normalisation³³.
- 116** Cela a un coût réel. Outre les préoccupations d'États Membres et de personnes, bon nombre de nos consultations ont révélé que les entreprises (notamment de petites et moyennes entreprises des pays en développement) craignaient qu'une gouvernance fragmentée et des normes incohérentes augmentent les coûts des activités commerciales dans un monde de plus en plus globalisé.
- 117** Le présent rapport ne vise pas à proposer que l'ONU contribue à cette multiplication de normes mais serve de centre d'échange de normes en matière d'IA applicables à l'échelle mondiale, en s'appuyant sur les compétences du groupe scientifique international (proposé dans la recommandation 1) et en intégrant des membres des différentes entités qui ont contribué à l'élaboration de normes, ainsi que des représentants d'entreprises de technologie et de la société civile³⁴.
- 118** La valeur ajoutée de l'ONU serait de favoriser les échanges entre le plus grand nombre d'organismes de normalisation afin de maximiser l'interopérabilité mondiale des normes techniques et d'intégrer les nouvelles connaissances sur l'élaboration de normes sociotechniques dans les débats sur les normes d'intelligence artificielle.
- 119** La collecte et la diffusion d'informations sur les normes en matière d'IA, qui s'appuient sur les travaux en cours, notamment d'AI Standards Hub³⁵, permettraient aux participants de tous les organismes de normalisation de converger vers un langage commun dans des domaines clés.
- 120** Soutenu par le bureau de l'IA proposé, l'échange de normes bénéficierait également de liens étroits avec le groupe scientifique international pour les questions techniques et avec la concertation pour les questions morales, éthiques, réglementaires, juridiques et politiques.
- 121** En cas d'accord, l'UIT, l'ISO/CEI et l'IEEE pourraient diriger conjointement un premier sommet sur les normes en matière d'IA, avec un suivi annuel pour maintenir l'intérêt et l'élan. Afin de jeter les bases d'une approche sociotechnique intégrant des considérations économiques, éthiques et relatives aux droits humains, il convient de collaborer avec l'OCDE, l'OMPI, l'OMC, le Haut-Commissariat aux droits de l'homme, l'OIT, l'UNESCO et d'autres entités compétentes de l'ONU³⁶.

33 Voir A/HRC/53/42 (Les droits de l'homme et les processus de normalisation technique relatifs aux nouvelles technologies numériques : Rapport du Haut-Commissariat des Nations Unies aux droits de l'homme) et résolution 53/29 du Conseil des droits de l'homme (Nouvelles technologies numériques et droits de l'homme).

34 Même cette tâche peut sembler difficile, mais les progrès réalisés en vue d'un accord sur la taxe minimale mondiale montrent qu'il est possible d'agir collectivement, même dans des domaines économiquement et politiquement complexes.

35 Voir <https://aistandardshub.org>.

36 Cela pourrait concerner des organismes de normalisation sectoriels, nationaux et régionaux compétents.

Encadré 10 : Normes applicables à la sûreté de l'IA

Une approche globale de la sûreté de l'IA implique de comprendre les capacités des modèles d'IA avancés, d'adopter des normes pour une conception et un déploiement sûrs, et d'évaluer à la fois les systèmes et leurs effets plus larges.

Auparavant, les normes relatives à l'IA étaient principalement axées sur les spécifications techniques, qui détaillaient la manière dont les systèmes devaient être construits et exploités. Toutefois, comme les technologies de l'IA ont de plus en plus d'effets sur la société, il convient de passer à un paradigme sociotechnique. Cette évolution tient compte du fait que les systèmes d'IA n'existent pas en vase clos mais qu'ils interagissent avec les utilisateurs humains et qu'ils ont des effets sur les structures sociétales. Les normes modernes en matière d'IA peuvent intégrer des considérations éthiques, culturelles et sociétales en plus des exigences techniques. Dans le contexte de la sûreté, il s'agit notamment de garantir la fiabilité et l'interprétabilité, ainsi que d'évaluer et d'atténuer les risques pour les droits individuels et collectifs^a, la sécurité nationale et internationale et la sûreté publique dans différents contextes.

L'un des principaux objectifs des organismes nationaux de sûreté de l'IA récemment créés est de garantir des approches cohérentes et efficaces de la sûreté de l'IA. L'harmonisation de ces approches permettrait aux systèmes d'IA de satisfaire à des critères de sûreté élevés au niveau international, ce qui favoriserait l'innovation et le commerce transfrontaliers et permettrait de maintenir des protocoles de sûreté rigoureux.

Dans la mesure où la « sûreté » est contextuelle, la participation de diverses parties prenantes et cultures à l'élaboration de ces normes renforce leur pertinence et leur efficacité et contribue à une vision commune des définitions et des notions. Lorsque les protocoles intègrent des perspectives diverses, ils peuvent traiter de manière plus approfondie le large éventail de risques et d'avantages qui peuvent être associés aux technologies de l'IA.

^a Voir A/HRC/53/42 [Les droits de l'homme et les processus de normalisation technique relatifs aux nouvelles technologies numériques: Rapport du Haut-Commissariat des Nations Unies aux droits de l'homme] et résolution 53/29 du Conseil des droits de l'homme (Nouvelles technologies numériques et droits de l'homme)].

122 Le centre d'échange devrait également appuyer les travaux de renforcement des capacités prévu dans la recommandation 4, en veillant à ce que les normes soutiennent les pratiques sur le terrain. Il pourrait mettre en commun des informations sur les outils élaborés au niveau national ou régional qui permettraient une auto-évaluation de la conformité aux normes.

123 Il n'est pas proposé à l'heure actuelle que l'ONU joue un rôle plus important que celui d'un forum de débat et d'accord sur les normes. Dans la mesure où les normes de sûreté sont formalisées au fil du temps, elles pourraient servir de base aux travaux de surveillance et de vérification éventuellement menés par un organisme.

C. Avantages communs

124 Le Programme 2030, qui contient 17 objectifs de développement durable, peut conférer un objectif unique à l'IA, celui de détourner l'arc des investissements du gaspillage et de l'utilisation nocive et de l'orienter vers les défis du développement mondial. Dans le cas contraire, les investissements continueront de servir les profits, même au prix d'externalités négatives pour d'autres. L'ONU peut également contribuer à lier l'application positive de l'IA à la garantie d'une répartition équitable de ses possibilités (encadré 11).

Encadré 11 : L'IA et les objectifs de développement durable

Les progrès scientifiques que l'IA peut permettre de réaliser (encadré 1) et les débouchés économiques qu'elle peut créer (encadré 2) laissent espérer qu'elle accélère les progrès dans la réalisation des objectifs de développement durable. Un examen des données pertinentes réalisé en 2023 a fait valoir que l'IA pouvait jouer un rôle de catalyseur pour 134 cibles (79 %) dans l'ensemble des objectifs de développement durable, généralement grâce à des améliorations technologiques susceptibles de permettre de surmonter certaines limites qui prévalent^a.

Un aperçu des perceptions actuelles des experts a été obtenu au moyen d'une analyse des opportunités réalisée dans le cadre de nos travaux, qui a consisté à interroger plus de 120 experts de 38 pays sur les effets positifs qu'ils attendaient de l'IA en termes de percées scientifiques, d'activités économiques et d'objectifs de développement durable. L'enquête menée n'a posé que des questions sur les éventuels effets positifs de l'IA.

Dans l'ensemble, les experts avaient des attentes mitigées quant au délai dans lequel l'IA pourrait avoir un effet positif majeur (voir également fig. 14) :

- Les experts se sont montrés plus optimistes quant à l'**accélération des découvertes scientifiques**, 7 sur 10 estimant que l'IA était susceptible d'avoir des effets positifs majeurs au cours des trois années suivantes ou plus tôt dans les pays à revenu élevé/intermédiaire supérieur, et 28 % prévoyant de même pour les pays à revenu faible/intermédiaire inférieur.
- Environ 5 experts sur 10 prévoient des effets positifs majeurs sur l'**augmentation de l'activité économique** dans les trois années suivantes ou plus tôt dans les pays à revenu élevé/intermédiaire supérieur, et 32 % prévoient de même dans les pays à revenu faible/intermédiaire inférieur.
- Au total, 46 % des experts s'attendaient à des effets positifs majeurs sur la **réalisation des objectifs de développement durable** au cours des trois années suivantes ou plus tôt dans les pays à revenu élevé ou intermédiaire de la tranche supérieure. Toutefois, seuls 21 % d'entre eux s'attendent à ce que ce soit le cas dans les pays à revenu faible/intermédiaire inférieur, 4 experts sur 10 estimant que des effets positifs majeurs sur les objectifs de développement durable ne se produiront probablement pas avant au moins 10 ans dans ces pays.

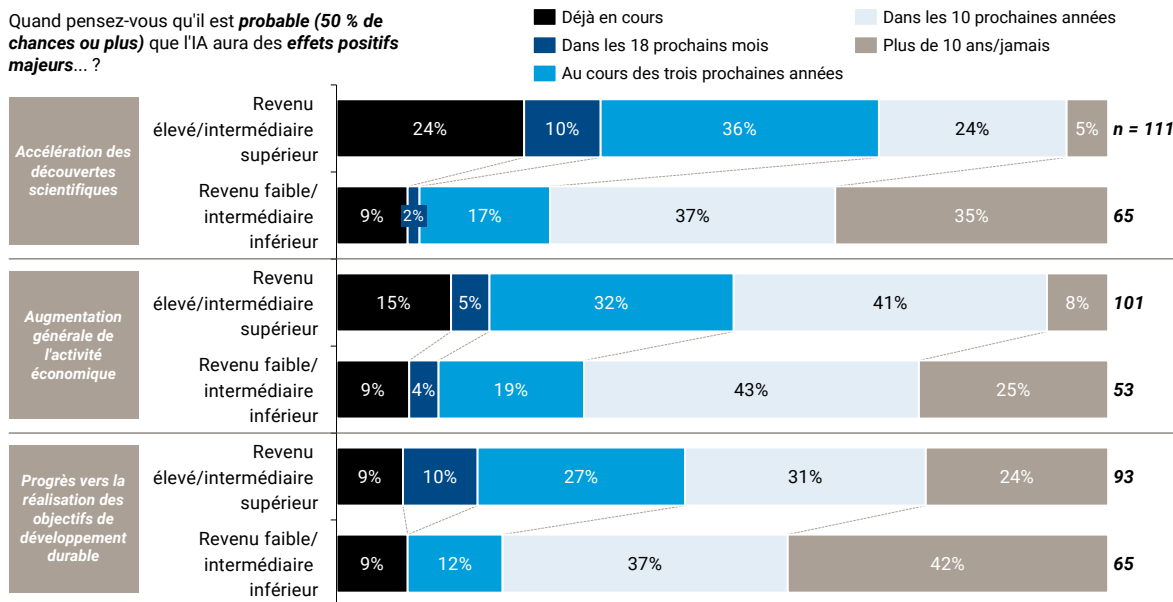
Les experts prévoient des effets positifs plus importants de l'IA au cours des trois prochaines années dans les pays à revenu élevé dans tous les domaines visés par l'enquête, notamment l'accélération des découvertes scientifiques, l'augmentation de l'activité économique^b et dans les 14 domaines des objectifs de développement durable sur lesquels portait la question (voir fig. 15). Ils se sont montrés les plus optimistes quant aux effets positifs de l'IA sur la santé et l'éducation (objectifs 3 et 4), 20 à 25 % d'entre eux prévoyant des effets positifs majeurs ou transformateurs de l'IA au cours des trois années suivantes dans les pays à revenu élevé/intermédiaire supérieur. Ils étaient moins optimistes quant aux effets positifs de l'IA sur l'égalité des genres et les inégalités (objectifs 5 et 10), 2 sur 3 prévoyant que l'IA n'aura aucun effet positif sur la réduction des inégalités au sein des pays ou entre les pays, que ce soit dans les pays à revenu élevé ou faible.

a Voir Ricardo Vinuesa et al., « Le rôle de l'intelligence artificielle dans la réalisation des objectifs de développement durable ». *Nature Communication*, vol. 11, n° 233 (janv. 2020). Cette étude a également fait valoir que 59 cibles (35 %, également sur l'ensemble des objectifs de développement durable) pourraient subir des effets négatifs du développement de l'IA.

b La proportion d'experts s'attendant à des « effets positifs majeurs » sur l'augmentation de l'activité économique et l'accélération de la découverte scientifique sur trois ans est plus élevée dans le premier graphique que dans le second. Cela peut être dû au qualificatif « d'ici quand pensez-vous qu'il est probable (50 % de chances ou plus) que l'IA aura des effets positifs majeurs » (non souligné dans l'original) dans les réponses à la question illustrée dans le premier graphique, qui est absent dans le second.

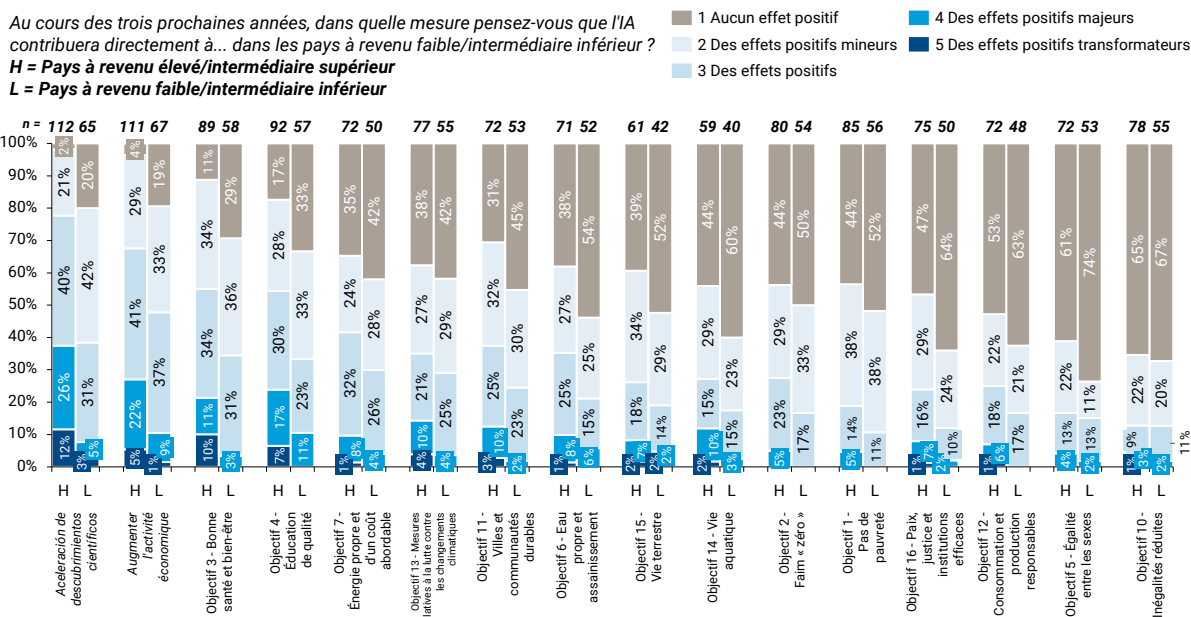
Encadré 11 : L'IA et les objectifs de développement durable (suite)

Figure 14 : Attentes des experts concernant le moment où l'IA aura des effets positifs majeurs, par région



Note : ne prend pas en compte les réponses « Ne sait pas »/« Pas d'opinion » et les réponses en blanc. Les n particuliers sont indiqués à chaque question.
Source : OSET Sondage des opportunités de l'IA, 9-21 août 2024 (n=121, dont 78 ont répondu aux questions sur les pays à revenu faible/intermédiaire inférieur).

Figure 15 : Attentes des experts concernant les effets positifs majeurs de l'IA au cours des trois prochaines années, par domaine et par objectif de développement durable

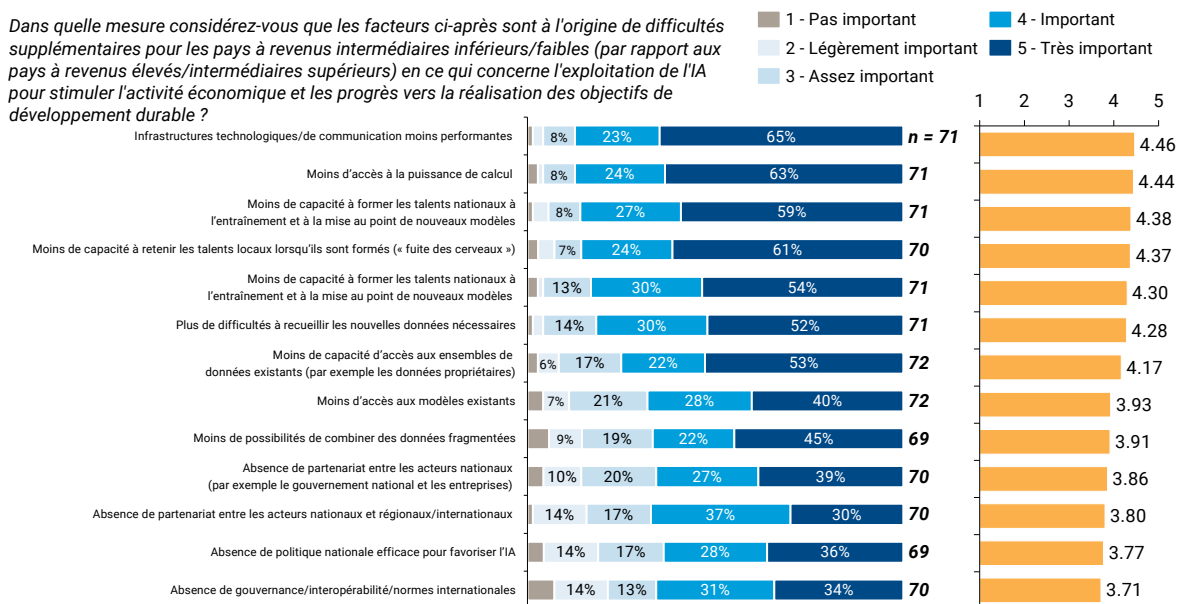


Note : ne prend pas en compte les réponses « Ne sait pas »/« Pas d'opinion » et les réponses en blanc. Les n particuliers sont indiqués à chaque question. N'a pas posé de questions sur les objectifs 8, 9 et 17. Source : OSET Sondage des opportunités de l'IA, 9-21 août 2024 (n=121, dont 78 ont répondu aux questions sur les pays à revenu faible/intermédiaire inférieur).

Encadré 11 : L'IA et les objectifs de développement durable (suite)

On peut s'attendre à ce que l'IA ait des effets plus précoces et plus importants dans les pays à revenu élevé, en partie en raison des obstacles qui freinent les pays à revenu intermédiaire inférieur et à faible revenu (voir fig. 16). Plus de la moitié des répondants ont cité l'absence de catalyseurs – allant d'une infrastructure insuffisante à l'absence de politique nationale et de gouvernance internationale – comme un facteur important des difficultés supplémentaires rencontrées par les pays à faible revenu pour mettre l'IA au service de l'activité économique et des progrès vers la réalisation des objectifs de développement durable.

Figure 16 : Évaluation par les experts des obstacles à l'exploitation de l'IA pour stimuler l'activité économique et progresser dans la réalisation des objectifs de développement durable dans les pays à revenu faible/intermédiaire inférieur



Note : ne prend pas en compte les réponses « Ne sait pas »/« Pas d'opinion » et les réponses en blanc. Seuls les répondants déclarant avoir des connaissances pertinentes ont été interrogés sur les pays à revenu moyen inférieur et à revenu faible. Source : OSET Sondage des opportunités de l'IA, 9-21 août 2024 (n=121, dont 78 ont répondu aux questions sur les pays à revenu faible/intermédiaire inférieur).

Ces résultats soulignent le caractère provisoire de la contribution que l'IA pourrait apporter aux objectifs de développement durable et le fait qu'elle reste fortement tributaire des éléments facilitateurs manquants. C'est particulièrement le cas dans les pays moins développés, qui ne disposent déjà pas de la plupart des éléments dont disposent les pays plus développés, qu'il s'agisse d'infrastructures ou de politiques. Sans coopération pour renforcer les capacités et faciliter l'accès aux outils clés, les fractures existantes en matière d'IA pourraient encore se creuser et s'enraciner, limitant la capacité de l'IA à contribuer efficacement aux progrès de la science, aux avantages économiques et à la réalisation des objectifs de développement durable d'ici à 2030.

- 125** Comme nous l'avons affirmé dans notre rapport d'étape, cela dépend dans une large mesure de l'accès aux talents, à la puissance de calcul et aux données, selon des modalités qui permettent à la diversité culturelle et linguistique de se développer. La gouvernance elle-même peut être un catalyseur clé dans la mesure où elle facilite les mesures incitatives, suscite la confiance et favorise les pratiques durables, en même temps qu'elle promeut la collaboration au-delà des frontières et des domaines. Sans une approche globale et inclusive de la gouvernance, l'IA risque de ne pas contribuer positivement à la réalisation des objectifs de développement durable et son déploiement pourrait, par inadvertance, renforcer les disparités et les préjugés.
- 126** Au cours des consultations approfondies menées par l'Organe consultatif sur des sujets tels que l'éducation, la santé, les données, le genre, les enfants, la paix et la sécurité, les industries créatives et le travail, il est devenu évident que, grâce à l'IA et à ses capacités à stimuler l'innovation et l'exécution dans divers domaines essentiels, il est possible d'accélérer considérablement la réalisation des objectifs de développement durable.
- 127** Cependant, l'IA n'est pas une solution miracle aux problèmes de développement ; il s'agit d'une composante d'un ensemble plus large de solutions, qui peut même exacerber certains de ces défis, tels que le changement climatique. Afin de tirer pleinement parti de la capacité de l'IA à relever les défis sociétaux, la collaboration entre les gouvernements, le milieu universitaire, l'industrie et la société civile est cruciale.
- 128** L'efficacité des solutions liées à l'IA dépend de la qualité et de la disponibilité des données, et la qualité et la représentativité des ensembles de données relatives aux objectifs de développement durable suscitent de vives inquiétudes car elles risquent de ne pas tenir compte des réalités de certaines populations. En outre, les solutions liées à l'IA conçues par des experts en IA qui n'ont pas une connaissance approfondie des domaines d'application croisés fonctionnent souvent en vase clos et ne sont pas suffisamment robustes ou efficaces dans des contextes de développement réels. Elles doivent donc être conçues en collaboration et mises en œuvre avec une compréhension approfondie de leur contexte social, économique et culturel. Elles doivent s'inscrire dans des stratégies locales et nationales plus larges de transformation numérique et de lutte contre la fracture numérique.
- 129** Par exemple, les capacités d'IA dans les pays à faible revenu et à revenu intermédiaire inférieur ne peuvent pas être exploitées sans garantir une électricité et une connectivité Internet fiables pour faire fonctionner les centres de données, maintenir des opérations informatiques cohérentes, accéder à des ensembles de données mondiales, mener des collaborations de recherche internationales et utiliser des outils d'IA basés sur l'informatique dématérialisée. C'est pourquoi nous nous associons aux appels à investir dans l'infrastructure numérique de base, qui est une condition préalable pour que les pays en développement puissent participer aux progrès de l'IA et en bénéficier.
- 130** Il est essentiel de renforcer les capacités en matière d'IA pour que les populations à travers le monde, quel que soit le stade de développement de leur région, puissent bénéficier des progrès de l'IA. Le renforcement stratégique des capacités, soutenu par un financement adéquat, est également essentiel pour rendre les technologies de l'IA efficaces, durables et axées sur l'intérêt général, et contribuer aux efforts de développement au niveau mondial. Nous examinons ci-après trois catalyseurs essentiels des capacités nationales en matière d'IA : la disponibilité des compétences technique, l'accès à la puissance de calcul et la disponibilité de données de qualité. Nous recommandons ensuite des actions concrètes.

Talent

- 131** La capacité des sociétés du monde entier à participer aux résultats positifs de l'IA dépend avant tout des personnes. Il convient de reconnaître que toutes les sociétés n'ont pas besoin de cadres d'informaticiens pour construire leurs propres modèles. Toutefois, que la technologie soit achetée, empruntée ou construite, des ressources humaines sont nécessaires pour comprendre les capacités et les limites de l'IA et l'utiliser au mieux.
- 132** Une telle capacité, principalement dans le secteur public, mais aussi dans le milieu universitaire, les entreprises et la société civile, renforcera l'efficacité des stratégies d'IA et leur mise en œuvre dans les différents secteurs. Le développement des capacités humaines liées à l'IA sera également essentiel pour préserver la diversité culturelle et linguistique du monde et constituer des ensembles de données de haute qualité pour le développement futur de l'IA. Il s'agit essentiellement d'un renforcement des capacités de l'IA d'intérêt général.
- 133** Le renforcement des ressources humaines dans des environnements diversifiés à la démographie jeune, comme en Afrique (un tiers de la main-d'œuvre mondiale sera africaine au cours de la première moitié de ce siècle) sera également vital pour la future réserve mondiale de talents. Le renforcement des capacités des femmes dans la technologie doit viser à combler le fossé existant entre les hommes et les femmes, d'une part, et à éviter le fossé entre les hommes et les femmes dans le domaine de l'intelligence artificielle, d'autre part. Le secteur de l'IA a également besoin d'un plus grand nombre de femmes à des postes de direction afin de tenir compte des questions de genre dans la gouvernance de l'IA. Il s'agit tout d'abord de permettre aux filles d'exploiter les atouts qu'elles portent en elles dans le domaine de l'IA.

Puissance de calcul

- 134** Malgré les efforts déployés en vue de mettre au point des approches de l'IA moins exigeantes en ressources informatiques, le besoin d'accès à des ressources informatiques abordables reste important pour l'entraînement de modèles d'IA performants³⁷. C'est l'un des principaux obstacles à l'entrée dans le domaine de l'IA pour les entreprises du monde du Sud, mais aussi pour de nombreuses startups et PME du monde du Nord. Sur les 100 premiers groupes de calcul haute performance au monde capables d'entraîner de grands modèles d'IA, aucun ne se trouve dans un pays en développement³⁸. Un seul pays d'Afrique figure parmi les 300 premiers. Deux pays hébergent la moitié des centres de données à grande échelle du monde³⁹.
- 135** La plupart des développeurs accèdent à l'infrastructure informatique au moyen de services en nuage ; nombre d'entre eux ont choisi de s'associer avec les grandes sociétés d'informatique en nuage pour s'assurer un accès fiable à la puissance de calcul. Il est possible que les problèmes liés à la chaîne d'approvisionnement soient résolus au fil du temps et que la concurrence conduise à une plus grande diversité des sources de matériel, y compris des puces à haute performance pour les modèles d'entraînement et des puces d'accélération de l'IA pour le déploiement sur les appareils mobiles. Toutefois, dans un avenir proche, cette contrainte restera un obstacle de taille à un écosystème d'innovation en matière d'IA plus inclusif à l'échelle mondiale.
- 136** Paradoxalement, la capacité de calcul peut rester inactive ou être rapidement dépassée. Utiliser cette capacité de manière optimale tout au long des cycles d'amortissement peut s'avérer utile. Toutefois, il reste des obstacles à surmonter en termes d'interopérabilité des différentes

37 L'organe consultatif a eu connaissance d'un cas récent où une entreprise basée dans le monde du Sud a dépensé un montant de 70 millions de dollars pour trois mois d'entraînement à un grand modèle de langage. Posséder des groupes de calcul haute performance au lieu de les louer à des fournisseurs de services en nuage aurait coûté beaucoup moins cher.

38 Voir <https://top500.org/statistics/sublist>; indicateur supplétif car la plupart des groupes de calcul haute performance n'ont pas de processeurs graphiques et sont d'une utilité limitée pour l'IA avancée.

39 CNUCED, *Rapport sur l'économie numérique 2021* (Genève, 2021).

configurations matérielles et d'ordonnement des tâches exigeantes, tout en préservant la priorité de l'utilisation en temps critique (comme pour les prévisions météorologiques).

- 137** De plus, sans talent et sans données, la puissance de calcul en soi n'a aucune valeur. Dans la proposition de fonds mondial pour l'IA, nous examinons comment traiter ces trois aspects moyennant une combinaison de soutien financier et de ressources en nature.

Données

- 138** Bien que de nombreux débats sur l'économie de l'IA se concentrent sur la « guerre des talents » et la concurrence pour le matériel, comme les unités de traitement graphique, les données ne sont pas moins vitales. Faciliter l'accès à des données d'entraînement de qualité à grande échelle pour l'entraînement des modèles d'IA par les startups et les PME, ainsi que des dispositifs de compensation des détenteurs de données et des créateurs de données d'entraînement dans le respect des droits, pourrait être le plus important catalyseur d'une économie de l'IA dynamique. La mise en commun des données dans l'intérêt général afin de réaliser certains objectifs de développement durable est un aspect essentiel (tel que décrit à l'encadré 12), mais cela ne suffit pas.

- 139** Dans le contexte de l'IA, il est courant de parler d'« utilisation abusive » des données (par exemple, violation de la vie privée) ou d'utilisations « manquées » des données (absence d'exploitation des ensembles de données existants), mais un problème connexe est celui des données « manquantes », qui englobe les vastes portions du monde où les données sont rares. Un exemple est celui des soins de santé, où environ la moitié des principaux ensembles de données peuvent être attribués à une douzaine d'organisations, dont une en Europe, une en Asie et le reste en Amérique du Nord⁴⁰.

- 140** Un autre exemple est celui de l'agriculture, où des données sont nécessaires pour une interaction complexe de facteurs (tels que le climat, le sol et les pratiques de gestion des cultures) afin d'obtenir des modèles d'IA utiles. L'agriculture souffre aussi souvent d'un manque de données et d'infrastructures de collecte de données. Il convient de mener des efforts ciblés afin de conserver les ensembles de données agricoles, en particulier dans le contexte de la résilience des systèmes alimentaires face aux changements climatiques.

- 141** À l'instar du problème du capital informel, les personnes dont les données ne sont pas enregistrées (actes de naissance, transactions financières) peuvent être dans l'incapacité de participer aux avantages de l'économie de l'IA, d'obtenir des prestations gouvernementales ou d'accéder au crédit. L'utilisation de données synthétiques ne peut compenser que partiellement le besoin de nouveaux ensembles de données.

- 142** Les commentaires reçus au sujet de notre rapport intérimaire ont montré que la manière dont les pratiques transnationales actuelles en matière d'approvisionnement, d'utilisation et de non-divulgaration des données d'entraînement à l'IA menaçaient les droits et aboutissaient à une concentration économique n'était pas suffisamment expliquée. Il a été recommandé d'examiner comment la gouvernance internationale de l'IA pourrait favoriser et catalyser une participation plus diversifiée à l'exploitation des données en matière d'IA.

40 Voir <https://2022.internethalthreport.org/facts>.

Encadré 12 : Mise en commun des données pour l'intérêt général dans les domaines des objectifs de développement durable

Les données collaboratives et les biens communs de l'IA, à savoir des modèles communs sont entraînés sur des données regroupées, peuvent jouer un rôle clé dans la promotion de l'intérêt général lorsque les données seraient autrement manquantes ou trop rares pour tirer profit de l'IA. Le regroupement de données interfonctionnelles et multi-domaines pourrait permettre l'établissement d'ensembles de données transdisciplinaires englobant divers domaines des objectifs de développement durable et provenant de diverses sources.

À titre d'exemple, nous pouvons considérer la question complexe de l'évaluation des effets des changements climatiques sur la santé. Afin de relever efficacement ce défi, il est essentiel d'adopter une approche transdisciplinaire intégrant des données épidémiologiques sur la prévalence des maladies et des données météorologiques sur les variations climatiques. En regroupant ces différents types de données provenant de pays du monde entier, tout en préservant la vie privée, les chercheurs pourraient être en mesure d'utiliser l'IA afin de définir des modèles et des corrélations qui ne sont pas évidents à partir d'ensembles de données isolés.

L'inclusion de données provenant de tous les pays garantit une couverture complète, tenant compte de la nature mondiale des changements climatiques et des divers effets environnementaux et les résultats en matière de santé dans les différentes régions. Les origines transdisciplinaires des données améliorent la précision prédictive des modèles qui visent à prévoir les futures crises de santé publique ou les catastrophes naturelles provoquées par les changements climatiques.

Renforcer les capacités publiques internationales en matière d'IA dans l'intérêt commun

143 Avec les trois catalyseurs susmentionnés, les économies avancées ont à la fois les moyens et le devoir de faciliter le renforcement des capacités en matière d'IA dans le cadre d'une collaboration internationale. En retour, elles bénéficieront d'une économie numérique plus étendue ainsi que de flux de talents et de données de qualité. Il est important de noter que tout le monde tirera profit de l'intégration de la bonne gouvernance de l'IA grâce à cette collaboration.

144 La coopération devrait se concentrer sur l'épanouissement des talents en matière d'IA, le renforcement des connaissances du public en matière d'IA, l'amélioration des capacités de gouvernance de l'IA, l'élargissement de l'accès à

l'infrastructure de l'IA, la promotion de plateformes de données et de connaissances adaptées aux divers besoins culturels et régionaux, et l'amélioration de l'adoption des applications de l'IA et des capacités de service. Seule une telle approche globale peut garantir un accès équitable aux avantages de l'IA, afin qu'aucune nation ne soit laissée de côté.

145 De nombreuses parties prenantes que nous avons consultées ont souligné qu'il fallait définir des stratégies détaillées pour mettre en commun les ressources mondiales afin de renforcer les capacités, de catalyser l'action collective en vue d'un partage équitable des possibilités offertes et de combler le fossé numérique.

Réseau de développement des capacités

Recommandation 4 : réseau de développement des capacités

Nous recommandons la création d'un réseau de développement des capacités en matière d'IA afin de relier un ensemble de centres de développement des capacités affiliés à l'ONU et collaborant entre eux, et mettre ainsi à la disposition des acteurs clés des compétences, la puissance de calcul et les données d'entraînement en matière d'IA. Les objectifs de ce réseau sont les suivants :

- a. Catalyser et harmoniser les initiatives régionales et mondiales liées aux capacités en matière d'IA au moyen de la mise en réseau ;
- b. Renforcer les capacités de gouvernance des fonctionnaires en matière d'IA afin d'encourager le développement et, en même temps, de favoriser le respect, la protection et la concrétisation des droits humains ;
- c. Mettre à disposition des formateurs, une puissance de calcul et des données d'entraînement à l'IA dans plusieurs centres à l'intention des chercheurs et des entrepreneurs sociaux qui visent à appliquer l'IA à des cas d'utilisation d'intérêt général local, notamment par les moyens suivants :
 - i. Des protocoles permettant aux équipes de recherche interdisciplinaires et aux entrepreneurs d'accéder à la puissance de calcul lorsqu'elle fait défaut aux fins de l'entraînement, de la mise au point et de l'application de leurs modèles en fonction des contextes locaux ;
 - ii. Des dispositifs expérimentaux pour mettre à l'essai d'éventuelles solutions d'IA et apprendre par la pratique ;

- iii. Une série de formations en ligne sur l'IA destinées aux étudiants, aux jeunes chercheurs, aux entrepreneurs sociaux et aux fonctionnaires du secteur public ;
- iv. Un programme de bourses destiné à des personnes prometteuses afin de leur permettre de passer du temps dans des établissements universitaires ou des entreprises technologiques.

146 Depuis les objectifs du Millénaire pour le développement jusqu'aux objectifs de développement durable, l'ONU contribue depuis longtemps à renforcer les capacités des personnes et des organismes⁴¹. Grâce aux travaux de l'UNESCO, de l'OMPI et d'autres organisations, l'ONU contribue à préserver la grande diversité des cultures et des traditions de production de connaissances dans le monde entier.

147 Dans le même temps, le développement des capacités en matière d'IA nécessiterait une nouvelle approche, en particulier une formation interdomaines à l'intention d'une nouvelle génération d'experts multidisciplinaires, notamment dans la santé publique et l'IA, ou les systèmes alimentaires et énergétiques et l'IA.

148 La capacité devrait également être liée aux résultats par une formation pratique au moyen de dispositifs expérimentaux⁴² et de projets collaboratifs regroupant des données et la puissance de calcul afin de résoudre des problèmes communs. Les évaluations des risques, les essais de sûreté et d'autres méthodes de gouvernance devraient être intégrés dans cette infrastructure de formation collaborative.

149 Compte tenu de l'urgence et de l'ampleur du défi, nous proposons d'adopter une approche stratégique qui mette en commun et facilite l'accès à la puissance de calcul au moyen d'un réseau de

41 L'ONU s'emploie depuis longtemps à renforcer les capacités par l'enseignement supérieur et la recherche, et l'Institut des Nations Unies pour la formation et la recherche contribue à former des fonctionnaires dans des domaines essentiels au développement durable. La méthode d'évaluation de l'état de préparation élaborée par l'UNESCO est un outil essentiel pour aider les États Membres à appliquer la recommandation de l'Organisation sur l'éthique de l'IA. Il existe également d'autres exemples, tels que l'Académie de l'OMS à Lyon, l'Institut virtuel de la CNUCED, le Programme de bourses d'études, de formation et de services consultatifs des Nations Unies en matière de désarmement et les programmes de développement des capacités menés par l'UIT et le PNUD.

42 Des dispositifs expérimentaux ont été mis en place par divers organismes nationaux, notamment des autorités financières et médicales, comme l'Infocomm Media Development Authority de Singapour.

pôles de calcul haute performance, encourage le développement d'ensembles de données critiques dans les domaines pertinents pour les objectifs de développement durable, favorise la mise à disposition des modèles d'IA, intègre les meilleures pratiques en matière de gouvernance de l'IA et crée des talents interdisciplinaires pour l'IA d'intérêt général, garantissant ainsi l'intégration transversale des compétences en matière de droits humains.

150 En d'autres termes, au lieu de rechercher les catalyseurs critiques un par un au moyen de projets distincts, nous proposons **une stratégie globale mise en œuvre en une seule fois dans le cadre d'une chaîne de centres de collaboration**. Les nouvelles initiatives sur le développement des capacités et l'IA pour les objectifs de développement durable, telles que l'initiative International Computation and Artificial Intelligence Network lancée par la Suisse, peuvent contribuer à créer la masse critique initiale pour cette stratégie.

151 Idéalement, il devrait y avoir au moins un ou deux pôles dans chaque région du monde. Les deux centres de compétence participant au PMIA pourraient se joindre à l'ONU pour soutenir le réseau. Les institutions universitaires et les contributeurs du secteur privé au développement des capacités pourraient s'affilier au pôle régional le plus proche ou à une organisation internationale soutenant le réseau.

152 Nous sommes particulièrement encouragés par la perspective d'une coopération entre les pays, par exemple au moyen d'un accès fédéré à l'infrastructure informatique et connexe. Comme indiqué dans notre rapport d'étape, l'Organisation européenne pour la recherche nucléaire (CERN) fournit des enseignements utiles. Un « CERN distribué » réimaginé pour l'IA et organisé en réseau entre différents États et régions pourrait élargir les possibilités d'accès aux outils et aux connaissances sur l'IA.

153 Nous envisageons le réseau de développement des capacités comme un catalyseur des capacités nationales et régionales et non comme un concentrateur de matériel, de talents et de données.

En accélérant l'apprentissage, le réseau pourrait catalyser les centres d'excellence nationaux afin de stimuler le développement d'écosystèmes locaux d'innovation en matière d'IA, comblant ainsi les lacunes sous-jacentes en matière de coordination et de mise en œuvre mentionnées aux paragraphes 73, 80 et 81. Les travaux au niveau national pourraient continuer à utiliser des outils de diagnostic tels que la méthode d'évaluation de l'état de préparation à l'IA de l'UNESCO pour déterminer la maturité initiale des pays, recenser les lacunes et orienter la manière dont les plans de renforcement des capacités peuvent être adaptées par pays et par région, le réseau international aidant à combler ces lacunes.

154 Le bureau de l'IA proposé pourrait être le mieux placé pour se concentrer sur la stratégie, les partenariats et l'affiliation afin de relier les pôles au réseau, en servant à connecter plutôt qu'à réinventer. Il pourrait également contribuer à faciliter l'accès à la puissance de calcul à travers le réseau. Un ou plusieurs pôles du réseau pourraient jouer le rôle de chef de file pour certains aspects de la formation et héberger des dispositifs expérimentaux ou des groupes de calcul haute performance pour la mise au point de modèles d'IA. Les pôles pourraient collaborer à des programmes de recherche sur des sujets tels que l'utilisation des données dans le respect de la vie privée, de nouvelles méthodes pour relier différents types de matériel ou d'ensembles de données pour l'apprentissage des modèles, ainsi que des moyens d'utiliser les modèles d'IA en combinaison les uns avec les autres.

155 Nous espérons qu'un tel réseau favorise également un autre paradigme de développement des technologies de l'IA : ascendant, interdisciplinaire, interrégional, ouvert et collaboratif. Compte tenu de l'augmentation des coûts énergétiques et d'autres coûts liés à la formation et au déploiement de modèles d'IA, et de la perspective d'une puissance de calcul inutilisée, il est logique de relier les ressources informatiques pour un accès en temps partagé, tout en tirant parti de cet accès pour faire progresser les talents, les données et les modèles d'IA interdomaines pour les objectifs de développement durable.

Fonds mondial pour l'IA

Recommandation 5 : fonds mondial pour l'IA

Nous recommandons la création d'un fonds mondial pour l'IA afin de réduire la fracture en la matière. Géré par une structure de gouvernance indépendante, le fonds recevrait des contributions financières et des ressources en nature de sources publiques et privées, qui seraient allouées, notamment par l'intermédiaire du réseau de développement des capacités, pour faciliter l'accès aux outils de l'IA en vue de favoriser la réalisation des objectifs de développement durable au niveau local, notamment :

- a. Des ressources de calcul mises en commun pour l'entraînement et la mise au point de modèles par les développeurs d'IA des pays qui ne disposent pas des capacités locales adéquates ou des moyens de se les procurer ;
- b. Des dispositifs expérimentaux et des outils d'analyse comparative et d'essai visant à généraliser les meilleures pratiques en matière de développement de modèles sûrs et fiables et de gouvernance des données ;
- c. Des solutions de gouvernance, de sûreté et d'interopérabilité applicables à l'échelle mondiale ;
- d. Des ensembles de données et des travaux de recherche sur la manière dont les données et les modèles pourraient être combinés afin de mettre en œuvre des projets liés aux objectifs de développement durable ;
- e. Un référentiel de modèles d'IA et d'ensembles de données conservés aux fins de la réalisation des objectifs de développement durable.

156 Le modèle de développement et d'utilisation de l'IA proposé ici est analogue à la vision initiale d'Internet : une infrastructure distribuée mais connectée, interopérable et responsabilisante. Un marché où les modèles d'IA, l'infrastructure et les données sur lesquelles ils reposent sont interopérables, bien gérés et dignes de confiance permettrait de mieux servir l'intérêt général. Cet objectif ne sera pas atteint automatiquement. Il serait essentiel de déployer des efforts soutenus par des ressources suffisantes.

157 Nous formulons cette recommandation avec humilité, conscients des puissantes forces du marché qui déterminent l'accès aux talents et à la puissance de calcul, et de la concurrence géopolitique qui s'oppose à la collaboration dans le domaine de la science et de la technologie. Malheureusement, de nombreux pays pourraient être incapables d'accéder à la formation, à la puissance de calcul, aux modèles et aux données d'entraînement sans un soutien international. Sans ce soutien, les efforts de financement existants pourraient également ne pas être suffisants.

158 L'égalesation des conditions de concurrence est, en partie, une question d'équité. Il est également dans notre intérêt collectif de créer un monde dans lequel tous contribuent à un écosystème partagé et en tirent profit. Cette constatation ne vaut pas seulement pour les États. Garantir un accès diversifié aux infrastructures de mise au point et de mise à l'essai des modèles d'IA permettrait également de répondre aux préoccupations concernant la concentration d'un pouvoir disproportionné entre les mains d'une poignée d'entreprises technologiques.

Buts et objectifs du Fonds

159 En proposant un fonds, notre intention n'est pas de garantir l'accès aux ressources et aux capacités de calcul que même les pays et les entreprises les plus riches peinent à acquérir. La réponse à apporter ne consiste pas toujours à augmenter la puissance de calcul. Nous pourrions également avoir besoin de nouvelles façons d'exploiter les infrastructures de calcul haute performance existantes, qui sont conçues pour un usage intensif et pas nécessairement pour l'IA. Il y aurait peut-être de meilleurs moyens pour relier les talents, la puissance de calcul et les données.

160 L'objectif est donc de combler les lacunes sous-jacentes en matière de coordination et de mise en œuvre visées aux paragraphes 73, 80 et 81 pour celles et ceux qui ne peuvent pas accéder aux éléments facilitateurs requis par d'autres moyens, afin d'obtenir les garanties suivantes :

- a. Permettre aux pays qui en ont besoin d'accéder à des outils d'aide à l'IA, ce qui réduirait la fracture en matière d'IA ;

- b. Favoriser la collaboration pour le développement des capacités en matière d'IA afin de créer des habitudes de coopération et d'atténuer la concurrence géopolitique ;
- c. Inciter les pays ayant des approches réglementaires divergentes à élaborer des modèles communs pour régir les données, les modèles et les applications afin de relever les défis sociétaux liés aux objectifs de développement durable et aux avancées scientifiques.

161 Les capacités renforcées grâce aux ressources du fonds mondial seraient orientées vers les objectifs de développement durable et la gouvernance mondiale partagée de l'IA (encadré 13). Elles pourraient, par exemple, intégrer différents niveaux de gouvernance pour les essais de sécurité et de sûreté. Cela permettrait d'intégrer les meilleures pratiques dans l'ensemble de la base d'utilisateurs tout en réduisant la charge de la validation pour les petits utilisateurs.

162 L'accent mis sur l'intérêt général rend le fonds mondial complémentaire de la proposition de réseau de développement des capacités en matière d'IA, vers lequel le fonds canaliserait des ressources. Celui-ci fournirait également une capacité indépendante de suivi de l'impact. Nous veillons ainsi à ce que de vastes régions du monde ne soient pas laissées pour compte et aient au contraire les moyens de mettre l'IA au service des objectifs de développement durable dans différents contextes.

163 Il est dans l'intérêt de tous qu'une coopération soit établie dans le monde tant numérique que physique. Des analogies peuvent être faites avec les efforts de lutte contre les changements climatiques, où les coûts de la transition, de l'atténuation ou de l'adaptation ne sont pas répartis de manière égale, et où l'assistance internationale est essentielle pour que les pays aux ressources limitées participe à l'effort mondial en vue de relever un défi planétaire.

Encadré 13 : Fonds mondial pour l'IA : exemples d'investissements possibles

Un fonds relativement modeste pourrait contribuer à créer une infrastructure de calcul minimale commune pour la formation de modèles de petite et moyenne taille. Ces modèles présentent d'importantes possibilités pour les objectifs de développement durable, par exemple pour la formation des agriculteurs dans leur langue locale.

Cet investissement créerait également un environnement de type expérimental permettant aux développeurs d'affiner les modèles à source ouverte existants à l'aide de leurs propres données contextuelles et de haute qualité. L'accès à l'infrastructure de calcul et aux dispositifs expérimentaux pourrait se faire sur la base d'un partage du temps et de frais d'utilisation raisonnables contribuant à couvrir les coûts de maintenance et d'exploitation.

Une troisième utilisation du financement consisterait à aider à constituer des ensembles de données de référence pour certains objectifs de développement durable pour lesquels il n'y a pas d'incitation commerciale. Les travaux de mise au point de modèles, d'essais et de collecte de données pourraient être regroupés de manière stratégique dans le cadre d'une approche pratique et puissante de l'autonomisation de l'IA, liée à des résultats concrets.

Enfin, le fonds pourrait stimuler la recherche et le développement non seulement pour les applications de l'IA liées au développement contextuel et aux objectifs de développement durable, mais aussi pour l'interconnexion du calcul et des modèles, ainsi que pour de nouvelles évaluations de la gouvernance.

164 Ici, l'accent est mis sur l'utilisation du financement pour aider à garantir la création d'une capacité minimale dans les pays de différentes régions afin de comprendre les possibilités qu'offre l'IA pour le développement durable, d'adapter et de construire des modèles pour les besoins locaux et de se joindre aux efforts de collaboration internationale sur l'IA.

Gouvernance du fonds

165 Le fonds rechercherait et regrouperait les contributions en nature, y compris celles d'entités du secteur privé. La coordination des contributions financières et des ressources en nature nécessite des niveaux appropriés de contrôle et de responsabilité indépendants. Les modalités de gouvernance doivent être inclusives et les membres du conseil d'administration doivent être issus de gouvernements, du secteur privé, du milieu philanthropique, de la société civile et d'organismes des Nations Unies. Elles devraient intégrer des contributions scientifiques et techniques, canalisées (par exemple) par le groupe scientifique international proposé, et susciter la neutralité et la confiance pour la collaboration autour de l'élaboration des données et des modèles.

Exploitation du fonds

166 Le modèle d'exploitation du fonds devrait s'inspirer des enseignements tirés des collaborations internationales en matière de recherche et de développement telles que le CERN et Gavi, l'Alliance pour les vaccins, ainsi que des enseignements tirés des plateformes commerciales d'infrastructures en temps partagé. Il devrait également s'inspirer d'organismes tels que le Fonds mondial (établi en 2002 pour regrouper les ressources visant à lutter contre le VIH, la tuberculose et le paludisme)⁴³ et le Complex Risk Analytics Fund (qui mettent à disposition des données pour soutenir toutes les parties prenantes dans l'anticipation, la prévention et la réponse aux crises).

Cadre mondial de données sur l'IA

Recommandation 6 : cadre mondial de données sur l'IA

Nous recommandons la création d'un cadre mondial de données sur l'IA, qui s'inscrirait dans un processus lancé par un organisme compétent tel que la CNUDCI, sur la base des travaux menés par d'autres organisations internationales, dans les buts suivants :

- a. Élaborer des définitions et des principes relatifs aux données pour la gouvernance mondiale des données d'entraînement de l'IA, notamment à partir des meilleures pratiques existantes, et promouvoir la diversité culturelle et linguistique ;
- b. Établir des normes communes concernant la provenance et l'utilisation des données d'entraînement de l'IA pour une responsabilité transparente et fondée sur les droits dans tous les pays ;
- c. Instaurer des dispositifs de gestion et d'échange de données qui façonnent le marché afin de faire prospérer des écosystèmes d'IA locaux à l'échelle mondiale, tels que :
 - i. Des référentiels de données ;
 - ii. Des places de marché mondiales bien gérées pour l'échange de données anonymes destinées à l'entraînement de modèles d'IA ;
 - iii. Des accords types visant à faciliter l'accès international aux données et l'interopérabilité mondiale, éventuellement en tant que protocoles technicojuridiques du cadre.

167 Lors de nos consultations, nous avons entendu que si de nombreuses propositions ont été faites pour promouvoir un accès plus large aux données et aux accords d'échange de données afin de créer des écosystèmes d'IA plus diversifiés, peu d'entre elles se sont concrétisées jusqu'à présent. Il s'agit d'une lacune majeure dans la mise au point d'écosystèmes d'IA inclusifs et dynamiques.

43 Voir <https://www.theglobalfund.org/fr/about-the-global-fund>.

- 168** Une partie de la réponse réside dans la transparence des caractéristiques, notamment culturelles et linguistiques, des données d'entraînement de l'IA. Il est également utile de recenser les données sous-représentées ou « manquantes ». Dans le même ordre d'idées, la promotion des « communs de données » encourage la conservation des données d'entraînement pour de multiples acteurs. De telles initiatives pourraient créer des bonnes pratiques en démontrant comment la conception peut intégrer des cadres technicojuridiques pour la vie privée, la protection des données, l'interopérabilité, l'utilisation équitable des données et les droits humains.
- 169** Les marchés de données pour l'IA sont aujourd'hui une sorte de Far West. L'un des principes de fonctionnement semble être de « saisir ce que l'on peut et de le dissimuler dans des algorithmes opaques » ; une autre consiste en des accords contractuels exclusifs pour l'accès à des données protégées, applicables dans certains pays. Ces relations exclusives sont à l'origine des préoccupations de l'autorité du Royaume-Uni chargée de la concurrence (Competition Market Authority), qui craint que le secteur [des modèles de pointe] ne se développe d'une manière qui risque d'avoir des conséquences négatives sur le marché⁴⁴.
- 170** Nous considérons donc qu'il est vital de lancer un processus mondial auquel participe une variété d'acteurs, notamment des nations à différents niveaux de développement, soutenues par des organisations internationales pertinentes du système des Nations Unies et au-delà (OCDE, OMPI et OMC), afin de créer des « garde-fous » et des « voies communes » pour des écosystèmes de données d'entraînement de l'IA dynamiques. Les résultats de ce processus ne doivent pas nécessairement être des lois contraignantes, mais des contrats types et des accords technicojuridiques. Ces dispositions facilitatrices peuvent être élaborées une à une, en tant que protocoles à un cadre de principes et de définitions.
- 171** Bien que les détails ne soient pas de notre ressort, les principes clés d'un cadre mondial de données sur l'IA incluraient l'interopérabilité, la gestion, la préservation de la vie privée, l'autonomisation et le renforcement des droits et l'activation de l'écosystème de l'IA.
- 172** Nous sommes conscients que la politique antitrust et la politique de concurrence restent du ressort des autorités nationales et régionales. Toutefois, une action collective internationale peut faciliter l'accès transfrontalier aux données d'entraînement pour les jeunes pousses locales de l'IA qui ne sont pas disponibles au niveau national.
- 173** L'ONU est particulièrement bien placée pour soutenir l'établissement de principes mondiaux et de dispositions pratiques régissant la gouvernance et l'utilisation des données d'entraînement de l'IA en s'appuyant sur des années de travail de la communauté des données et en les intégrant aux développements récents en matière d'éthique et de gouvernance de l'IA. Cela est comparable aux travaux menés par la CNUDCI sur le commerce international, notamment sur les cadres juridiques et non juridiques transfrontaliers, et la facilitation du commerce et de l'investissement numériques au moyen de lois types sur le commerce électronique, l'informatique en nuage et la gestion de l'identité.
- 174** De même, la Commission de la science et de la technique au service du développement et la Commission de statistique ont inscrit à leur ordre du jour l'examen des données pour le développement et les objectifs de développement durable. L'OMPI traite également d'importantes questions relatives au contenu, aux droits d'auteur et à la protection des connaissances autochtones et de l'expression culturelle.

- 175** Le cadre proposé ici serait sans préjudice des cadres nationaux ou régionaux de protection des données et ne créerait pas de nouveaux droits liés aux données ni ne prescrirait la manière dont les droits existants s'appliquent au niveau international, mais il devrait être conçu de manière à empêcher l'appropriation par des intérêts, notamment commerciaux, qui pourraient saper ou empêcher la protection des droits. En revanche, un cadre mondial pour les données d'IA permettrait de traiter les questions transversales de la disponibilité, de l'interopérabilité et de l'utilisation des données d'entraînement de l'IA. Il contribuerait à élaborer une vision commune de la manière d'harmoniser les différents cadres nationaux et régionaux de protection des données.
- 176** Les mesures prises pour résoudre ces problèmes aux niveaux national et régional sont prometteuses, les secteurs public et privé accordant davantage d'attention aux meilleures pratiques. Pourtant, en l'absence d'un cadre mondial régissant les ensembles de données d'entraînement de l'IA, la concurrence commerciale invite à un nivellement par le bas entre les pays en ce qui concerne les exigences d'accès et d'utilisation, ce qui rend difficile la gouvernance de la chaîne de valeur de l'IA au niveau international. Seule une action collective mondiale peut promouvoir un nivellement par le haut dans la gouvernance de la collecte, de la création, de l'utilisation et de la monétisation des données d'entraînement de l'IA, de manière à favoriser l'interopérabilité, la gestion, la préservation de la vie privée, la responsabilisation et le renforcement des droits.
- 177** De même, une telle action est nécessaire pour promouvoir des écosystèmes locaux d'IA dynamiques et limiter la poursuite de la concentration économique. Ces mesures pourraient être complétées par la promotion des communs de données et par des dispositions relatives à l'hébergement de référentiels de données dans des domaines en rapport avec les objectifs de développement durable (voir encadré 14). L'élaboration de ces modèles, ainsi que le stockage et l'analyse des données détenues en communs ou en fiducies, pourraient être soutenus par le réseau de développement des capacités et le fonds mondial pour l'IA.

Encadré 14 : Sécurisation des données pour l'entraînement des modèles d'IA : autonomisation des données, référentiels de données et accords de flux de données transfrontaliers

Dans de nombreuses circonstances, les données doivent être protégées (notamment pour des raisons de protection de la vie privée, de confidentialité commerciale, de propriété intellectuelle, de sûreté et de sécurité), mais leur mise à disposition pour l'entraînement de modèles d'IA présenterait également des avantages pour les personnes et la société.

En droit, les droits relatifs aux données permettent généralement d'empêcher des actions en rapport avec les données. Les droits relatifs à la confidentialité des données sont également personnels. La constitution des droits relatifs aux données peut rendre difficile l'exercice de ces droits d'une manière souple qui permette d'utiliser les données à certaines fins sans perdre les droits, et de le faire collectivement en tant que groupe. Même lorsqu'il est possible de contrôler les autorisations de manière souple et positive, cela nécessite généralement plus de temps, de compétences techniques et de confiance que la plupart des personnes et des organisations n'en possèdent.

Les dispositifs qui permettent aux propriétaires et aux personnes concernées par les données d'autoriser une utilisation sûre et limitée de leurs données, tout en préservant leurs droits, peuvent être décrits comme des moyens d'autonomisation à l'égard des données. Cette autonomisation peut amener beaucoup plus de personnes et de groupes de la société à devenir des partenaires et des acteurs actifs de l'IA, et pas seulement des sujets de données. Des outils sont déjà en cours d'élaboration afin de gérer l'accès en toute sécurité, notamment des référentiels de données et des applications de protection de la vie privée pour piloter les flux de données transfrontaliers.

Les référentiels de données sont des dispositifs qui permettent aux personnes aux organisations de donner accès à leurs données collectivement, l'accès étant contrôlé par des fiduciaires. Les propriétaires des données peuvent fixer les conditions d'accès, d'utilisation et de finalité, que les fiduciaires exercent. Les propriétaires et les sujets des données conservent leurs droits juridiques tout en contribuant à des objectifs communs. On peut s'attendre à ce qu'un modèle d'IA entraîné à partir de ces données soit plus précis qu'un modèle ne disposant pas de ces données spécifiques, et qu'il réponde mieux au bien-être de ce groupe particulier, ou de la société en général.

Les dispositifs de gestion de l'accès et de l'utilisation, et de l'accès transfrontalier en particulier, reposent tous sur des cadres juridiques précis. L'utilisation de ces mécanismes dans la pratique nécessite également une adaptation aux besoins et aux contextes des secteurs et des populations. Il convient de recenser et de combler les lacunes dans la gestion des données. Le succès et la généralisation de ces dispositifs à l'avenir dépendront de l'assurance technique et du maintien de la confiance des fournisseurs de données.

Nous proposons donc de soutenir davantage la mise au point de ces outils et de recenser les domaines dans lesquels leur utilisation pour l'entraînement de l'IA pourrait apporter la plus grande valeur publique.

D. Stratégie cohérente

- 178** En favorisant une vision commune, un socle commun et des avantages communs, les propositions ci-dessus visent à combler les lacunes recensées dans le régime international émergent de gouvernance de l'IA. Des partenariats et une collaboration avec les organismes et les dispositifs existants peuvent permettre de combler les lacunes en matière de représentation, de coordination et de mise en œuvre.
- 179** Toutefois sans un point focal dédié au sein de l'ONU pour soutenir et permettre une coordination souple entre les travaux menés, et garantir que le système des Nations Unies parle d'une seule voix sur l'IA, le monde ne disposera pas de l'approche inclusive en réseau, agile et cohérente nécessaire à une gouvernance efficace et équitable de l'IA.
- 180** C'est pourquoi nous proposons la création d'une petite capacité agile, sous la forme d'un bureau de l'IA au Secrétariat de l'ONU.

Bureau de l'AI au Secrétariat

Recommandation 7 : bureau de l'IA au Secrétariat

Nous recommandons la création d'un bureau de l'IA au Secrétariat, qui relèverait du Secrétaire général. Sa structure devrait être légère et agile et s'appuyer, dans la mesure du possible, sur les entités de l'ONU existantes. Agissant comme le « ciment » qui soutient et catalyse les propositions contenues dans le présent rapport, en partenariat et en interface avec d'autres processus et organismes, le bureau aurait le mandat suivant :

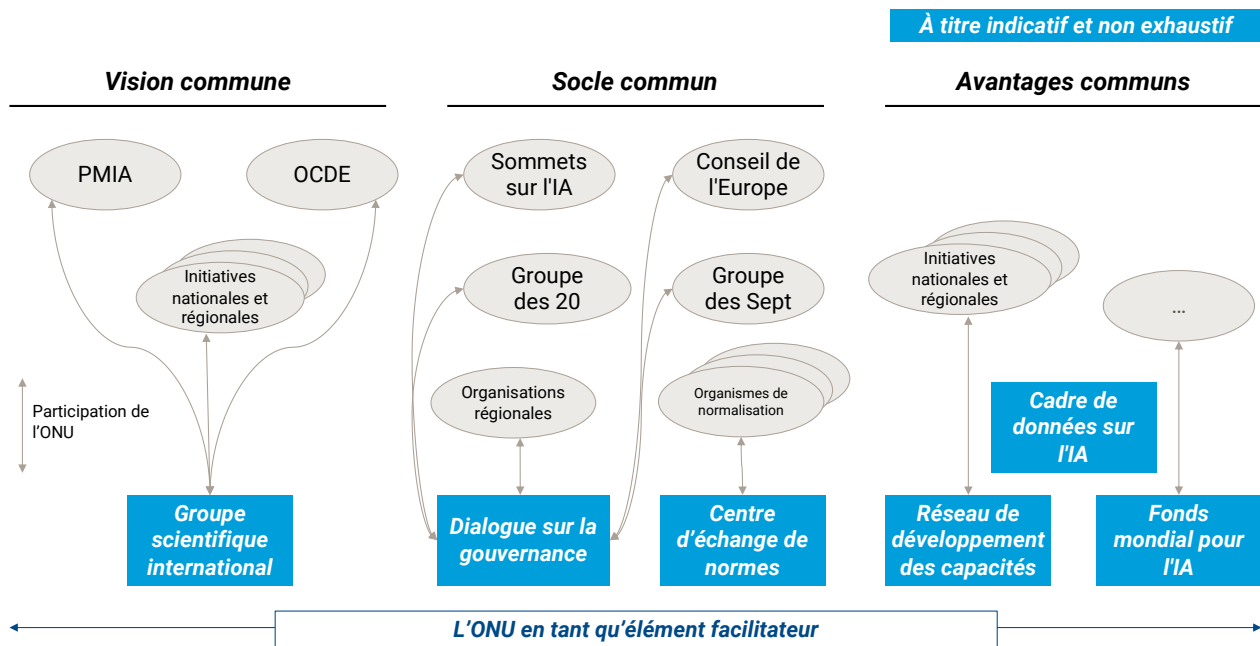
- a. Soutenir le groupe scientifique international, la concertation, l'échange de normes, le réseau de développement des capacités et, dans la mesure où cela est nécessaire, le fonds mondial et le cadre mondial de données sur l'IA ;
- b. Sensibiliser diverses parties prenantes, notamment les entreprises technologiques,

- la société civile et le milieu universitaire, aux questions qui se posent en matière d'IA ;
- c. Conseiller le Secrétaire général sur les questions liées à l'IA, en coordination avec d'autres parties concernées du système des Nations Unies, afin de fournir une réponse à l'échelle du système.

- 181** Au cours de nos consultations, il est apparu clairement que les arguments en faveur de la création d'un organisme doté de pouvoirs de notification, de contrôle, de vérification et d'exécution n'avaient pas encore convaincu et que les États Membres n'étaient pas vraiment disposés à créer une nouvelle organisation coûteuse.
- 182** Nous nous concentrons donc sur la valeur que l'ONU peut apporter, en gardant à l'esprit les lacunes du système des Nations Unies, et sur ce qui pourrait être réalisé de manière réaliste en l'espace d'un an. À cet égard, nous proposons un dispositif léger et agile qui servirait de « ciment » pour fédérer les processus qui favorisent une vision commune, un socle commun et des avantages communs, et qui permettrait au système des Nations Unies de parler d'une seule voix dans l'écosystème international de la gouvernance de l'IA en pleine évolution.
- 183** Tout comme les pays ont mis en place des instituts et des bureaux spécialisés dans la gouvernance nationale, régionale et internationale de l'IA⁴⁵, nous estimons qu'il est nécessaire de disposer d'une capacité qui serve et soutienne le groupe scientifique international sur l'IA et la concertation sur l'IA, et qui catalyse l'échange de normes sur l'IA et le réseau de développement des capacités – avec des frais généraux et des coûts de transaction moins élevés que si chacun d'entre eux était soutenu par des organisations différentes.
- 184** La création d'un bureau de l'IA au Secrétariat, qui rendrait compte au secrétaire général, aurait l'avantage d'avoir des liens avec l'ensemble du système des Nations Unies, sans être lié à une seule partie de celui-ci. Cela est important en raison de l'avenir incertain de l'IA et de la forte probabilité qu'elle s'infilte dans tous les aspects de l'activité humaine.

⁴⁵ Notamment le Canada, les États-Unis, le Japon, la République de Corée, le Royaume-Uni, Singapour et l'Union européenne.

Figure 17 : Rôle que l'ONU pourrait assurer dans l'écosystème international de gouvernance de l'IA



Abréviations : PMIA, Partenariat mondial sur l'intelligence artificielle ; OCDE, Organisation de coopération et de développement économiques ; organismes de normalisation.

185 Un bureau de l'IA, petit et agile, serait bien placé pour relier plusieurs domaines et organisations sur les questions de gouvernance de l'IA afin de contribuer à combler les lacunes de manière dynamique, en s'efforçant d'amplifier les efforts existants au sein de l'ONU et au-delà. En reliant d'autres initiatives, telles que celles menées par des organisations régionales et d'autres initiatives plurilatérales, il peut contribuer à réduire les coûts de la coopération entre elles.

186 Un organe devrait promouvoir l'inclusion et s'associer rapidement pour accélérer la coordination et la mise en œuvre, en s'appuyant en priorité sur les ressources et les fonctions existantes au sein du système des Nations Unies. Il pourrait être composé en partie de membres du personnel de l'ONU détachés par les organismes spécialisés concernés et d'autres parties du système des Nations Unies. Il devrait collaborer avec de nombreuses parties prenantes, notamment la société civile, l'industrie et le milieu universitaire, et nouer des partenariats avec des organisations de premier plan en dehors de l'ONU, telles que l'OCDE.

187 Le bureau de l'IA assurerait l'échange d'informations au sein du système des Nations Unies, qui s'exprimerait ainsi avec autorité et d'une seule voix. L'encadré 15 présente les fonctions possibles et les premiers résultats attendus d'un bureau de l'IA.

188 La recommandation est formulée sur la base d'une évaluation lucide des domaines dans lesquels l'ONU peut apporter une valeur ajoutée, notamment les domaines dans lesquels elle peut jouer un rôle moteur, combler les lacunes, contribuer à la coordination et assurer un rôle de deuxième plan, en étroite collaboration avec les initiatives en cours (voir fig. 17). Elle apporte également les avantages des dispositions institutionnelles existantes, notamment les financements prénégociés et les processus administratifs qui sont bien compris.

189 Il convient de tenir compte des caractéristiques évolutives de la technologie de l'IA. Il est fort probable que des progrès techniques modifient radicalement le paysage actuel des modèles d'IA. Un bureau de l'IA devrait être effectivement mis en place afin d'adapter les cadres de gouvernance à l'évolution du paysage et de répondre aux évolutions imprévues de la technologie de l'IA.

Encadré 15 : Fonctions possibles et résultats attendus pour la première année du bureau de l'IA

Le bureau de l'IA devrait être doté d'une structure légère et faire preuve d'agilité, susciter la confiance et fonctionner en réseau. Le cas échéant, il devrait adopter un modèle en étoile afin d'assurer la liaison avec d'autres parties du système des Nations Unies et au-delà.

Les activités de sensibilisation pourraient notamment consister à servir de relais clé dans une « architecture de coordination souple » entre les États Membres, les réseaux plurilatéraux, les organisations de la société civile, le milieu universitaire et les entreprises technologiques dans le cadre d'un régime complexe qui permet de résoudre les problèmes en collaboration grâce à la mise en réseau, et de constituer un lieu sûr et fiable pour se réunir sur des sujets pertinents. Idéalement, il pourrait servir de ciment pour consolider d'autres réseaux en constante évolution.

Afin de soutenir les diverses initiatives proposées dans le présent rapport, il convient de garantir des résultats rapides et inclusifs, tels que des rapports scientifiques, un dialogue sur la gouvernance et la désignation d'entités de suivi adéquates.

Vision commune :

- Faciliter le recrutement et soutenir le groupe scientifique international.

Socle commun :

- Assurer des concertations sur les services avec des contributions de multiples parties prenantes en vue de soutenir l'interopérabilité et l'apprentissage. L'articulation des seuils de risque et des cadres de sûreté entre les différents pays constitue une première priorité.
- Aider l'UIT, l'ISO/CEI et l'IEEE à mettre en place le centre d'échange de normes sur l'IA.

Avantages communs :

- Soutenir le réseau de développement des capacités en matière d'IA, en particulier le renforcement des capacités des fonctionnaires et des entrepreneurs sociaux en matière d'IA d'intérêt général. Définir la vision initiale du réseau, ses résultats, sa structure de gouvernance, ses partenariats et ses dispositifs opérationnels.
- Définir la vision, les résultats, la structure de gouvernance et les dispositifs opérationnels du fonds mondial pour l'IA, et demander aux États Membres, à l'industrie et aux parties prenantes de la société civile de donner leur avis sur la proposition, dans le but de financer les premiers projets dans les six mois suivant la création du Fonds.
- Préparer et publier une liste annuelle de domaines d'investissement prioritaires pour orienter le fonds mondial pour l'IA et les investissements réalisés en dehors de cette structure.

Encadré 15 : Fonctions possibles et résultats attendus pour la première année du bureau de l'IA (suite)

Stratégie cohérente :

- Mettre en place des dispositifs légers qui aident les États Membres et d'autres organisations concernées à être mieux connectés, coordonnés et efficaces dans leurs travaux sur la gouvernance mondiale de l'IA :
- Élaborer des cadres initiaux pour guider et suivre le travail du bureau de l'IA, notamment une taxonomie mondiale des risques de gouvernance, un examen du paysage politique mondial de l'IA et une carte mondiale des parties prenantes.
- Élaborer et présenter à intervalles réguliers des rapports trimestriels et des exposés en personne aux États Membres sur l'état d'avancement du plan de travail du bureau de l'IA et mettre en place des voies de retour d'information à l'appui des ajustements nécessaires.
- Créer un comité directeur dirigé conjointement par le bureau de l'IA, l'UIT, la CNUCED, l'UNESCO et d'autres entités et organismes compétents des Nations Unies afin d'accélérer les travaux des Nations Unies au service des fonctions susmentionnées, et examiner tous les trois mois l'état d'avancement des efforts accélérés.
- Promouvoir des possibilités d'apprentissage et de développement communs à l'intention des représentants des États Membres afin de les aider à assumer leurs responsabilités en matière de gouvernance mondiale de l'IA, en coopération avec les entités et organisations compétentes des Nations Unies telles que l'Institut des Nations Unies pour la formation et la recherche et l'UNU.

E. Réflexions sur les modèles institutionnels

190 Les débats sur l'IA sont souvent polarisés. Lors des consultations que nous avons menées à travers le monde, nous nous sommes entretenus avec ceux qui entrevoient un avenir de biens illimités fournis par des systèmes d'IA toujours moins chers et toujours plus utiles. Nous nous sommes également entretenus avec ceux qui craignent un avenir plus sombre, la division et le chômage, voire l'extinction.

191 Nous ne savons pas ce que l'avenir nous réserve. Nous sommes conscients que la technologie peut évoluer dans une direction qui supprime cette dualité. Le présent rapport se concentre sur les possibilités et les risques à court terme, sur la base de données scientifiques et factuelles. Les recommandations qui y sont formulées constituent notre meilleur espoir de tirer profit des avantages de l'IA, tout en réduisant au minimum et en atténuant les risques. Nous sommes également conscients des difficultés pratiques que pose le

renforcement des organismes internationaux à plus grande échelle. C'est pourquoi nous proposons une approche institutionnelle en réseau, avec un soutien léger et agile.

192 Toutefois, si ou lorsque les risques s'aggravent et que les enjeux liés aux possibilités offertes augmentent, ces calculs changeront. Les guerres mondiales ont donné naissance au système international moderne ; la mise au point d'armes toujours plus puissantes a conduit à l'établissement de régimes qui limitent leur diffusion et favorisent l'utilisation pacifique des technologies sous-jacentes.

193 L'évolution de la vision de notre humanité commune a conduit au système moderne des droits humains et à nos engagements actuels en faveur des objectifs de développement durable pour tous. Les changements climatiques ne sont plus une préoccupation de niche mais un défi mondial. De même, l'évolution de l'IA peut exiger plus de ressources et de limites que ce qui est proposé dans le présent rapport.

194 Notre mandat comprenait l'examen des fonctions, des formes et du calendrier d'un nouvel organisme international pour l'intelligence artificielle. Nous concluons le présent rapport par quelques réflexions sur la question, même si nous ne recommandons pas la création de ce type d'organisme pour le moment.

Un organisme international de l'IA ?

195 Si les risques liés à l'IA continuent de s'aggraver et de se concentrer, les États Membres pourraient être amenés à envisager la création d'un organisme international plus solide, doté de pouvoirs de surveillance, d'information, de vérification et d'exécution.

196 Il existe des précédents concernant cette évolution. Depuis les Conventions de La Haye de 1899 et 1907, puis le protocole de Genève de 1925 et la Convention sur les armes chimiques, les produits chimiques à double usage sont depuis longtemps soumis à des limites d'accès, à des protocoles de stockage et d'utilisation, ainsi qu'à une interdiction de la militarisation.

197 Les armes biologiques ont également été interdites, de même que les limites périodiques imposées à la recherche, telles que celles imposées en 1975 à l'ADN recombinant ou à l'épissage des gènes. Ils ont mis l'accent sur la limitation en tant que considération essentielle dans la conception de l'expérience, le niveau de limitation étant lié au risque estimé. Certaines classes d'expériences à haut risque pour lesquelles le confinement ne pouvait être garanti ont été essentiellement interdites. D'autres exemples comprenaient la recherche qui menace de franchir des limites éthiques fondamentales, telles que les restrictions actuelles sur le clonage humain, qui est un exemple du type de « ligne rouge » qui pourrait un jour être nécessaire dans le contexte de la recherche sur l'IA, ainsi qu'une coopération efficace en ce qui concerne l'application.

198 Certains de ces cadres prévoient également des évaluations scientifiques continues, par exemple le Conseil consultatif scientifique de l'Organisation pour l'interdiction des armes chimiques (OIAC) et l'article XII de la Convention sur les armes biologiques.

199 La comparaison entre l'IA et l'énergie nucléaire est bien connue. Dès le jour de la fission de l'atome, les scientifiques ont compris que cette technologie pouvait être utilisée à bon escient, même si leurs recherches visaient à construire une arme nouvelle et terrible. À l'époque, comme aujourd'hui, il n'est pas anodin que ceux qui réclament avec le plus d'insistance une limitation de cette nouvelle technologie comptent parmi eux d'éminents scientifiques.

200 Le Grand Bargain (pacte relatif au financement de l'action humanitaire) défendu par l'AIEA prévoyait que les avantages de l'énergie nucléaire soient mis en commun dans la production d'énergie, l'agriculture et la médecine, en échange de garanties que l'énergie nucléaire ne serait pas militarisée. Comme le montre le régime de non-prolifération nucléaire, de bonnes normes sont nécessaires mais pas suffisantes pour une réglementation efficace.

201 Les limites de l'analogie sont claires. L'énergie nucléaire implique un ensemble bien défini de processus liés à des matériaux spécifiques qui sont inégalement répartis, et une grande partie des matériaux et de l'infrastructure nécessaires pour créer une capacité nucléaire est contrôlée par des États-nations. L'IA est un terme amorphe ; ses applications sont très vastes et ses capacités les plus puissantes s'étendent à l'industrie et aux États. Le Grand Bargain de l'AIEA s'est concentré sur les armes coûteuses à fabriquer et difficiles à cacher ; la militarisation de l'IA ne sera ni l'un ni l'autre.

202 Une première idée – la mise en commun du combustible nucléaire à des fins pacifiques – n'a pas fonctionné comme prévu. S'agissant de la mise en commun des ressources afin de partager les bénéfices de la technologies, une analogie plus adaptée à l'IA pourrait être le CERN, qui met en commun les financements, les talents et les

infrastructures. Toutefois, il y a des limites à la comparaison étant donné la différence entre la physique fondamentale expérimentale et l'IA, qui nécessite une approche plus globale.

203 Une autre analogie imparfaite est celle d'organisations telles que l'OACI et l'Organisation maritime internationale (OMI). Les technologies sous-jacentes du transport sont bien établies et leurs applications civiles peuvent être facilement distinguées des applications militaires, ce qui n'est pas le cas de l'IA à usage général. Le réseau des autorités réglementaires nationales qui appliquent les normes internationales élaborées dans le cadre de l'OACI et de l'OMI est également bien établi. La sûreté, la facilitation de l'activité commerciale et l'interopérabilité sont au centre des préoccupations. La conformité n'est pas traitée de manière descendante.

204 D'autres approches de la conformité peuvent servir d'exemples. La gestion des risques financiers tire profit de dispositifs tels que le Conseil de stabilité financière et le Groupe d'action financière (GAFI), sans avoir recours à des traités.

205 À terme, un dispositif au niveau mondial pourrait s'avérer essentiel pour formaliser les lignes rouges si la réglementation de l'IA doit être opposable. Il pourrait inclure des engagements formels tels que ceux du CERN pour la mise en commun des ressources en vue d'une collaboration sur la recherche en matière d'IA et la mise à disposition des bénéfices en tant que partie intégrante du marché.

206 Toutefois, compte tenu de la rapidité, de l'autonomie et de l'opacité des systèmes d'IA, attendre l'urgence d'une menace peut signifier que toute réponse arrivera trop tard. La poursuite des évaluations scientifiques et de la concertation permettrait de ne pas prendre le monde par surprise. Toute décision d'entamer un processus formel relèverait naturellement des États Membres.

207 Les seuils possibles pour une telle évolution pourraient inclure la perspective du développement de systèmes d'IA incontrôlables ou incontrôlés, ou

le déploiement de systèmes incapables de remonter jusqu'à des acteurs humains, des entreprises ou des acteurs étatiques. Ils pourraient également inclure des indications selon lesquelles les systèmes d'IA présentent des qualités suggérant l'apparition d'une « superintelligence », bien que celle-ci ne soit pas présente dans les systèmes d'IA d'aujourd'hui.

208 La mise en place d'un mandat de surveillance, faisant appel à des experts divers et éminents pour surveiller l'horizon, est une première étape raisonnable. Le groupe scientifique pourrait être chargé de mener des travaux des recherches sur cette question, dans le cadre de sa série Quarterly Research Digest. Au fil du temps, la concertation pourrait constituer une enceinte utile pour l'échange d'informations sur les problèmes liés à l'IA, tels que ceux qui sollicitent ou dépassent les capacités des organismes existants, par analogie avec les pratiques de l'AIEA en matière de réassurance mutuelle sur la sûreté et la sécurité nucléaires, ou de l'OMS en matière de surveillance des maladies.

209 Les fonctions de l'organisme international de l'IA proposé pourraient s'appuyer sur l'expérience des organismes concernés, telles que l'AIEA, l'OIAC, l'OMI, le CERN et la Convention sur les armes biologiques, notamment :

- Élaborer et promulguer des normes et des critères de sûreté de l'IA ;
- Surveiller les systèmes d'IA susceptibles de menacer la paix et la sécurité internationales ou de provoquer de graves violations des droits humains ou du droit international humanitaire ;
- Recevoir et examiner les rapports de problèmes ou d'abus, et signaler les infractions graves ;
- Vérifier le respect des obligations internationales ;
- Coordonner la responsabilité, les interventions d'urgence et les mesures correctives en cas de problèmes liés à la sûreté de l'IA ;
- Promouvoir la coopération internationale pour l'utilisation pacifique de l'IA.

210 Une approche personnalisée de la conception de tout futur organisme de l'IA serait nécessaire, basée le cas échéant sur des enseignements tirés d'autres organismes (voir encadré 16).

Encadré 16 : Enseignements tirés d'anciens organismes de gouvernance mondiale

L'IA est un ensemble unique de technologies dont les risques et les effets sociétaux dépassent les frontières. Toutefois, ce n'est pas la première série de technologies qui a donné lieu à des accords de gouvernance de l'IA à l'échelle mondiale. L'aviation civile, les changements climatiques, l'énergie nucléaire et le financement du terrorisme sont également des domaines complexes et multidimensionnels qui ont justifié une action mondiale.

Certains de ces domaines, comme l'aviation civile, les changements climatiques et l'énergie nucléaire, ont conduit à la création de nouveaux organismes des Nations Unies. D'autres, notamment la protection des flux financiers mondiaux, ont donné naissance à des organismes qui ne sont pas fondés sur des traités et qui ont pourtant fourni des cadres normatifs solides, des dispositifs d'application efficaces fondés sur le marché et des partenariats public-privé efficaces.

En établissant des parallèles entre ces mesures institutionnelles et des travaux initiaux pour faire de même avec l'IA, nous ne devrions pas trop nous concentrer sur le modèle analogue institutionnel le mieux adapté à l'ensemble des problèmes de l'IA. Comme annoncé dans notre rapport d'étape, nous devrions plutôt nous pencher sur les fonctions de gouvernance nécessaires à une gouvernance mondiale de l'IA efficace et inclusive, et sur les enseignements que nous pouvons tirer des efforts passés en matière de gouvernance mondiale.

Le premier enseignement à retenir est qu'il convient d'élaborer une vision scientifique et technique commune du problème afin de déclencher une stratégie communément acceptée. Le GIEC, qui continue de s'intéresser aux risques des changements climatiques, est un modèle utile à cet égard. Il montre qu'une approche inclusive de la rédaction de rapports et de la recherche d'un consensus scientifique dans un domaine en constante évolution peut uniformiser les règles du jeu pour les chercheurs et les décideurs et créer la vision commune qui est essentielle à l'élaboration de politiques efficaces. Le processus de rédaction et de diffusion des rapports du GIEC et des bilans mondiaux, bien qu'il n'ait pas été exempt de difficultés, a joué un rôle essentiel dans la construction d'une vision partagée et d'une base de connaissances commune, la réduction des coûts de coopération et l'orientation de la Conférence des Parties à la Convention-cadre des Nations Unies sur les changements climatiques vers des résultats stratégiques concrets.

S'agissant de l'IA, à mesure que la technologie évolue, il sera tout aussi important d'élaborer une vision scientifique commune. Étant donné que les capacités des systèmes d'IA continuent de progresser et que les risques susceptibles de se poser peuvent dépasser les approches efficaces connues pour les atténuer, le groupe scientifique international pourrait être modifié de manière à répondre aux nouveaux besoins.

Un deuxième enseignement à retenir est que la collaboration entre les différentes parties prenantes peut permettre d'établir des normes strictes et de promouvoir des mesures rapides. L'OACI et le GAFI fournissent ici des exemples utiles sur la manière de gérer une question hautement technique au niveau transfrontalier. Dans le domaine de l'aviation civile, les normes de sûreté et de sécurité de l'OACI, élaborées par des experts de l'industrie et des gouvernements et appliquées au moyen de restrictions d'accès au marché, garantissent qu'un avion décollant, par exemple, de New York peut atterrir à Genève sans déclencher de nouveaux audits de sécurité. Une combinaison d'audits de sécurité menés par l'OACI et d'audits menés par les États Membres garantit une mise en œuvre cohérente, même si la technologie évolue.

Créé par le G7 en 1989 en vue de lutter contre le blanchiment d'argent, le GAFI montre également que les institutions de droit non contraignant peuvent promouvoir des normes et une mise en œuvre communes. Il est doté d'un système d'examen par les pairs souple en matière de suivi ; et l'acceptation généralisée de ses recommandations a terni la réputation des entreprises et des États Membres qui ne s'y conformaient pas. Même si les risques pesant sur les flux financiers internationaux ont évolué, notamment avec la montée du terrorisme et du financement de la prolifération, la structure souple et le cadre normatif du GAFI lui ont permis de réagir rapidement et de faire face à des défis complexes.

Encadré 16 : Enseignements tirés d'anciens organismes de gouvernance mondiale (suite)

L'OACI et le GAFI ont, chacun à leur manière, créé des normes internationales largement reconnues, des cadres nationaux pour mesurer la conformité et des systèmes interopérables pour répondre à certaines catégories de risques et de défis qui se manifestent d'un pays à l'autre. L'OACI impose des incitations et des restrictions à l'accès au marché, tandis que le GAFI crée un risque de perte de réputation en cas de non-conformité. Tous deux offrent des modèles utiles pour l'IA car ils montrent comment les gouvernements et d'autres parties prenantes peuvent collaborer pour créer un réseau de normes et de réglementations interconnectées et causer des désagréments en cas de non-conformité.

Un troisième enseignement à retenir est que la coordination mondiale est souvent vitale pour surveiller et prendre des mesures face à des risques graves susceptibles d'avoir de larges effets. Les modèles du Conseil de stabilité financière et de l'AIEA constituent des exemples clés. Établi en 2009, le Conseil a été créé par le G20 afin de surveiller et de prévenir les risques systémiques du système financier international. Sa composition unique de fonctionnaires des finances du G20 et d'organisations financières et de développement internationales lui a permis de faire preuve de souplesse, d'habileté et d'ouverture dans la coordination des travaux visant à détecter les risques financiers mondiaux.

L'approche de l'AIEA en matière de garanties nucléaires fournit un modèle différent. Ses accords de garanties généralisées, signés par 182 États, font partie du régime le plus étendu des Nations Unies pour garantir le respect des règles. En combinant les inspections et la surveillance, et en évoquant la menace d'une action du Conseil de sécurité, l'AIEA fournit peut-être le moyen le plus visible de censurer les États Membres qui ne respectent pas les règles.

Le Conseil de stabilité financière et l'AIEA démontrent tous deux que la coordination internationale est essentielle pour surveiller les risques graves. À mesure que les risques liés à l'IA se précisent et s'accroissent, il pourrait s'avérer nécessaire de créer un nouvel organisme axé sur l'IA afin de maximiser les efforts de coordination et de surveiller les risques graves et systémiques, de sorte que les États Membres puissent, dans la mesure du possible, prévenir ces risques.

Un quatrième enseignement à retenir est qu'il est important de créer un accès inclusif aux ressources nécessaires à la recherche et au développement, ainsi qu'aux avantages qui en découlent. Les expériences du CERN et de l'AIEA sont toutes deux instructives. Le CERN réunit des chercheurs et des physiciens de renommée mondiale chargés de mener des travaux de recherche complexes sur les accélérateurs de particules et d'autres projets destinés à servir l'humanité. Il fournit également une formation aux physiciens et aux ingénieurs.

De même, l'AIEA facilite l'accès à la technologie, en l'occurrence l'énergie nucléaire et les rayonnements ionisants. Le compromis de base est simple : les États Membres respectent les garanties nucléaires et l'AIEA fournit une assistance technique pour l'utilisation de l'énergie nucléaire à des fins pacifiques. À cet égard, l'AIEA propose une approche globale pour faire bénéficier les pays en développement des avantages de la technologie. La facilitation d'un réseau de centres d'excellence sur la sécurité nucléaire va dans le sens de notre recommandation d'une approche en réseau du renforcement des capacités.

Pour les raisons expliquées ci-dessus, l'IA est un ensemble de technologies dont les avantages doivent être mis en commun de manière plus inclusive et équitable, en particulier avec le monde du Sud. C'est pourquoi nous avons recommandé la création d'un réseau de développement des capacités en matière d'IA et d'un fonds mondial pour l'IA. À mesure que nous en apprenons davantage sur l'IA grâce aux travaux du groupe scientifique international et que le déploiement responsable de l'IA à l'appui des objectifs de développement durable devient encore plus pressant, les États Membres de l'ONU pourraient vouloir institutionnaliser cette fonction à plus grande échelle. Dans ce cas, ils devraient s'inspirer du CERN et de l'AIEA, qui constituent des modèles utiles pour favoriser un accès plus large aux ressources dans le cadre d'une structure globale de gouvernance de l'IA.

5. Conclusion : appel à l'action

- 211** En tant qu'experts, nous restons optimistes quant à l'avenir de l'IA et à ses effets positifs. Cet optimisme dépend toutefois de la réalité des risques et de l'inadéquation des structures et des incitations actuellement en place. Nous devons également être réalistes quant aux suspicions internationales qui pourraient entraver l'action collective mondiale nécessaire à une gouvernance efficace et équitable. La technologie est trop importante et les enjeux sont trop élevés pour que l'on s'en remette uniquement aux forces du marché et à une mosaïque fragmentée de mesures nationales et multilatérales.
- 212** Nous devons nous montrer actifs et déterminés. Au-delà de la dualité opportunité/risque, il faut faire face aux difficultés posées par un changement rapide et transversal. Les effets en aval de l'IA concerneront quasiment tout le monde. Placer la gouvernance de l'IA entre les mains de quelques développeurs ou des pays qui l'hébergent créera une situation profondément injuste où les conséquences de la mise au point, du déploiement et de l'utilisation de l'IA seront imposées à la plupart des personnes sans qu'elles aient eu leur mot à dire dans les décisions prises à cet effet.
- 213** L'attention et les débats mondiaux sur la gouvernance de l'IA qui ont eu lieu l'année dernière nous ont donné de l'espoir en tant qu'êtres humains. Il existe des divergences entre les pays et les secteurs, mais aussi une forte volonté de dialogue. La participation de divers experts, décideurs, hommes d'affaires, chercheurs et défenseurs, à travers les régions, les genres et les disciplines, nous a montré que la diversité n'était pas nécessairement synonyme de discorde, et que le dialogue pouvait déboucher sur un socle commun et une collaboration.
- 214** Nous avons parfois hésité : devrions-nous être pragmatiques et nous concentrer sur ce qui semble faisable ? Ou viser haut avec de grandes ambitions ? Finalement, nous avons décidé de faire les deux. Nos propositions traduisent une vision globale d'un régime mondial équitable et efficace de gouvernance de l'IA, avec une réflexion approfondie sur la manière dont il peut être mis en œuvre, étape par étape.
- 215** Nous remercions les nombreuses personnes, organisations et États Membres qui ont contribué à nos délibérations, notamment les représentants des organismes des Nations Unies et le personnel du Secrétariat qui ont fourni des évaluations perspicaces des capacités et des limites de l'ONU dans ce domaine complexe. La question de la gouvernance de l'IA ne se limite pas à la gestion des retombées de la technologie. C'est l'avenir de la coopération multilatérale et multipartite qui est également en jeu.
- 216** Dans cinq ans, le paysage technologique pourrait être radicalement différent de ce qu'il est aujourd'hui. Toutefois, si nous maintenons le cap et surmontons les hésitations et les doutes, nous pourrions, dans cinq ans, contempler un paysage de la gouvernance de l'IA qui soit inclusif et responsabilisant pour les personnes, les collectivités et les États, où qu'ils se trouvent. En fin de compte, ce n'est pas le changement technologique en soi, mais la façon dont l'humanité y réagit qui importe.

217 Nous considérons que les fonctions et les modalités recommandées dans le présent rapport, si elles sont mises en œuvre de bonne foi, peuvent permettre de créer un régime agile et adaptable qui reste en phase avec les progrès de l'IA et aide à en récolter les fruits et à gérer les risques. Ces recommandations peuvent nous aider à repérer à temps les problèmes à résoudre et les chances à saisir, à utiliser des principes et des cadres communs pour harmoniser les initiatives internationales, à promouvoir la coopération internationale et à renforcer la capacité des personnes et des organismes à faire face au changement.

218 La mise en œuvre des recommandations formulées dans le présent rapport peut également encourager de nouvelles façons de penser – un état d'esprit de collaboration et d'apprentissage, une collaboration multipartite et un engagement public à grande échelle. L'ONU peut être le vecteur d'un nouveau contrat social pour l'IA qui garantisse l'adhésion mondiale à un régime de gouvernance qui nous protège et nous donne à toutes et à tous les moyens d'agir. Un tel contrat garantira que les chances sont équitablement accessibles et réparties et que les risques ne sont pas reportés sur les plus vulnérables – ou sur les générations futures, comme nous l'avons vu, tragiquement, avec les changements climatiques.

219 En tant que groupe et en tant que personnes représentant de nombreux domaines de compétences, organisations et régions du monde, nous nous réjouissons de poursuivre cette réflexion cruciale. En collaboration avec les nombreuses personnes que nous avons rencontrées au cours de ce voyage, et la communauté mondiale qu'elles représentent, nous espérons que le présent rapport favorisera les travaux que nous menons conjointement en vue de gouverner l'IA au bénéfice de l'humanité.

Annexes

Annexe A : membres de l'Organe consultatif de haut niveau sur l'intelligence artificielle

Anna Abramova

Omar Sultan Al Olama

Latifa Al-Abdulkarim

Estela Aranha

Carme Artigas (Coprésidente)

Ran Balicer

Paolo Benanti

Abeba Birhane

Ian Bremmer (Corapporteur)

Anna Christmann

Natasha Crampton

Nighat Dad

Vilas Dhar

Virginia Dignum

Arisa Ema

Mohamed Farahat

Amandeep Singh Gill

Wendy Hall

Rahaf Harfoush

Ruimin He

Hiroaki Kitano

Haksoo Ko

Andreas Krause

James Manyika (Coprésidente)

Maria Vanina Martinez Posse

Seydina Moussa Ndiaye

Mira Murati

Petri Myllymäki

Alondra Nelson

Nazneen Rajani

Craig Ramlal

Emma Ruttkamp-Bloem

Marietje Schaake (Corapporteur)

Sharad Sharma

Jaan Tallinn

Philip Thigo

Jimena Sofia Viveros Álvarez

Zeng Yi

Zhang Linghan

Annexe B : mandat de l'Organe consultatif de haut niveau sur l'intelligence artificielle

L'Organe consultatif de haut niveau sur l'intelligence artificielle, créé par le Secrétaire général de l'ONU, analysera la gouvernance internationale de l'intelligence artificielle et formulera des recommandations à ce sujet. Les premiers rapports de l'Organe consultatif apporteront des contributions d'experts de haut niveau et des contributions indépendantes aux débats actuellement menés aux niveaux national, régional et multilatéral.

L'Organe consultatif sera composé de 38 membres issus de gouvernements, du secteur privé, de la société civile et du milieu universitaire, ainsi que d'un membre du Secrétariat. Sa composition sera équilibrée en termes de genre, d'âge, de représentation géographique et de domaine de compétence lié aux risques et aux applications de l'intelligence artificielle. Les membres de l'Organe consultatif siègent en leur nom personnel.

L'Organe consultatif mènera une collaboration et des consultations approfondies avec les gouvernements, le secteur privé, le milieu universitaire, la société civile et les organisations internationales. Il sera agile et innovant dans ses échanges avec les processus et les plateformes existants, ainsi que dans l'exploitation des contributions de diverses parties prenantes. Il peut créer des groupes de travail ou des groupes sur des sujets particuliers.

Les membres de l'Organe consultatif sont sélectionnés par le Secrétaire général sur la base des candidatures proposées par les États Membres et d'un appel public à candidatures. L'Organe consultatif sera composé de deux coprésidents et d'un Comité exécutif. Tous les groupes de parties prenantes seront représentés au Comité exécutif.

L'Organe consultatif est créé pour une période initiale d'un an, le Secrétaire général ayant la possibilité de décider d'une prolongation. Il tiendra des réunions en personne et en ligne.

L'Organe consultatif élaborera un premier rapport, d'ici le 31 décembre 2023, pour examen par le Secrétaire général et les États Membres de l'ONU. Le premier rapport présentera une analyse des solutions possibles en matière de gouvernance internationale de l'IA.

Sur la base des commentaires reçus sur le premier rapport, l'Organe consultatif soumettra un deuxième rapport, d'ici le 31 août 2024, qui pourrait donner des recommandations détaillées sur les fonctions, la forme et le calendrier d'un nouvel organisme international chargé de la gouvernance de l'intelligence artificielle.

L'Organe consultatif doit éviter tout double emploi avec les forums et processus existants dans le cadre desquels les questions relatives à l'intelligence artificielle sont examinées. Il s'efforcera plutôt de tirer parti des plateformes et des partenaires existants, y compris les entités du système des Nations Unies, qui travaillent dans des domaines connexes. Il respectera pleinement les structures actuelles des Nations Unies ainsi que les prérogatives nationales, régionales et sectorielles en matière de gouvernance de l'intelligence artificielle.

Les travaux de l'Organe consultatif seront appuyés par un petit secrétariat basé au Bureau de l'Envoyé du Secrétaire général pour les technologies et financés par des ressources extrabudgétaires fournies par des donateurs.

Annexe C : liste des consultations tenues en 2024

Consultations	Date, 2024	Région
UNESCO Slovénie	5 janv.	Europe
Conseil scientifique consultatif du Secrétaire général	8 janv.	Monde
Présentation du rapport d'étape aux États Membres	12 janv.	Monde
Forum économique mondial de Davos	24 janv.	Europe
Réunion de hauts fonctionnaires de l'ASEAN sur le numérique	30 janv.	Asie
Sommet mondial des gouvernements	12 févr.	Moyen-Orient
Institut des algorithmes d'apprentissage de Montréal (Mila - Institut québécois d'IA)	14 févr.	Amérique du Nord
Consultation à Berlin	15 févr.	Europe
Forum euro-asiatique des technologies de l'information	20 févr.	Monde
Congrès mondial de la téléphonie mobile	26 févr.	Europe
Institut d'État des relations internationales de Moscou	28 févr.	Europe
Atelier de la Royal Society sur la gouvernance internationale de l'IA	28 févr.	Europe
Réseau de conseil en science et technologie des ministères des affaires étrangères	28 févr.	Monde
Dialogue OCDE-Union africaine sur l'IA	4 mars	Europe
Consultation à Bruxelles	5 mars	Europe
Banque mondiale, Sommet mondial du numérique	5 mars	Amérique du Nord
Science ouverte et intelligence artificielle : webinaire sur les questions éthiques	5 mars	Europe de l'Est
Dialogue de l'UNESCO sur la transformation numérique	6 mars	Europe
Union interparlementaire	6 mars	Monde
Quarante-septième session du Comité de haut niveau sur les programmes	11 mars	Monde
Sommet mondial de la jeunesse sur les droits numériques	13 mars	Amérique latine
Sommet du G7 sur l'IA à Trente (Italie)	15 mars	Europe
Réunions de lancement du réseau consultatif, 18 et 19 mars	18 mars	Monde
Soixante-huitième session de la Commission de la condition de la femme	21 mars	Amérique du Nord
Mise à jour de l'Organe consultatif aux États Membres	25 mars	Monde
Observatoire africain sur l'IA responsable	25 mars	Afrique
Conférence sur l'IA pour un avenir durable et inclusif - Agence française de développement	26 mars	Europe
Élaboration de normes mondiales : retour d'information collectif	28 mars	Afrique
Innovate Switzerland	2 avril	Europe
Visite du Bureau de l'Envoyé du Secrétaire général pour les technologies en Chine, 9-12 avril	9 avril	Asie
Forum russe sur la gouvernance d'Internet	9 avril	Europe de l'Est
Wharton Cypher Days - Finance	12 avril	Amérique du Nord
Visite dans la Silicon Valley	15 avril	Amérique du Nord
Stanford, symposium AI+Policy : a global stocktaking	16 avril	Amérique du Nord
Commission de la science et de la technique au service du développement	16 avril	Europe
G20 économie numérique, 16-18 avril, Brésil	17 avril	Amérique latine
Mise à jour de l'Organe consultatif aux États Membres	22 avril	Monde
UNU, Conférence sur l'IA à Macao, 24 et 25 avril	24 avril	Asie
Visite du Bureau de l'Envoyé du Secrétaire général pour les technologies à Bruxelles et à Paris, 25 et 26 avril	26 avril	Europe
Présentation de l'Organe consultatif au Comité consultatif national sur l'IA (États-Unis)	2 mai	Amérique du Nord
Assemblée mondiale de l'intelligence artificielle à Riyad, avec l'Organisation du monde islamique pour l'éducation, les sciences et la culture (53 pays, quatre régions)	14 mai	Moyen-Orient
L'IA au service du développement durable : la contribution du Kazakhstan au Programme 2030	20 mai	Asie
Groupe des États d'Amérique latine et des Caraïbes	21 mai	Amérique latine
Forum académique des pays du groupe des BRICS	22 mai	Monde
Session sur la gouvernance de l'IA à Séoul	23 mai	Asie
Tech Summit Asia, Singapour, 29-31 mai	29 mai	Asie
Sommet mondial « AI for Good », du 29 au 31 mai	29 mai	Europe

Annexe D : liste des études approfondies

Domaine	Date (heure d'été de New York)
Éducation	29 mars
Propriété intellectuelle et contenu	2 avril
Enfants	4 avril
Paix et sécurité (1)	12 avril
Paix et sécurité (2)	29 avril
Agriculture (session 1)	30 avril
Agriculture (session 2)	30 avril
Basé sur la foi	1er mai
Logiciels libres et orientation technologique	1er mai
Impact sur la société	3 mai
Genre	7 mai
Données	13 mai
Avenir du travail	13 mai
Normes (session 1)	14 mai
Normes (session 2)	14 mai
Paix et sécurité (3)	20 mai
Environnement	20 mai
Santé	22 mai
État de droit, droits humains et démocratie	24 mai

Annexe E : réponses à l'enquête sur les risques

À la demande de l'Organe consultatif de haut niveau sur l'intelligence artificielle, le Bureau de l'Envoyé du Secrétaire général pour les technologies a mené une enquête sur les risques liés à l'IA dans le cadre d'un tour d'horizon prospectif sur l'IA, afin de recueillir les perceptions des experts du monde entier en la matière. Les experts ont été invités à donner leur avis à titre personnel (et non au nom de leur organisme ou employeur). On leur a demandé d'indiquer dans quelle mesure ils s'attendaient à une accélération ou à un ralentissement du changement technique de l'IA et (séparément) de l'adoption et de l'application de l'IA.

On a également demandé aux experts d'évaluer à quel point ils craignaient que les préjudices (existants ou nouveaux) résultant de l'IA s'aggravent ou se répandent, et dans quelle mesure cette préoccupation avait récemment augmenté ou diminué. Les répondants ont reçu une liste de 14 exemples de types de préjudice (tels que « l'utilisation malveillante intentionnelle de l'IA par des acteurs non étatiques ») afin d'évaluer leur niveau de préoccupation. Enfin, ils ont été invités, au moyen de relances, à commenter les nouvelles tendances, à préciser quelles personnes, groupes et (éco)systèmes étaient particulièrement menacés par l'IA et à expliquer les notes qu'ils avaient attribuées.

L'enquête a été menée du 13 au 25 mai 2024. La liste des invités a été établie à partir des réseaux du Bureau de l'Envoyé du Secrétaire général pour les technologies et de l'Organe consultatif, et comprenait des participants aux études approfondies de l'Organe. Au cours de la période d'enquête, des experts supplémentaires ont été invités régulièrement, en particulier en provenance de régions souvent moins représentées dans les débats sur l'IA, sur la base des références fournies par les répondants initiaux et des travaux de communication menés auprès des réseaux régionaux. Plus de 340 personnes ont répondu à l'enquête, apportant ainsi une perspective riche et diversifiée (notamment sur la base des régions et du genre) sur les risques posés par l'IA.

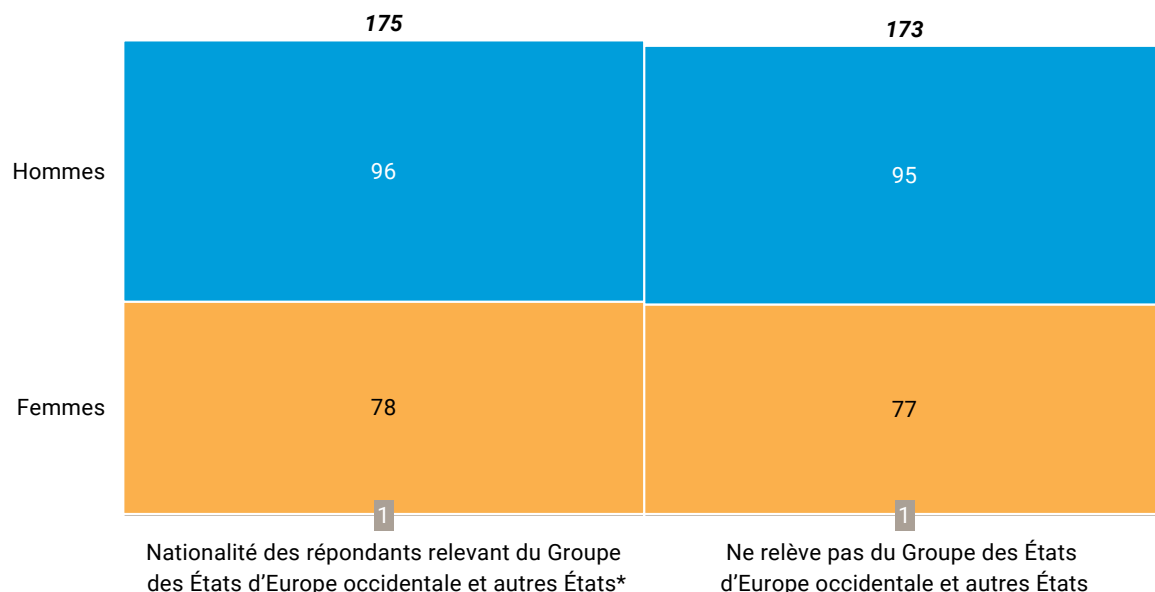
Vu d'ensemble de l'échantillon

La répartition par genre et par région est équilibrée

L'analyse univariée par genre et région n'est pas immédiatement contaminée par l'autre variable

% de répondants déclarant une affiliation/compétence par région de nationalité* (n = 348)

Par région et par genre



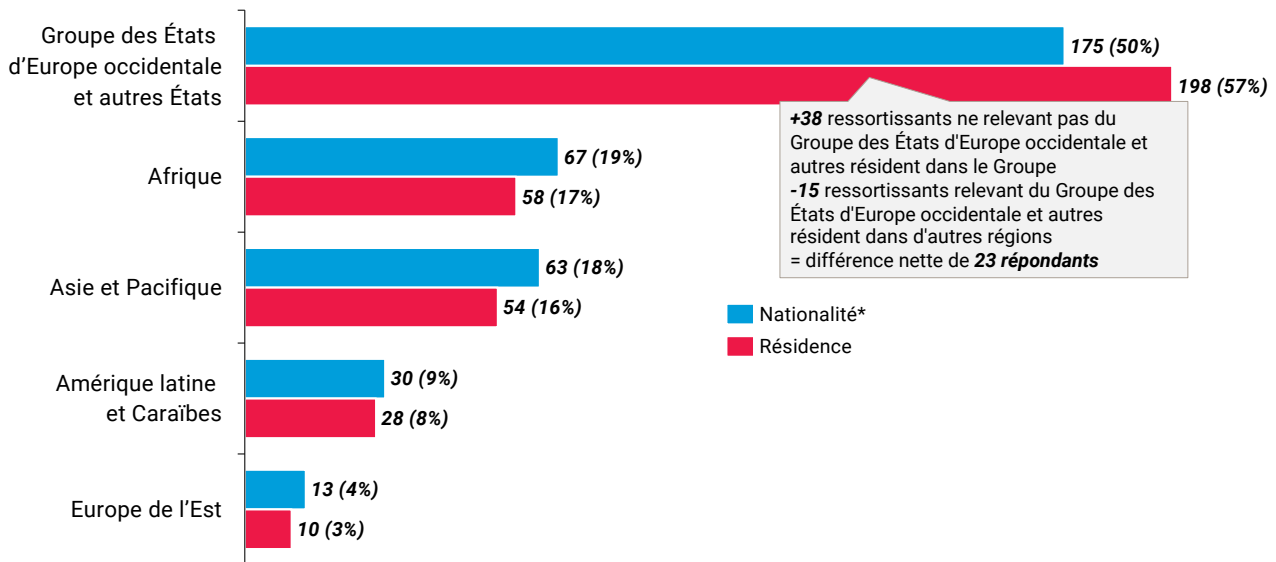
* 43 répondants (12 %) ont indiqué plusieurs nationalités. Si les répondants résidaient dans l'un de leurs pays de nationalité, c'est cette nationalité qui a été retenue pour l'analyse (34 sur 43). Dans le cas contraire, c'est la nationalité la moins représentée qui a été retenue (9 sur 43). Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

L'échantillon reste global s'il est examiné par lieu de résidence

84% des répondants résident dans la même région que leur nationalité

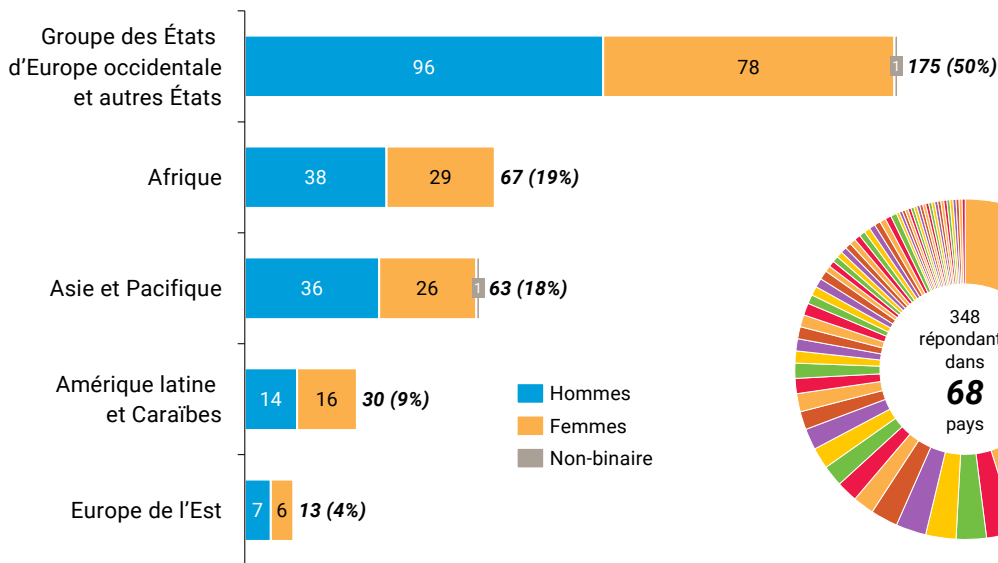
Répondants par région de nationalité* (n = 348)

Par région

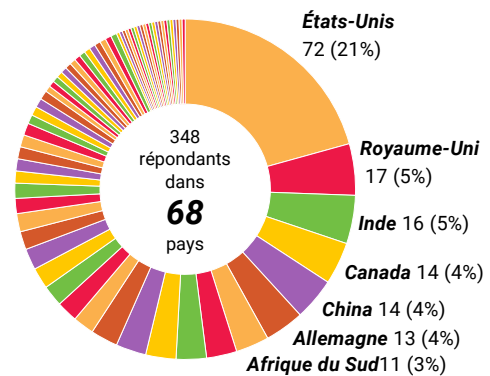


* 43 répondants (12 %) ont indiqué plusieurs nationalités. Si les répondants résidaient dans l'un de leurs pays de nationalité, c'est cette nationalité qui a été retenue pour l'analyse (34 sur 43). Dans le cas contraire, c'est la nationalité la moins représentée qui a été retenue (9 sur 43). Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

Répondants par région de nationalité* (n = 348)



Le Groupe des États d'Europe occidentale et autres États comprend l'Europe occidentale ainsi que l'Australie, le Canada, les États-Unis d'Amérique, Israël, la Nouvelle-Zélande et la Türkiye.



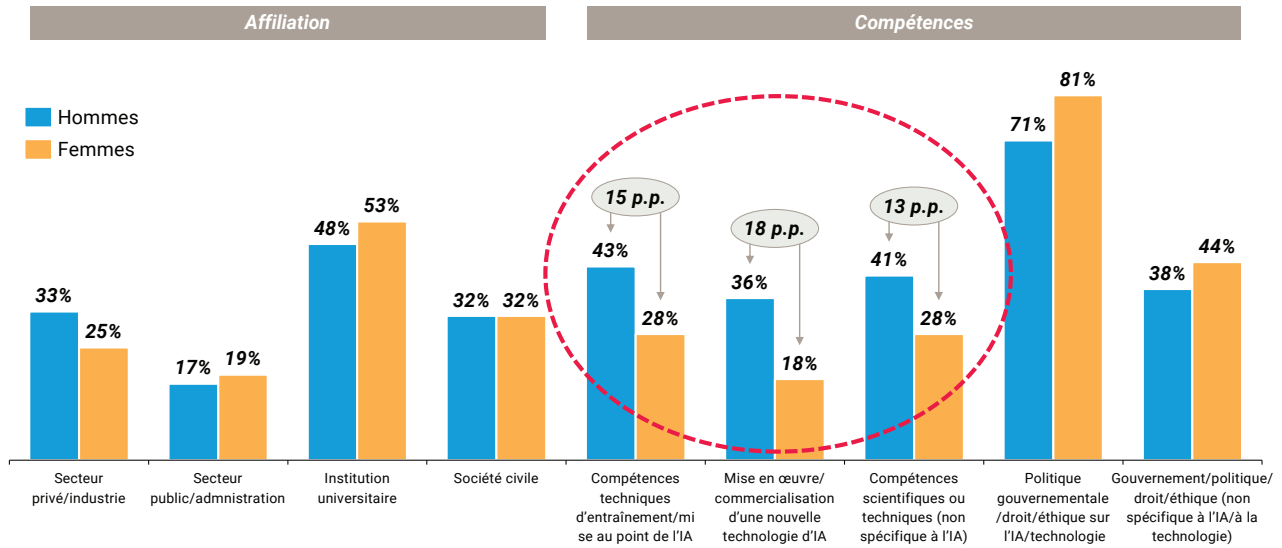
* 43 répondants (12 %) ont indiqué plusieurs nationalités. Si les répondants résidaient dans l'un de leurs pays de nationalité, c'est cette nationalité qui a été retenue pour l'analyse (34 sur 43). Dans le cas contraire, c'est la nationalité la moins représentée qui a été retenue (9 sur 43). Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

Les profils des hommes et des femmes interrogés présentent quelques différences

Les hommes sont plus nombreux à faire état d'une compétence technique ; les femmes sont plus nombreuses à mentionner la gouvernance, la politique, le droit et l'éthique

% de répondants déclarant une affiliation/compétence par genre (n = 348)

Par genre



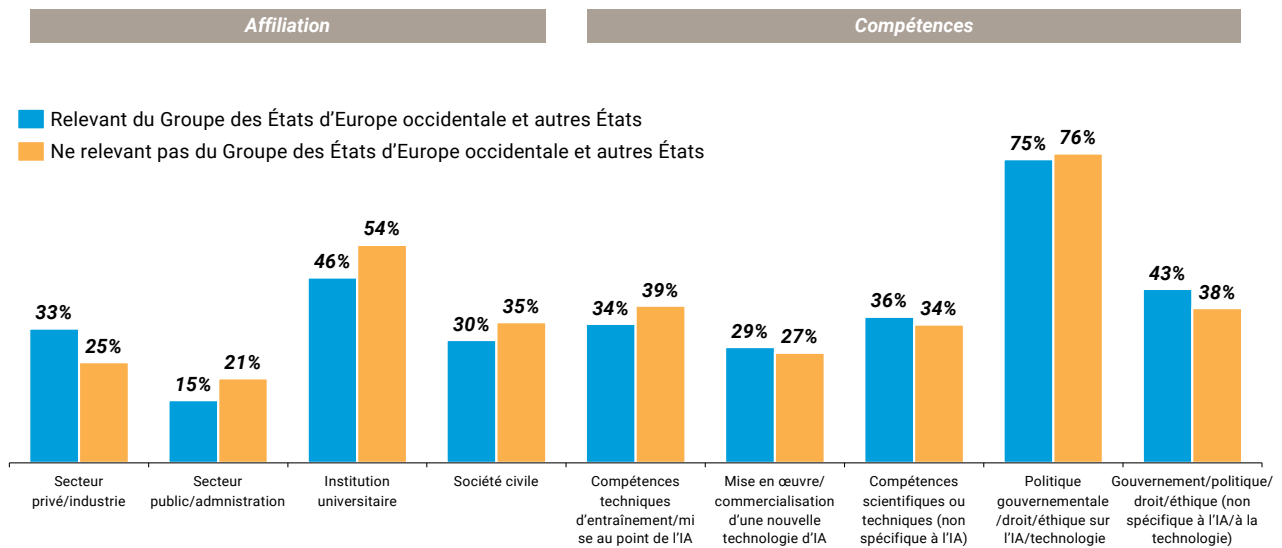
Note : p.p. = points de pourcentage. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

Les profils des répondants relevant du Groupe des États d'Europe occidentale et autres États et ceux relevant d'autres groupes sont relativement semblables

Les répondants ne relevant pas du Groupe des États d'Europe occidentale et autres États sont plus susceptibles d'être dans le secteur public ou universitaire que dans le secteur privé ou industriel

% de répondants déclarant une affiliation/compétence par région de nationalité* (n = 348)

Par région



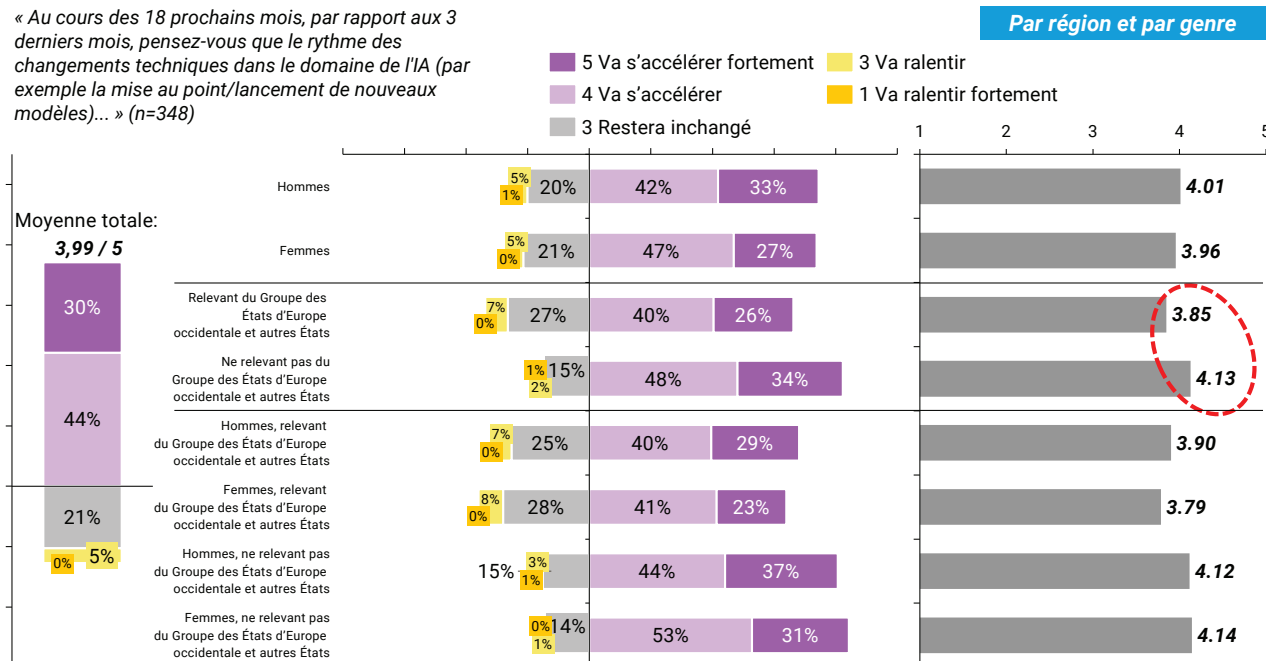
* 43 répondants (12 %) ont indiqué plusieurs nationalités. Si les répondants résidaient dans l'un de leurs pays de nationalité, c'est cette nationalité qui a été retenue pour l'analyse (34 sur 43). Dans le cas contraire, c'est la nationalité la moins représentée qui a été retenue (9 sur 43). Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

Perception de l'accélération de l'IA

74 % des répondants s'attendent à une accélération du changement technique

Un pourcentage plus élevé de répondants ne relevant pas du Groupe des États d'Europe occidentale et autres États s'attend à une accélération par rapport à ceux relevant du Groupe des États d'Europe occidentale et autres États

« Au cours des 18 prochains mois, par rapport aux 3 derniers mois, pensez-vous que le rythme des changements techniques dans le domaine de l'IA (par exemple la mise au point/lancement de nouveaux modèles)... » (n=348)

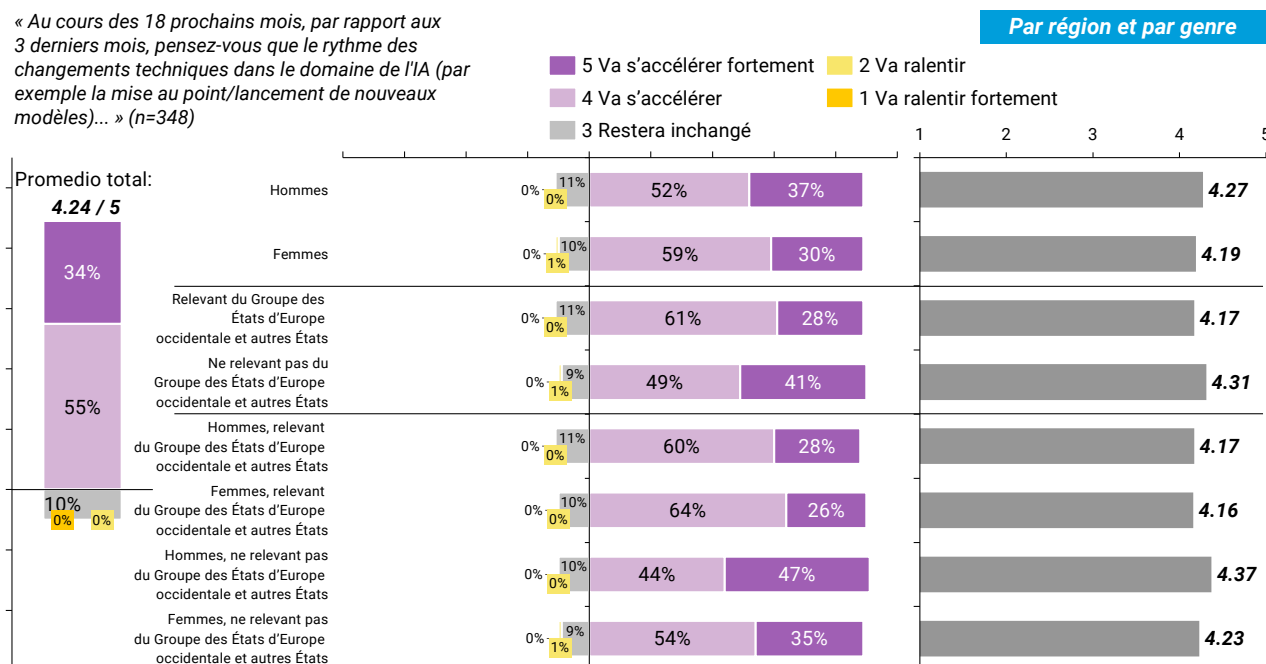


Note : Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

89 % des répondants s'attendent à une accélération de l'adoption et de l'application

Un peu plus de répondants ne relevant pas du Groupe des États d'Europe occidentale et autres États s'attendent à une forte accélération (en particulier les hommes)

« Au cours des 18 prochains mois, par rapport aux 3 derniers mois, pensez-vous que le rythme des changements techniques dans le domaine de l'IA (par exemple la mise au point/lancement de nouveaux modèles)... » (n=348)



Note : Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

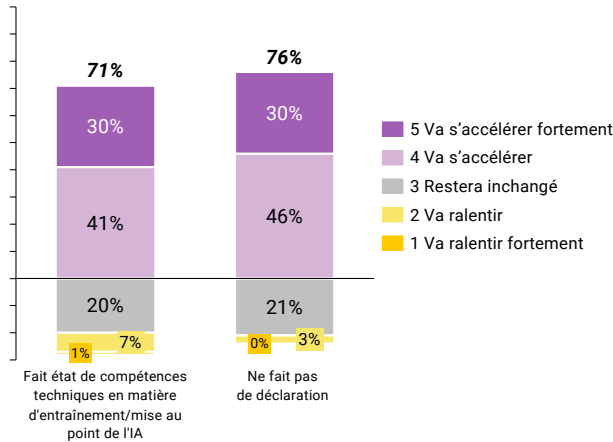
Effets limités des compétences techniques (entraînement/mise au point de l'IA)

Les répondants sont légèrement plus pessimistes en ce qui concerne les changements techniques et légèrement plus optimistes en ce qui concerne l'adoption et l'application

Par compétence

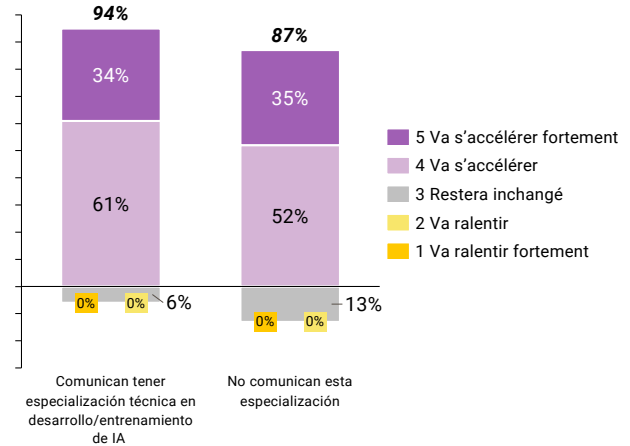
Changement technique

« Au cours des 18 prochains mois, par rapport aux 3 derniers mois, pensez-vous que le rythme des changements techniques dans le domaine de l'IA (par exemple la mise au point/lancement de nouveaux modèles)... » (n=348)



Adoption et application

« Au cours des 18 prochains mois, par rapport aux 3 derniers mois, pensez-vous que le rythme d'adoption et d'application de l'IA (par exemple, les nouvelles utilisations de l'IA dans les entreprises/l'administration)... » (n=348)



Note : les chiffres ayant été arrondis, leur somme peut ne pas correspondre à 100 %. Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

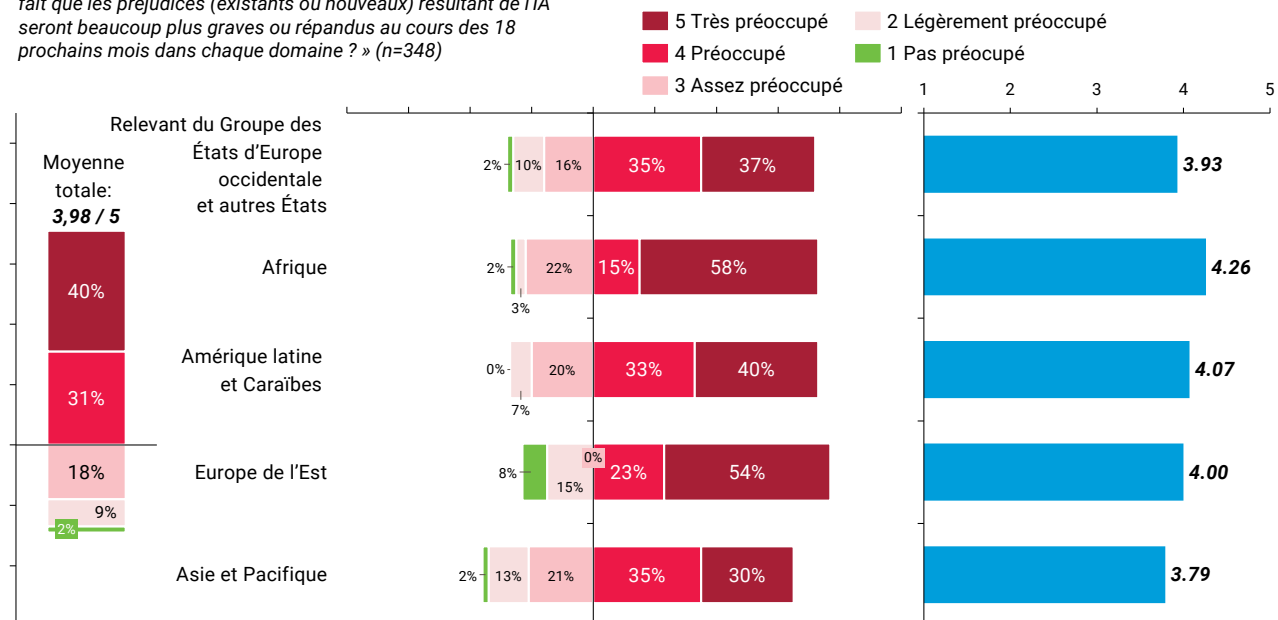
Perceptions des risques de préjudices liés à l'IA sur 18 mois (à compter de mai 2024)

71 % se disent préoccupés/très préoccupés par les préjudices pouvant résulter de l'IA dans les 18 prochains mois

Les répondants de la région d'Afrique sont plus inquiets que les autres ; les répondants de la région Asie-Pacifique sont moins préoccupés que ceux relevant du Groupe des États d'Europe occidentale et autres États

« Dans quelle mesure êtes-vous actuellement préoccupé par le fait que les préjudices (existants ou nouveaux) résultant de l'IA seront beaucoup plus graves ou répandus au cours des 18 prochains mois dans chaque domaine ? » (n=348)

Par région

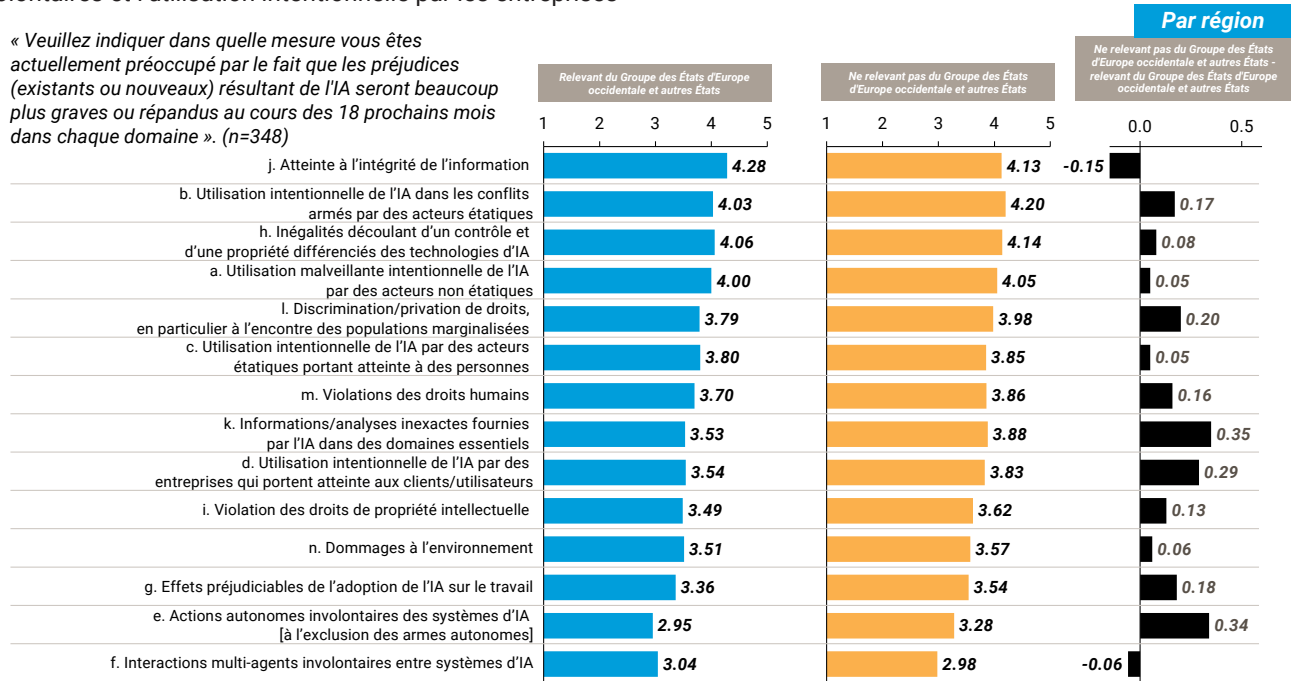


Note : les chiffres ayant été arrondis, leur somme peut ne pas correspondre à 100 %. Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

Les répondants ne relevant pas du Groupe des États d'Europe occidentale et autres États sont plus préoccupés que ceux qui relèvent du Groupe dans la plupart des domaines cités en exemple

Les lacunes sont particulièrement importantes en ce qui concerne les informations inexactes, les actes autonomes involontaires et l'utilisation intentionnelle par les entreprises

« Veuillez indiquer dans quelle mesure vous êtes actuellement préoccupé par le fait que les préjudices (existants ou nouveaux) résultant de l'IA seront beaucoup plus graves ou répandus au cours des 18 prochains mois dans chaque domaine ». (n=348)

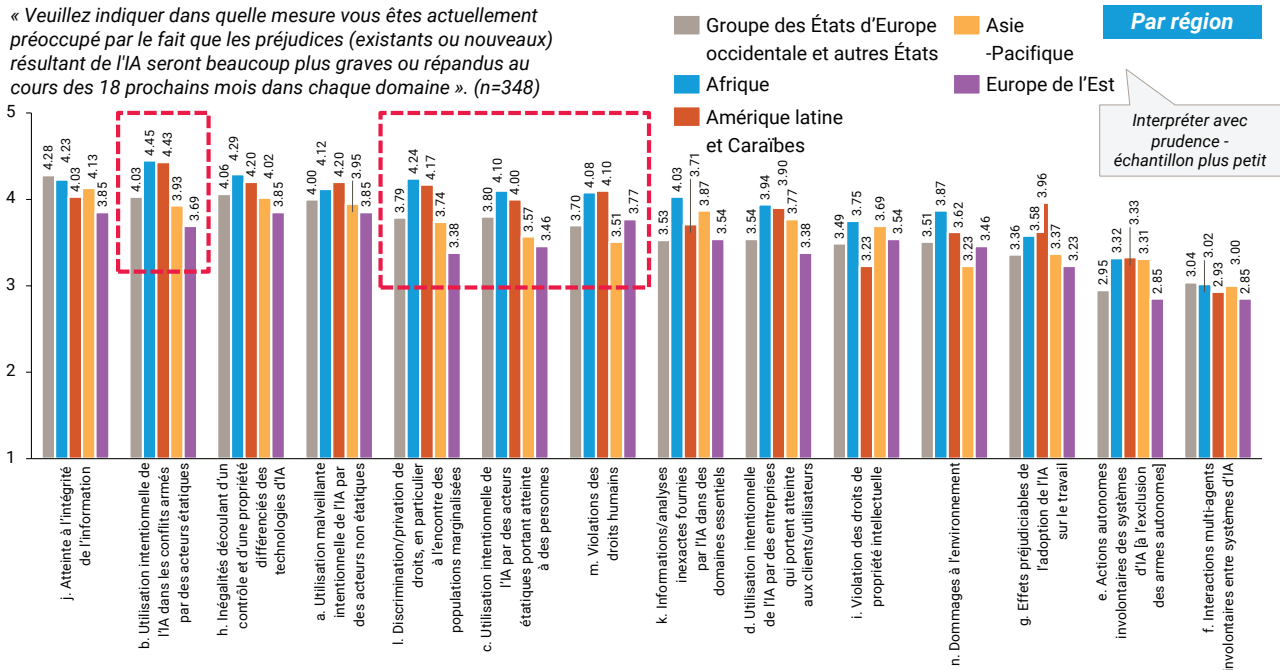


Moyenne, où : 1 = Pas préoccupé, 2 = Légèrement préoccupé, 3 = Assez préoccupé, 4 = Préoccupé, 5 = Très préoccupé. Note : Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

Les niveaux de préoccupation sont les plus élevés en Afrique, en Amérique latine et dans les Caraïbes

En particulier l'utilisation par l'État dans les conflits armés, ce qui entraîne une discrimination ou des violations des droits humains

« Veuillez indiquer dans quelle mesure vous êtes actuellement préoccupé par le fait que les préjudices (existants ou nouveaux) résultant de l'IA seront beaucoup plus graves ou répandus au cours des 18 prochains mois dans chaque domaine ». (n=348)

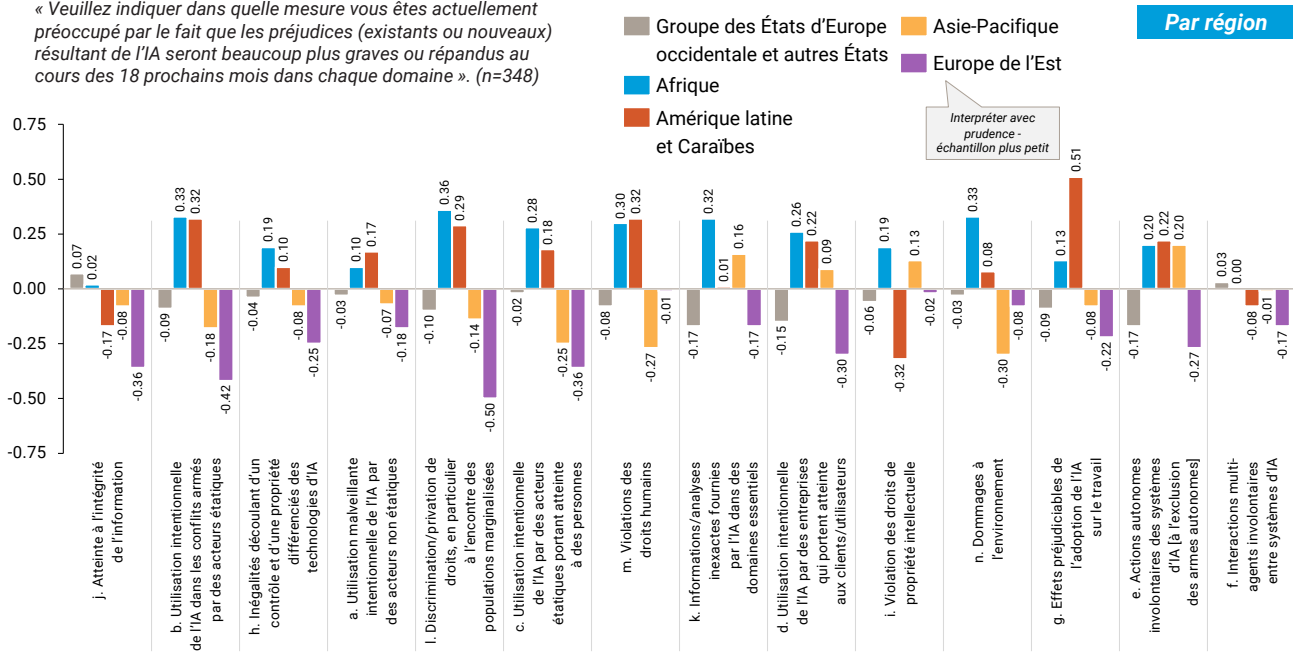


Moyenne, où : 1 = Pas préoccupé, 2 = Légèrement préoccupé, 3 = Assez préoccupé, 4 = Préoccupé, 5 = Très préoccupé. Note : Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

Les niveaux de préoccupation sont les plus élevés en Afrique, en Amérique latine et dans les Caraïbes

En particulier l'utilisation par l'État dans les conflits armés, ce qui entraîne une discrimination ou des violations des droits humains.

« Veuillez indiquer dans quelle mesure vous êtes actuellement préoccupé par le fait que les préjudices (existants ou nouveaux) résultant de l'IA seront beaucoup plus graves ou répandus au cours des 18 prochains mois dans chaque domaine ». (n=348)

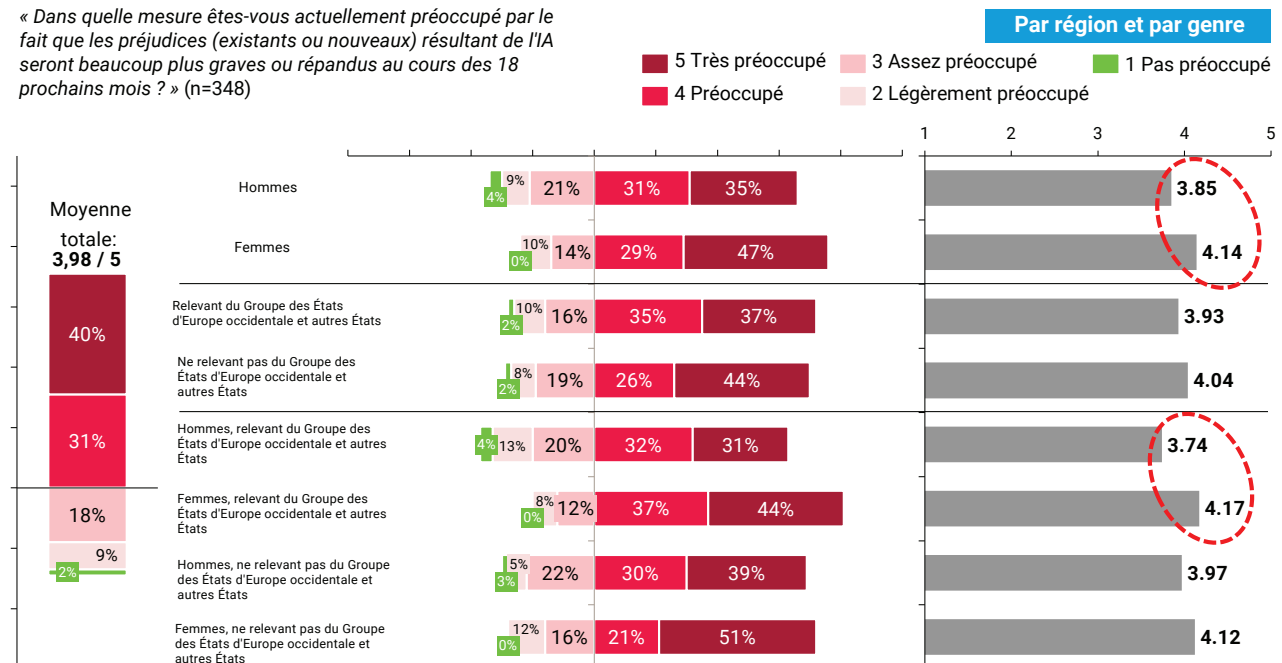


Représenté : Différence entre l'évaluation globale (tous les répondants) et l'évaluation du groupe indiqué où : 1 = Pas préoccupé, 2 = Légèrement préoccupé, 3 = Assez préoccupé, 4 = Préoccupé, 5 = Très préoccupé. Note : Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

71 % se disent préoccupés/très préoccupés par des préjudices pouvant résulter de l'IA dans les 18 prochains mois

Les femmes sont plus préoccupées que les hommes, en particulier celles relevant du Groupe des États d'Europe occidentale et autres États

« Dans quelle mesure êtes-vous actuellement préoccupé par le fait que les préjudices (existants ou nouveaux) résultant de l'IA seront beaucoup plus graves ou répandus au cours des 18 prochains mois ? » (n=348)

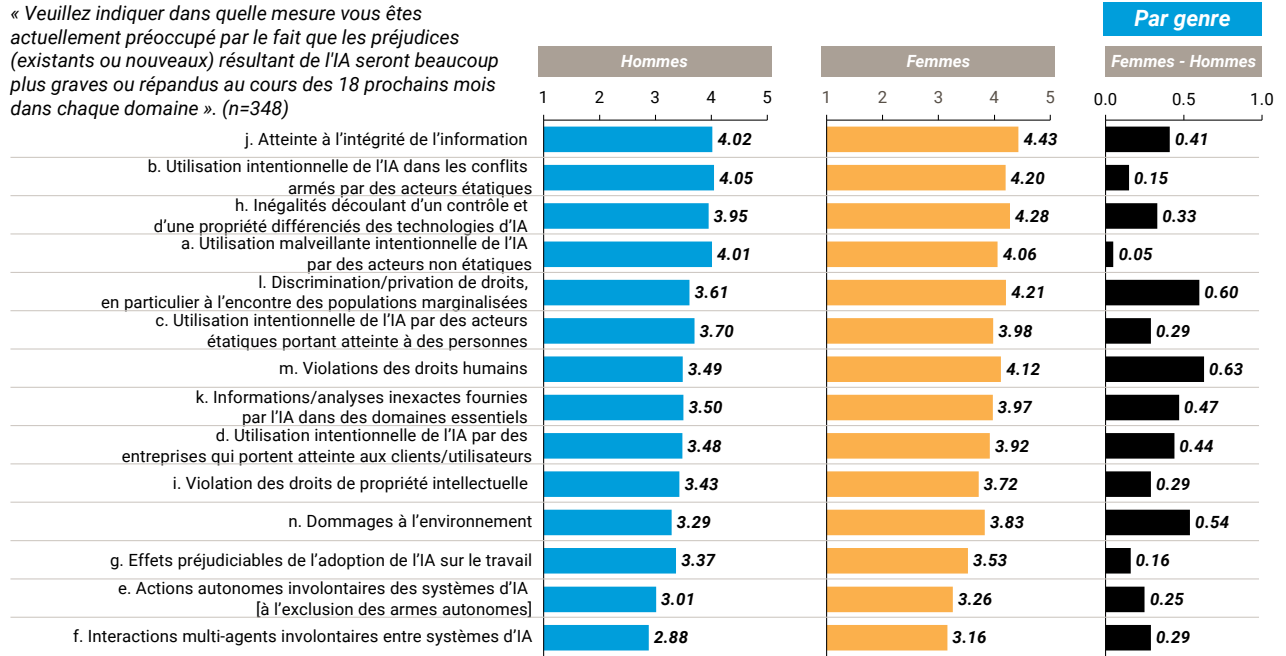


Note : Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

Les femmes sont plus préoccupées que les hommes dans tous les domaines donnés en exemple

Les lacunes sont particulièrement importantes en ce qui concerne les violations des droits humains, la discrimination et l'environnement

« Veuillez indiquer dans quelle mesure vous êtes actuellement préoccupé par le fait que les préjudices (existants ou nouveaux) résultant de l'IA seront beaucoup plus graves ou répandus au cours des 18 prochains mois dans chaque domaine ». (n=348)

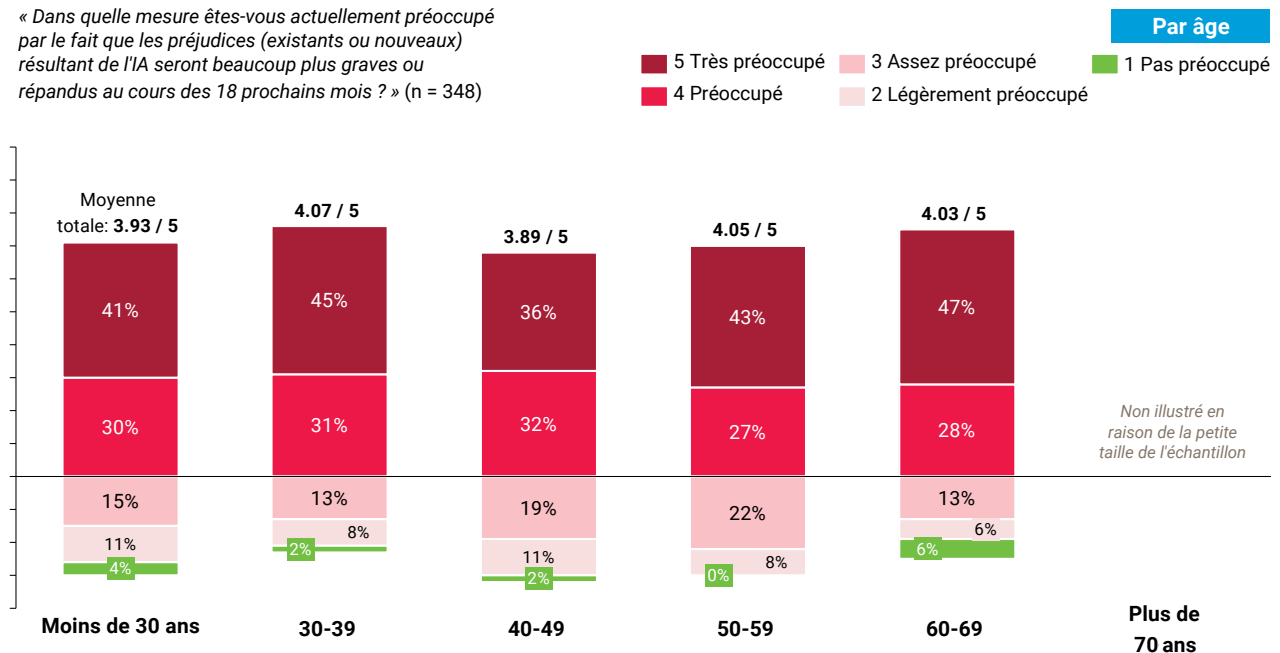


Moyenne, où : 1 = Pas préoccupé, 2 = Légèrement préoccupé, 3 = Assez préoccupé, 4 = Préoccupé, 5 = Très préoccupé. Note : Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

71 % se disent préoccupés/très préoccupés par des préjudices pouvant résulter de l'IA dans les 18 prochains mois

Différences relativement faibles dans les préoccupations en fonction de l'âge du répondant

« Dans quelle mesure êtes-vous actuellement préoccupé par le fait que les préjudices (existants ou nouveaux) résultant de l'IA seront beaucoup plus graves ou répandus au cours des 18 prochains mois ? » (n = 348)

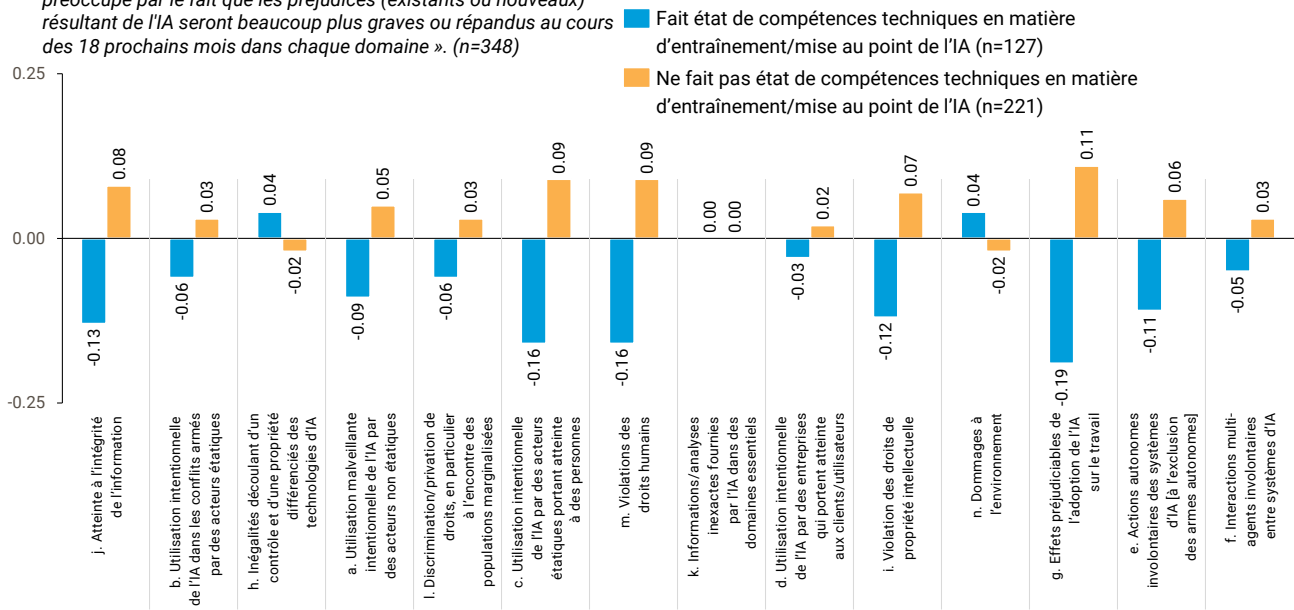


Note : Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

Les répondants qui font état de compétences techniques (entraînement/mise au point de l'IA) sont moins préoccupés par la plupart des domaines donnés en exemple

« Veuillez indiquer dans quelle mesure vous êtes actuellement préoccupé par le fait que les préjudices (existants ou nouveaux) résultant de l'IA seront beaucoup plus graves ou répandus au cours des 18 prochains mois dans chaque domaine ». (n=348)

Par compétence



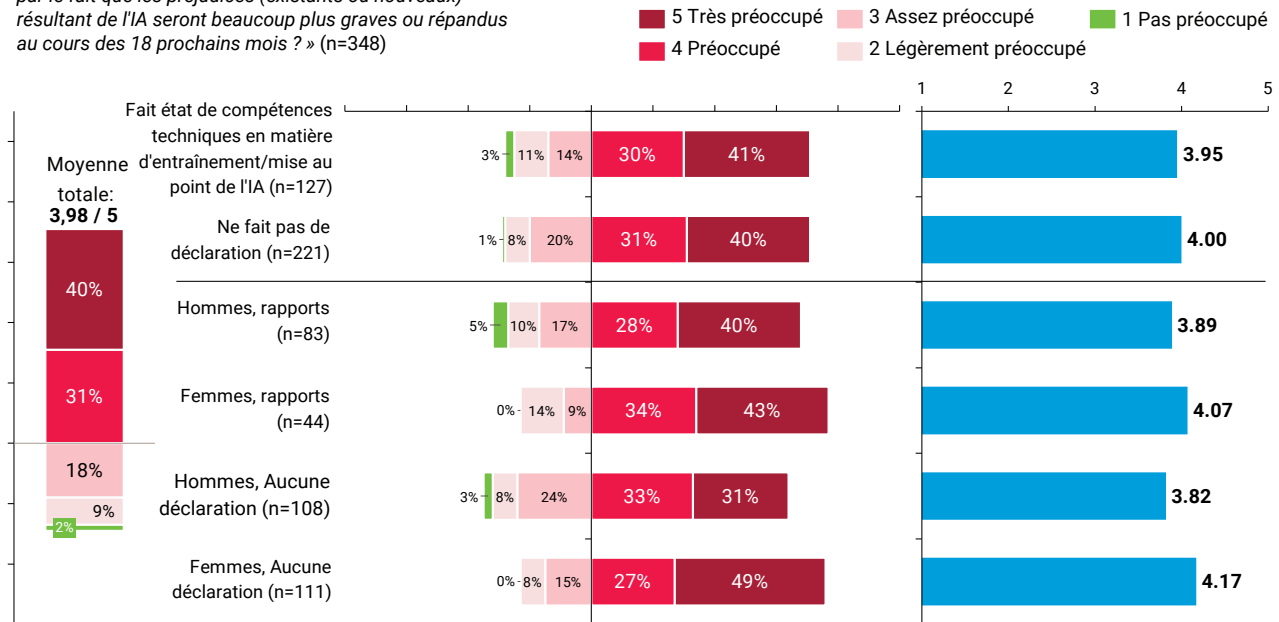
Représenté : Différence entre l'évaluation globale (tous les répondants) et l'évaluation du groupe indiqué où : 1 = Pas préoccupé, 2 = Légèrement préoccupé, 3 = Assez préoccupé, 4 = Préoccupé, 5 = Très préoccupé. Note : Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

Effets limités des compétences techniques (entraînement/mise au point de l'IA)

Les hommes sont moins préoccupés que les femmes, quelle que soit leur situation

« Dans quelle mesure êtes-vous actuellement préoccupé par le fait que les préjudices (existants ou nouveaux) résultant de l'IA seront beaucoup plus graves ou répandus au cours des 18 prochains mois ? » (n=348)

Par genre et par compétence



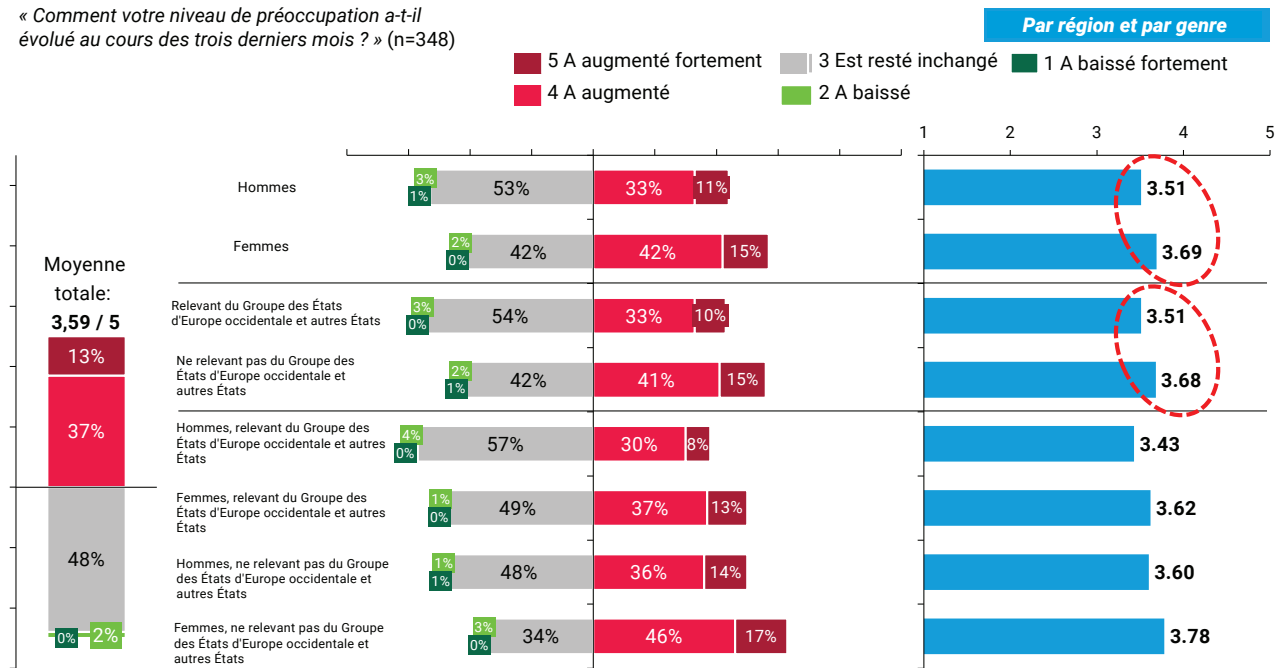
Note : Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

Évolution de la perception du niveau de préoccupation au cours des 3 derniers mois concernant les risques de préjudices liés à l'IA

50 % des répondants se disent plus préoccupés au cours des trois derniers mois ; 48 % n'indiquent pas de changements

Presque personne s'est dit moins préoccupé ; plus de femmes, ne relevant pas du Groupe des États d'Europe occidentale et autres États, ont un niveau de préoccupation plus élevé

« Comment votre niveau de préoccupation a-t-il évolué au cours des trois derniers mois ? » (n=348)



Note : Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Source : Enquête sur les risques liés à l'IA menée par le Bureau de l'Envoyé du Secrétaire général pour la technologie, 13-25 mai 2024.

Annexe F: réponses à l'analyse des opportunités

À la demande de l'Organe consultatif de haut niveau sur l'intelligence artificielle, le Bureau de l'Envoyé du Secrétaire général pour la technologie a réalisé une analyse mondiale des opportunités sur l'IA. Les experts ont été invités à répondre en donnant leur avis à titre personnel (et non au nom de leur organisme ou employeur). L'enquête a été divisée en sections couvrant les opportunités dans les pays à revenu élevé/intermédiaire supérieur et dans les pays à revenu faible/intermédiaire inférieur, seuls les répondants faisant état de connaissances sur les contextes des pays à revenu faible/intermédiaire inférieur ayant répondu à ces questions. L'enquête menée n'a porté que sur les éventuels effets positifs de l'IA.

On a demandé aux répondants dans quelle mesure ils avaient connaissance, à ce jour, d'exemples où l'IA avait augmenté l'activité économique, accéléré les découvertes scientifiques et contribué aux progrès dans la réalisation des objectifs de développement durable¹. On leur a demandé de fournir des détails, y compris des études de cas, des noms d'organisations, des données et des liens vers des articles, publications ou papiers pertinents. On leur a ensuite demandé dans quelle mesure ils s'attendaient à des progrès au cours des trois prochaines années selon les mêmes critères.

En outre, on a demandé aux répondants à quel moment l'IA aurait des effets majeurs dans ces domaines, avec un degré de confiance ou de probabilité de 50 %. D'autres questions ont été posées, notamment sur les acteurs qui participent à saisir les occasions offertes par l'IA, les obstacles qui contribuent à la fracture de l'IA entre les pays ainsi que les limites auxquelles certains groupes font face pour exploiter les avantages qui découlent de l'IA et comment y remédier.

L'enquête a été menée du 9 au 21 août 2024. La liste des invités a été établie à partir des réseaux du Bureau de l'Envoyé du Secrétaire général pour la technologie et de l'Organe consultatif, et comprenait des participants aux études approfondies de l'Organe. En outre, la réunion AI for Good de l'UIT et les réseaux de la CNUCED ont été généreusement mis à dispositions aux fins de l'enquête. Au total, plus d'un millier de personnes ont été invitées et plus de 120 personnes ont répondu à l'enquête, fournissant ainsi une perspective riche et diversifiée (notamment sur la base des régions et du genre) concernant les avantages qui découlent de l'IA.

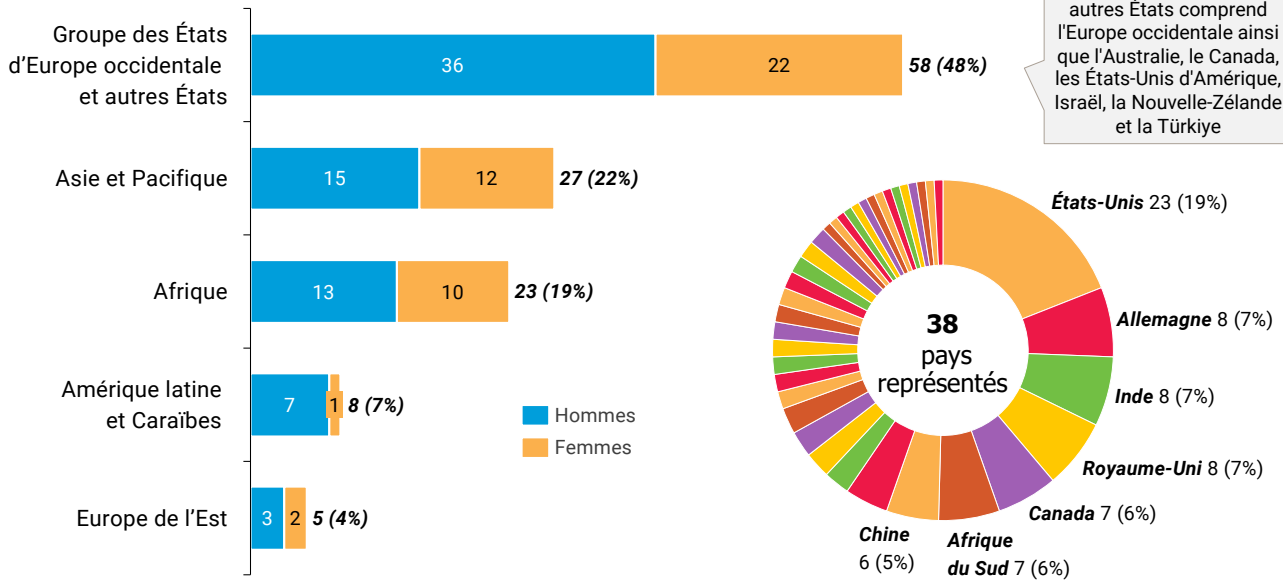
¹ Les objectifs 8 (Travail décent et croissance économique) et 9 (Innovation, industrie et infrastructure) n'ont pas fait l'objet d'une question distincte, étant donné leur lien étroit avec l'augmentation de l'activité économique. L'objectif 17 (Partenariats pour les objectifs) n'a pas non plus fait l'objet d'une question distincte.

Vu d'ensemble de l'échantillon

Représentation régionale : forte participation mondiale

Permet de comparer les réponses entre le Groupe des États d'Europe occidentale et autres États et d'autres régions

Répondants par région de nationalité* (n = 121)

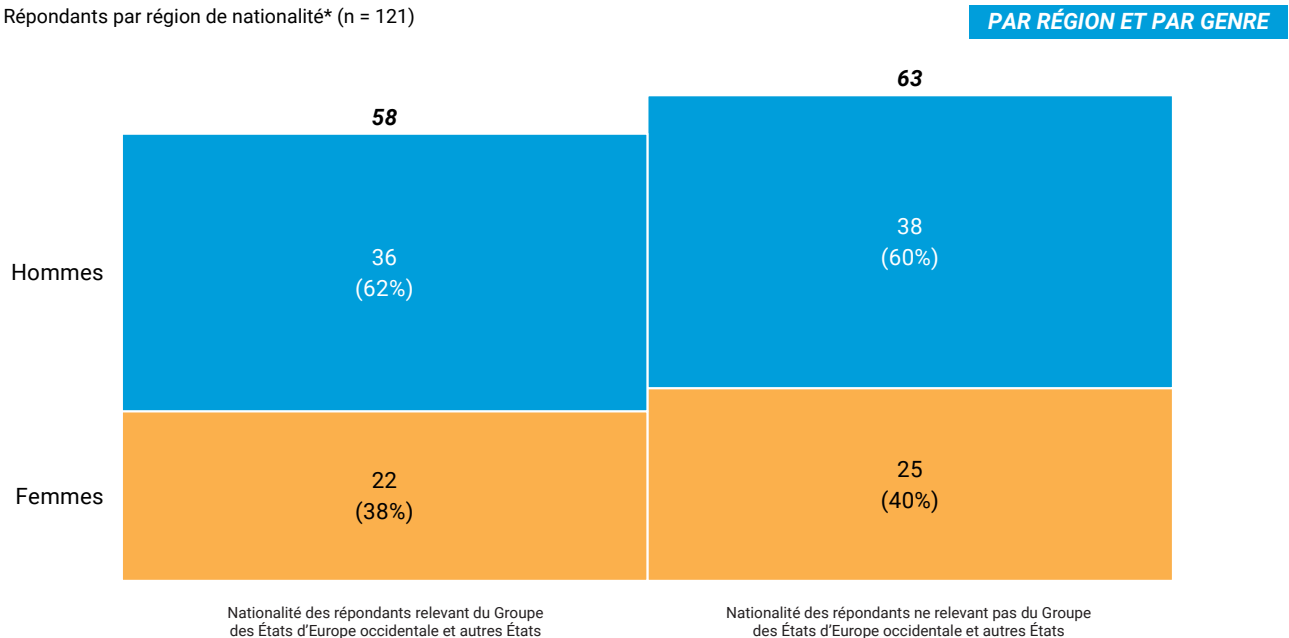


* 9 répondants (7%) ont indiqué plusieurs nationalités. Si les répondants résidaient dans l'un de leurs pays de nationalité, c'est cette nationalité qui a été retenue pour l'analyse (8 sur 9). Dans le cas contraire, c'est la nationalité la moins représentée qui a été retenue (1 sur 9). Seuls les répondants déclarant avoir des connaissances pertinentes ont été interrogés sur les pays à revenu moyen inférieur et à revenu faible. Source : OSET Sondage des opportunités de l'IA, 9-21 août 2024

Les hommes représentent ~60 % des répondants relevant du Groupe des États d'Europe occidentale et autres États et de ceux ne relevant pas du Groupe

La cohérence signifie que l'analyse univariée par genre et par région n'est pas immédiatement contaminée

Répondants par région de nationalité* (n = 121)

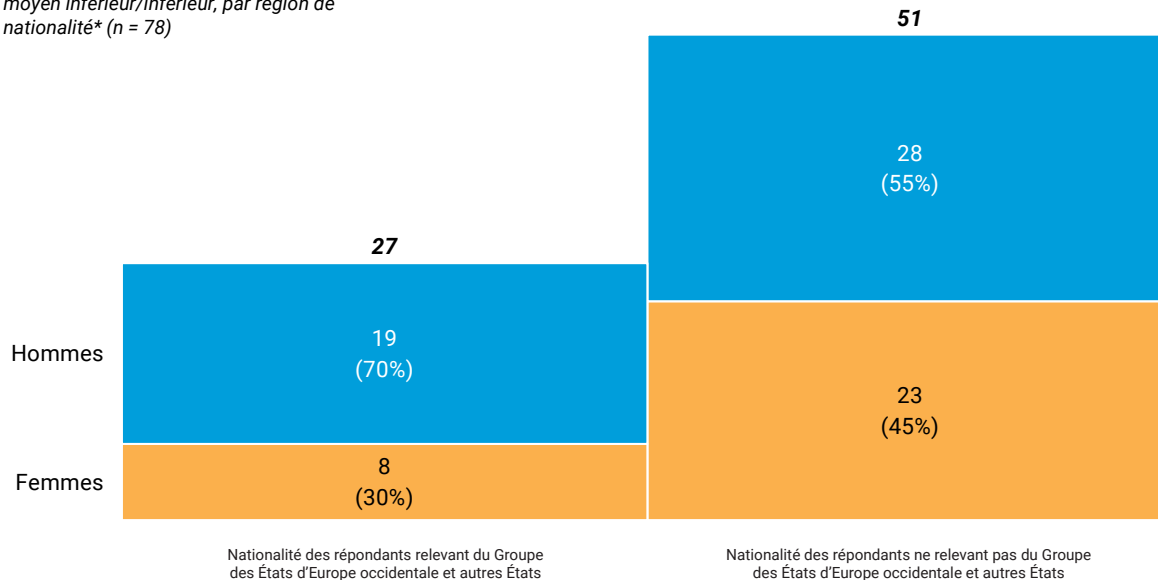


*9 répondants (7%) ont indiqué plusieurs nationalités. Si les répondants résidaient dans l'un de leurs pays de nationalité, c'est cette nationalité qui a été retenue pour l'analyse (8 sur 9). Dans le cas contraire, c'est la nationalité la moins représentée qui a été retenue (1 sur 9). Source : OSET Sondage des opportunités de l'IA, 9-21 août 2024

L'échantillon des experts issus de pays en développement est moins équilibré

Répondants confirmant leurs connaissances spécifiques sur les contextes des pays à revenu moyen inférieur/inférieur, par région de nationalité* (n = 78)

PAR RÉGION ET PAR GENRE



* 9 répondants (7%) ont indiqué plusieurs nationalités. Si les répondants résidaient dans l'un de leurs pays de nationalité, c'est cette nationalité qui a été retenue pour l'analyse (8 sur 9). Dans le cas contraire, c'est la nationalité la moins représentée qui a été retenue (1 sur 9). Seuls les répondants déclarant avoir des connaissances pertinentes ont été interrogés sur les pays à revenu moyen inférieur et à revenu faible. Source : OSET Sondage des opportunités de l'IA, 9-21 août 2024

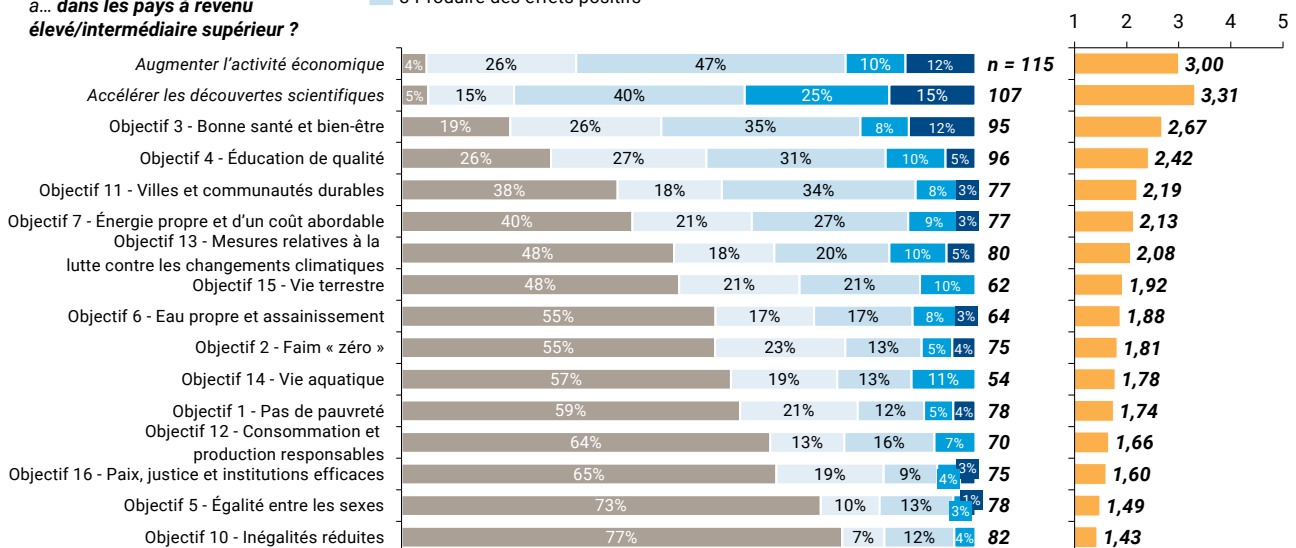
Perceptions des effets positifs de l'IA à ce jour

Effets positifs à ce jour sur la croissance et la science, mais moindre sur la plupart des objectifs de développement durable

Effets à ce jour dans les pays à revenu élevé/intermédiaire supérieur

Connaissez-vous des exemples où l'IA contribue actuellement ou a récemment contribué directement à... dans les pays à revenu élevé/intermédiaire supérieur ?

1 Produire aucun effet positif
2 Produire des effets positifs mineurs
3 Produire des effets positifs
4 Produire des effets positifs majeurs
5 Produire des effets positifs transformateurs



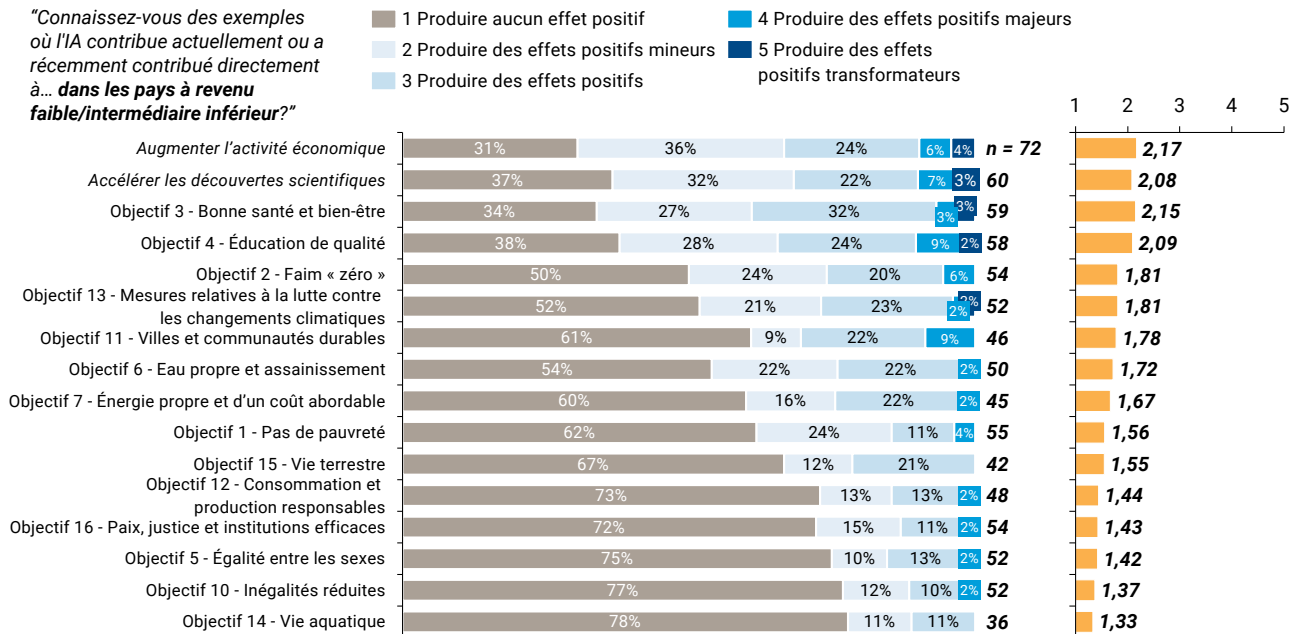
Note : ne prend pas en compte les réponses « Ne sait pas »/« Pas d'opinion » et les réponses en blanc. Les n particuliers sont indiqués à chaque question. N'a pas posé de questions sur les objectifs 8, 9 et 17.

Source : OSET Sondage des opportunités de l'IA, 9-21 août 2024

Moins d'effets signalés sur tous les fronts dans les pays à faible revenu

Effets à ce jour dans les pays à revenu faible/intermédiaire inférieur

"Connaissez-vous des exemples où l'IA contribue actuellement ou a récemment contribué directement à... dans les pays à revenu faible/intermédiaire inférieur?"



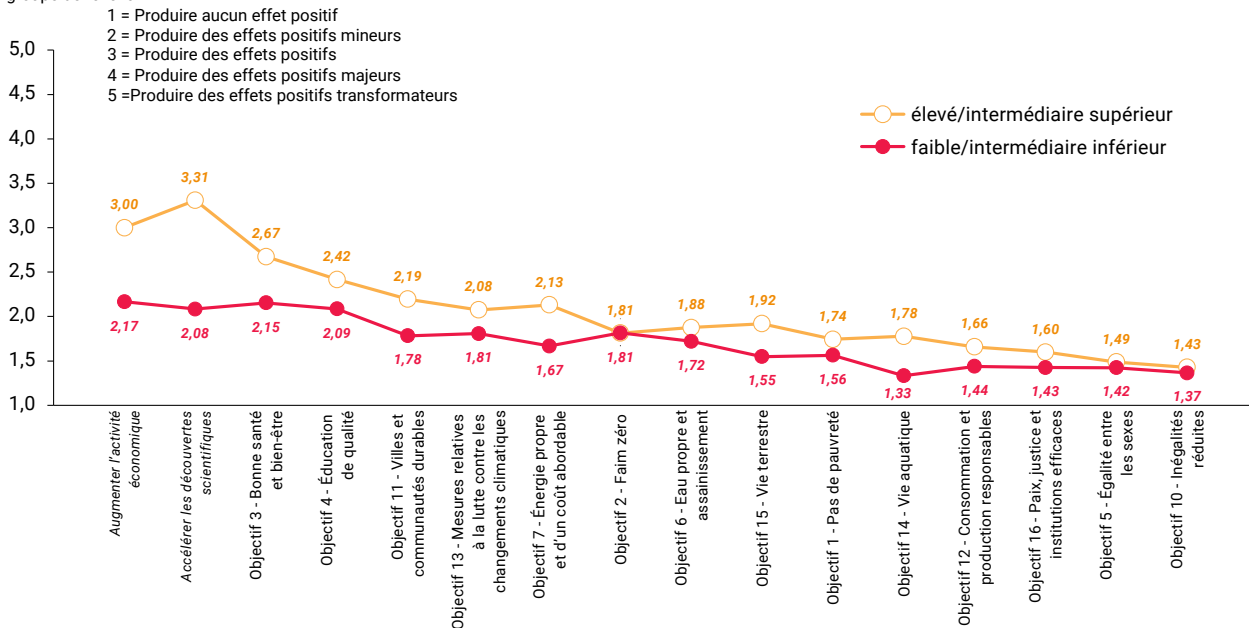
Note : ne prend pas en compte les réponses « Ne sait pas »/« Pas d'opinion » et les réponses en blanc. Les n particuliers sont indiqués à chaque question. N'a pas posé de questions sur les objectifs 8, 9 et 17.

Source : OSET Sondage des opportunités de l'IA, 9-21 août 2024

Moins d'effets signalés sur tous les fronts dans les pays à faible revenu

L'écart est le plus prononcé en ce qui concerne la croissance économique et la science

Note moyenne à la question "Connaissez-vous des exemples où l'IA contribue actuellement ou a récemment contribué directement à produire... ?" Par pays et groupe de revenu:



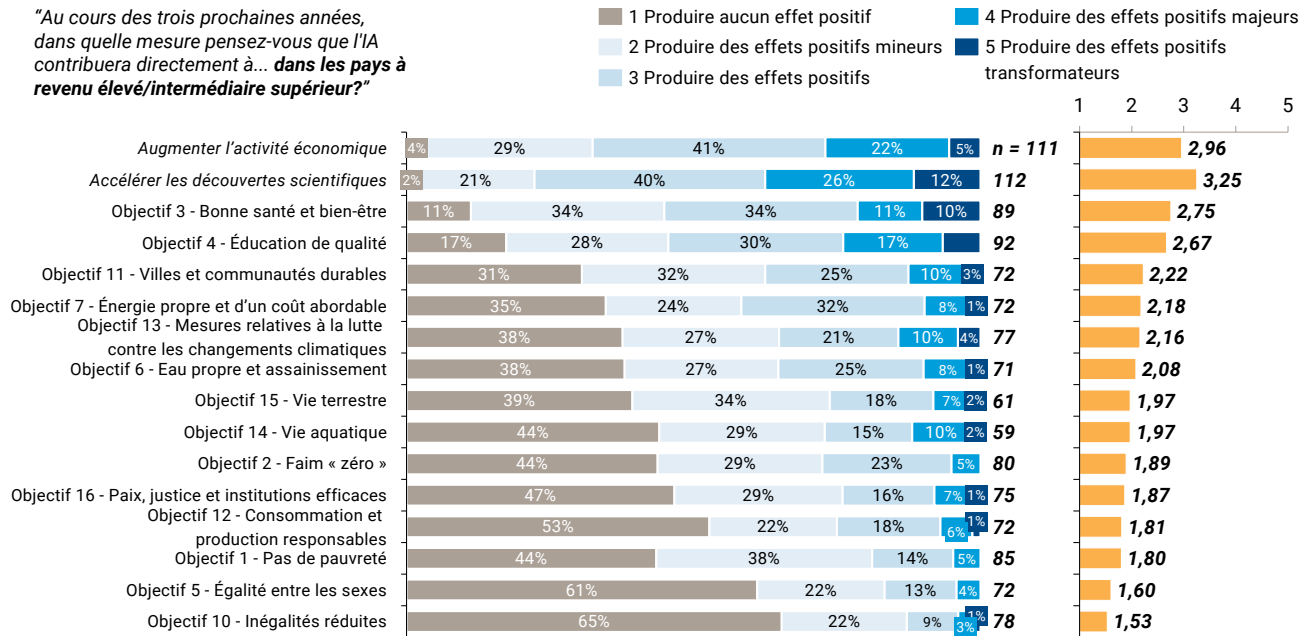
Note: Exclut les réponses « Je ne sais pas » / « Sans opinion » et les réponses en blanc. Seuls les répondants déclarant avoir des connaissances pertinentes ont été interrogés sur les pays à revenu moyen inférieur et à revenu faible. N'a pas posé de questions sur les ODD 8, 9 et 17. Source: OSET Sondage des opportunités de l'IA, 9-21 août 2024

Perception des effets positifs attendus de l'IA au cours des trois prochaines années

Effets attendus sur la croissance, la science, la santé, l'éducation – moins sur d'autres domaines

Effets attendus au cours des trois prochaines années dans les pays à revenu élevé/intermédiaire supérieur

"Au cours des trois prochaines années, dans quelle mesure pensez-vous que l'IA contribuera directement à... dans les pays à revenu élevé/intermédiaire supérieur?"



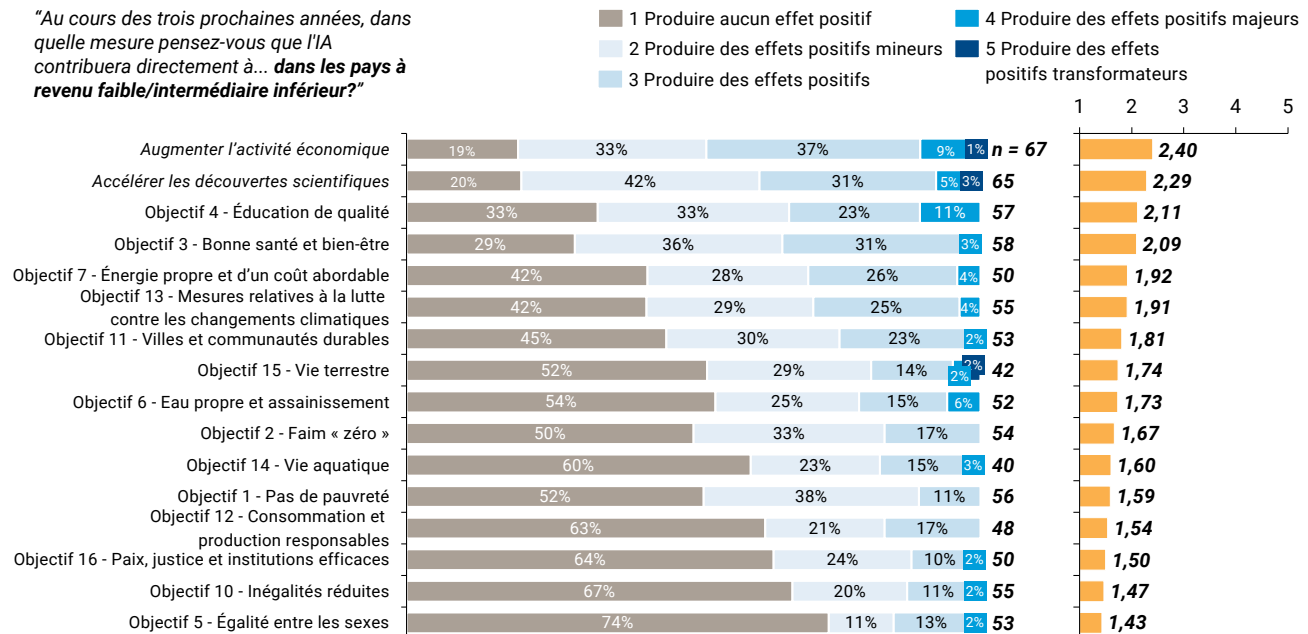
Note : ne prend pas en compte les réponses « Ne sait pas »/« Pas d'opinion » et les réponses en blanc. Les n particuliers sont indiqués à chaque question. N'a pas posé de questions sur les objectifs 8, 9 et 17.

Source : OSET Sondage des opportunités de l'IA, 9-21 août 2024

Effets sont attendus dans les pays à faible revenu, mais de nouveau plus limité

Effets attendus au cours des trois prochaines années dans les pays à revenu faible/intermédiaire inférieur

"Au cours des trois prochaines années, dans quelle mesure pensez-vous que l'IA contribuera directement à... dans les pays à revenu faible/intermédiaire inférieur?"



Note : ne prend pas en compte les réponses « Ne sait pas »/« Pas d'opinion » et les réponses en blanc. Les n particuliers sont indiqués à chaque question. N'a pas posé de questions sur les objectifs 8, 9 et 17.

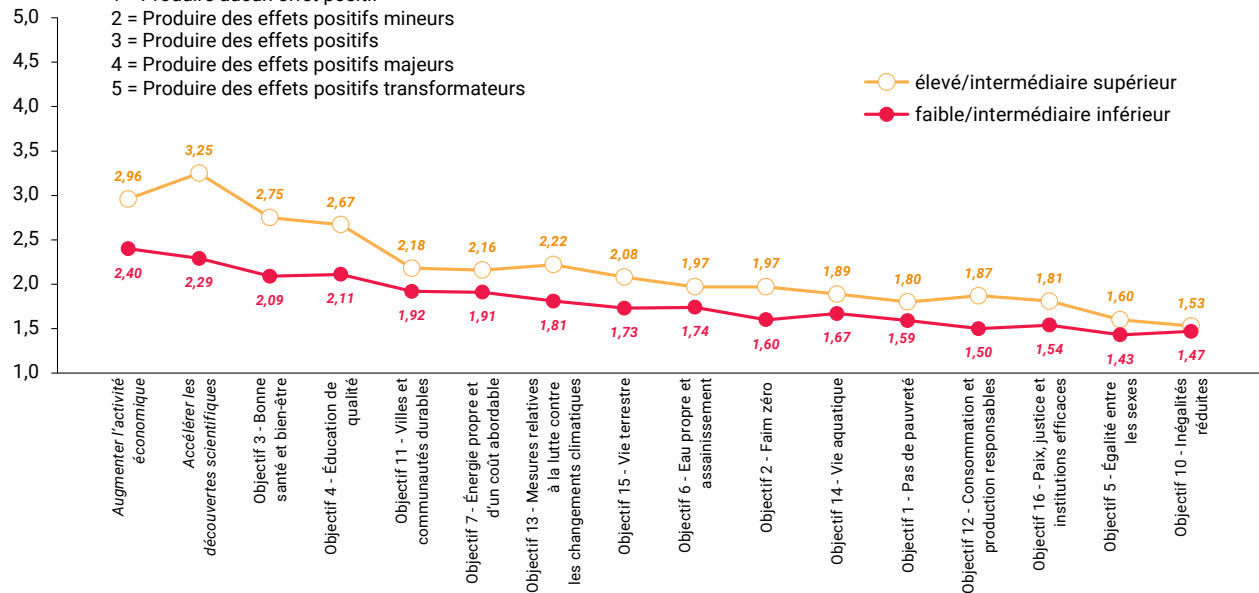
Source : OSET Sondage des opportunités de l'IA, 9-21 août 2024

Moins d'effets attendus sur tous les fronts dans les pays à faible revenu

L'écart est le plus prononcé en ce qui concerne la croissance économique, la science, la santé et l'éducation

Note moyenne à la question « Au cours des trois prochaines années, dans quelle mesure pensez-vous que l'IA contribuera directement à... ? » Par pays et groupe de revenu, où:

- 1 = Produire aucun effet positif
- 2 = Produire des effets positifs mineurs
- 3 = Produire des effets positifs
- 4 = Produire des effets positifs majeurs
- 5 = Produire des effets positifs transformateurs



Note : ne prend pas en compte les réponses "Ne sait pas"/ "Pas d'opinion" et les réponses en blanc. Les n particuliers sont indiqués à chaque question. N'a pas posé de questions sur les objectifs 8, 9 et 17.
 Source : OSET Sondage des opportunités de l'IA, 9-21 août 2024
 Graphiques préparés dans think-cell

Annexe G: liste des abréviations

ADN	Acide désoxyribonucléique
AIEA	Agence internationale de l'énergie atomique
ANSI	Institut national américain de normalisation
ASEAN	Association des nations de l'Asie du Sud-Est
BSI	Institut britannique des normes
CEI	Commission électrotechnique internationale
CEN	Comité européen de normalisation
CENELEC	Comité européen de normalisation électrotechnique
CERN	Organisation européenne pour la recherche nucléaire
CNUCED	Conférence des Nations Unies sur le commerce et le développement
CNUDCI	Commission des Nations Unies pour le droit commercial international
ETSI	Institut européen de normalisation des télécommunications
FAO	Organisation des Nations Unies pour l'alimentation et l'agriculture
FMI	Fonds monétaire international
G20	Groupe des 20
G7	Groupe des Sept
GAFI	Groupe d'action financière
GIEC	Groupe d'experts intergouvernemental sur l'évolution du climat
IA	Intelligence artificielle
IEEE	Institute of Electrical and Electronics Engineers
ISO	Organisation internationale de normalisation
LGBT+	Personnes lesbiennes, gays, bisexuelles et transgenres
OACI	Organisation de l'aviation civile internationale
OCDE	Organisation de coopération et de développement économiques
OEA	Organisation des États américains
OIAC	Organisation pour l'interdiction des armes chimiques
OIT	Organisation internationale du Travail
OMC	Organisation mondiale du commerce
OMI	Organisation maritime internationale
OMM	Organisation météorologique mondiale
OMPI	Organisation mondiale de la propriété intellectuelle
OMS	Organisation mondiale de la Santé
ONUDI	Organisation des Nations Unies pour le développement industriel
ONU-HABITAT	Programme des Nations Unies pour les établissements humains
PAM	Programme alimentaire mondial
PME	Petites et moyennes entreprises
PMIA	Partenariat mondial sur l'intelligence artificielle
PNUD	Programme des Nations Unies pour le développement
UA	Union africaine
UE	Union européenne
UIT	Union internationale des télécommunications
UNESCO	Organisation des Nations Unies pour l'éducation, la science et la culture
UNICEF	Fonds international des Nations Unies pour l'enfance
UNICRI	Institut interrégional de recherche des Nations Unies sur la criminalité et la justice
UNSCEAR	Comité scientifique des Nations Unies pour l'étude des effets des rayonnements ionisants
UNU	Université des Nations Unies

Donateurs

L'Organe remercie les gouvernements et partenaires suivants pour leurs contributions financières et en nature, sans lesquelles il n'aurait pas été en mesure de s'acquitter de ses responsabilités :

Gouvernement de l'Allemagne
Gouvernement du Royaume d'Arabie Saoudite
Gouvernement des Émirats arabes unis
Gouvernement de la Finlande
Gouvernement de l'Italie
Gouvernement du Japon
Gouvernement du Royaume des Pays-Bas
Gouvernement du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord
Gouvernement de la République tchèque
Gouvernement de Singapour
Gouvernement de la Suisse
Union européenne
Omidyar Network Fund
L'Organisation internationale de la Francophonie

Secrétariat

Simon Chesterman
Quintin Chou-Lambert
Eleonore Fournier-Tombs
Sebastian Frank
David Michael Kelly
Brian Shung Seun Lau
Andrew Morritt
Filippo Pierozzi
Mehdi Snène
Isabel de Sola
Lucia Velasco
Rebekah Hayoung Woo



**United
Nations**



**AI
Advisory
Body**