

5. Politiques et initiatives dans le domaine de l'IA

Les politiques et initiatives liées à l'intelligence artificielle (IA) se multiplient, que ce soit au niveau des pouvoirs publics, des entreprises, des organismes techniques, de la société civile et des syndicats. Des initiatives intergouvernementales en matière d'IA sont aussi en cours de développement. Ce chapitre recense les politiques, initiatives et stratégies d'IA élaborées par diverses parties prenantes dans le monde entier, aux niveaux tant national qu'international. Il observe que, d'une manière générale, les initiatives adoptées par les gouvernements nationaux envisagent d'abord l'IA comme un moyen d'améliorer la productivité et la compétitivité et s'appuient sur des plans d'action visant à renforcer : i) les facteurs, tels que les capacités de recherche en IA ; ii) la demande ; iii) les industries amont et apparentées ; iv) la stratégie, la structure et la rivalité des entreprises ; et v) la gouvernance et la coordination au plan national. Les initiatives internationales incluent notamment la Recommandation du Conseil de l'OCDE sur l'intelligence artificielle, qui constitue le premier instrument d'orientation intergouvernemental sur l'IA et énonce un certain nombre de principes et de priorités d'action pour une approche responsable à l'appui d'une IA digne de confiance.

Intelligence artificielle et compétitivité économique : stratégies et plans d'action

L'intelligence artificielle (IA) apparaît comme une priorité croissante des institutions gouvernementales, aux niveaux tant national qu'international. Nombre des initiatives lancées à ce jour par des gouvernements nationaux envisagent le recours à l'IA comme un moyen d'améliorer la productivité et la compétitivité. Les priorités énoncées dans les plans d'action nationaux relatifs à l'IA peuvent être regroupées en cinq grandes catégories, qui recoupent en partie celles du cadre de compétitivité économique de Porter. Ces priorités concernent : i) les facteurs, tels que les capacités de recherche en IA, y compris les compétences ; ii) la demande ; iii) les industries amont et apparentées ; iv) la stratégie, la structure et la rivalité des entreprises ; et v) les questions de gouvernance et de coordination au plan national (Encadré 5.1). Par ailleurs, parmi les considérations politiques propres à l'IA, les pouvoirs publics portent une attention croissante à la transparence, au respect des droits de l'homme et à l'éthique.

Plusieurs pays et économies partenaires de l'OCDE, notamment l'Allemagne, le Canada, la République populaire de Chine (ci-après « la Chine »), les États-Unis, la France, l'Inde, le Royaume-Uni et la Suède, se sont dotés de stratégies d'IA dédiées. Certains pays, comme la Corée, le Danemark et le Japon, ont inscrit des mesures en matière d'IA dans des programmes de portée plus générale. De nombreux autres – dont l'Australie, l'Espagne, l'Estonie, la Finlande, Israël et l'Italie – travaillent actuellement à l'élaboration d'une stratégie. L'ensemble des stratégies gouvernementales visent à : accroître le nombre de chercheurs et de diplômés qualifiés en IA ; renforcer les capacités nationales de recherche en IA ; et faire en sorte que les résultats de la recherche en IA débouchent sur des applications dans les secteurs public et privé. Dans leur manière d'aborder les conséquences économiques, sociales, éthiques, politiques et juridiques des progrès de l'IA, ces initiatives nationales reflètent les différences de cultures, de systèmes juridiques, de taille de pays et de degré d'adoption de l'IA, bien que la mise en œuvre des politiques en soit encore à un stade précoce. Le présent chapitre examine par ailleurs l'évolution récente de la réglementation et des politiques en matière d'IA ; il s'abstient cependant d'analyser ou d'évaluer la réalisation des objectifs des initiatives nationales ou le succès des différentes approches adoptées à cet égard.

La question de l'IA a également été abordée au niveau d'instances internationales telles que le Groupe des sept (G7), le Groupe des vingt (G20), l'OCDE, l'Union européenne et les Nations Unies. La Commission européenne met l'accent sur le rôle de l'IA pour promouvoir l'efficacité et la flexibilité, les échanges et la coopération, la productivité, la compétitivité et la croissance, ainsi que la qualité de vie des citoyens. Suite à la réunion des ministres des TIC du G7 qui s'était tenue au Japon en avril 2016, la réunion des ministres des TIC et de l'industrie du G7 organisée à Turin (Italie) en septembre 2017 a mis en avant une vision de l'IA « centrée sur l'humain ». À cette occasion, les ministres ont décidé d'encourager la collaboration internationale et le dialogue multipartite autour de l'IA, et de promouvoir une meilleure compréhension de la coopération dans ce domaine avec l'appui de l'OCDE. On notera également l'attention continue que le G20 porte à l'IA, en particulier avec la proposition du Japon de mettre l'intelligence artificielle à l'ordre du jour de sa présidence du G20 en 2019 (G20, 2018^[1]).

Principes concernant l'utilisation de l'intelligence artificielle dans la société

Plusieurs groupes de parties prenantes réfléchissent activement aux moyens de guider le développement et le déploiement de l'IA, afin que celle-ci bénéficie à l'ensemble de la société. En avril 2016, par exemple, l'Institute of Electrical and Electronics Engineers (IEEE)

a lancé une Initiative mondiale sur l'éthique dans la conception des systèmes autonomes et intelligents. Il a également publié en décembre 2017 la deuxième version de ses Principes d'intégration de l'éthique dès la conception (*Ethically Aligned Design*). La version finale était prévue pour le début de l'année 2019. Le Partenariat pour l'intelligence artificielle au service des citoyens et de la société (*Partnership on Artificial Intelligence to Benefit People and Society*), lancé en septembre 2016 sur la base d'une série de principes généraux, a commencé à travailler à l'élaboration de principes portant sur des questions spécifiques comme celle de la sécurité. Les Principes d'Asilomar formulent une série d'orientations de recherche, de normes éthiques et de valeurs pour assurer un développement sûr et socialement bénéfique de l'IA à court et à long terme. L'Initiative sur l'intelligence artificielle (*AI Initiative*), qui regroupe des experts, des praticiens et des citoyens du monde entier, s'est fixée pour objectif l'élaboration d'une définition commune de notions comme celle d'explicabilité de l'IA.

Plusieurs initiatives ont abouti à l'élaboration de séries de principes utiles pour guider le développement de l'IA (Tableau 5.1). Nombre de ces principes s'adressent aux communautés techniques qui mènent des activités de recherche et développement (R-D) ayant trait aux systèmes d'IA. Bien qu'ayant été développés pour une grande part dans le cadre de processus multipartites, ils ciblent cinq grandes communautés d'acteurs : les experts techniques, le secteur privé, les administrations, les universités et les syndicats. La communauté des experts techniques comprend le Future of Life Institute, l'IEEE, l'association JSAI (Japanese Society for Artificial Intelligence), la conférence FATML (Fairness, Accountability and Transparency in Machine Learning) et l'ACM (Association for Computing Machinery). Pour le secteur privé, on citera par exemple le Partenariat sur l'IA, le Conseil de l'industrie des technologies de l'information et Satya Nadella, président-directeur général de Microsoft. Du côté des administrations, on mentionnera le ministère des Affaires intérieures et des Communications du Japon, la Commission mondiale d'éthique des connaissances scientifiques et des technologies, et le Conseil britannique de la recherche en ingénierie et en sciences physiques. Pour les universités, on évoquera l'Université de Montréal et Nicolas Economou, président-directeur général de H5 et conseiller spécial pour l'AI Initiative lancée par l'incubateur The Future Society, à la Harvard Kennedy School. Enfin, du côté des syndicats, la fédération syndicale internationale UNI Global Union a elle aussi élaboré des principes éthiques sur l'intelligence artificielle.

Des thèmes communs se dégagent de ces initiatives. De fait, les différentes parties prenantes ont élaboré des lignes directrices sur le respect des valeurs humaines et des droits de l'homme, la non-discrimination, l'information et le contrôle, l'accès aux données, la protection de la vie privée et le contrôle des informations à caractère personnel, la sûreté et la sécurité, les compétences, la transparence et l'explicabilité, la responsabilité et la redevabilité, le dialogue à l'échelle de l'ensemble de la société, et la mesure de l'IA.

En mai 2018, le Comité de la politique de l'économie numérique de l'OCDE a créé le Groupe d'experts sur l'intelligence artificielle à l'OCDE (AIGO) chargé d'élaborer des principes à mettre en œuvre dans le cadre des politiques publiques et de la coopération internationale et ayant vocation à promouvoir la confiance dans l'IA et son adoption (OCDE, 2019^[2]). Les travaux du groupe d'experts ont nourri l'élaboration de la *Recommandation du Conseil de l'OCDE sur l'intelligence artificielle* (OCDE, 2019^[3]), à laquelle 42 pays ont adhéré le 22 mai 2019.

Tableau 5.1. Liste non exhaustive de lignes directrices, principes ou déclarations sur l'IA émanant de diverses parties prenantes

Désignation	Principes et lignes directrices sur l'IA élaborés par diverses parties prenantes
ACM	ACM (2017), « 2018 ACM Code of Ethics and Professional Conduct: Draft 3 », Association for Computing Machinery Committee on Professional Ethics, https://ethics.acm.org/2018-code-draft-3/ USACM (2017), « Statement on Algorithmic Transparency and Accountability », Association for Computing Machinery US Public Policy Council, www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf
AI safety	Amodei D. et al. (2016), « Concrete Problems in AI Safety », 25 juillet, https://arxiv.org/pdf/1606.06565.pdf
Asilomar	FLI (2017), « Asilomar AI Principles », Future of Life Institute, https://futureoflife.org/ai-principles/
COMEST	COMEST (2017), « Rapport de la COMEST sur l'éthique de la robotique », Commission mondiale d'éthique des connaissances scientifiques et des technologies, https://unesdoc.unesco.org/ark:/48223/pf0000253952_fre
Economou	Economou N. (2017), « A "Principled" Artificial Intelligence Could Improve Justice », 3 octobre, <i>Aba Journal</i> , www.abajournal.com/legalrebels/article/a_principled_artificial_intelligence_could_improve_justice
EPSRC	EPSRC (2010), « Principles of Robotics », Engineering and Physical Sciences Research Council (Royaume-Uni), https://epsrc.ukri.org/research/ourportfolio/themes/engineering/activities/principlesofrobotics/
FATML	FATML (2016), « Principles for Accountable Algorithms and a Social Impact Statement for Algorithms », Fairness, Accountability and Transparency in Machine Learning, www.fatml.org/resources/principles-for-accountable-algorithms
FPF	FPF (2018), « Beyond Explainability: A Practical Guide to Managing Risk in Machine Learning Models », The Future of Privacy Forum, https://fpf.org/wp-content/uploads/2018/06/Beyond-Explainability.pdf
GEE	GEE (2018), « Déclaration sur l'intelligence artificielle, la robotique et les systèmes "autonomes" », Groupe européen d'éthique des sciences et des nouvelles technologies, https://publications.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1/language-fr/format-PDF
Google	Google (2018), « AI at Google: Our Principles », https://www.blog.google/technology/ai/ai-principles/
HLEG AI (CE)	Groupe d'experts indépendants de haut niveau sur l'intelligence artificielle (AI HLEG) de la CE (2019), « Lignes directrices en matière d'éthique pour une IA digne de confiance », https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60427
IEEE	IEEE (2017), « Ethically Aligned Design: Version 2 », IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, Institute of Electrical and Electronics Engineers, http://standards.ieee.org/develop/indconn/ec/ead_v2.pdf
Intel	Intel (2017), « AI - The Public Policy Opportunity », https://blogs.intel.com/policy/files/2017/10/Intel-Artificial-Intelligence-Public-Policy-White-Paper-2017.pdf
ITI	ITI (2017), « AI Policy Principles », Information Technology Industry Council, www.itic.org/resources/AI-Policy-Principles-FullReport2.pdf
JSAI	JSAI (2017), « The Japanese Society for Artificial Intelligence Ethical Guidelines », The Japanese Society for Artificial Intelligence, http://ai-elsi.org/wp-content/uploads/2017/05/JSAI-Ethical-Guidelines-1.pdf
MIC	MIC (2017), « Draft AI R&D Guidelines for International Discussions », Ministère des Affaires intérieures et des Communications du Japon, www.soumu.go.jp/main_content/000507517.pdf
MIC	MIC (2018), « Draft AI Utilization Principles », Ministère des Affaires intérieures et des Communications du Japon, www.soumu.go.jp/main_content/000581310.pdf
Montréal	UdeM (2017), « Déclaration de Montréal pour un développement responsable de l'intelligence artificielle », Université de Montréal, https://www.declarationmontreal-iaresponsable.com/
Nadella	Nadella S. (2016), « The Partnership of the Future », 28 juin, <i>Slate</i> , www.slate.com/articles/technology/future_tense/2016/06/microsoft_ceo_satya_nadella_humans_and_a_i_can_work_together_to_solve_society.html
PAI	PAI (2016), « TENETS », Partenariat sur l'IA, www.partnershiponai.org/tenets/
Polonski	Polonski V. (2018), « The Hard Problem of AI Ethics: Three Guidelines for Building Morality into Machines », 28 février, Forum Network on Digitalisation and Trust, www.oecd-forum.org/users/80891-dr-vyacheslav-polonski/posts/30743-the-hard-problem-of-ai-ethics-three-guidelines-for-building-morality-into-machines
Taddeo et Floridi	Taddeo M. et L. Floridi (2018), « How IA Can Be a Force for Good », <i>Science</i> , 24 août, vol. 61/6404, p. 751-752, http://science.sciencemag.org/content/361/6404/751
The Public Voice Coalition	UGAI (2018), « Universal Guidelines on Artificial Intelligence », The Public Voice Coalition, https://thepublicvoice.org/ai-universal-guidelines/
Déclaration de Tokyo	Next Generation Artificial Intelligence Research Center (2017), « The Tokyo Statement – Co-operation for Beneficial AI », www.ai.u-tokyo.ac.jp/tokyo-statement.html
Twomey	Twomey P. (2018), « Toward a G20 Framework for Artificial Intelligence in the Workplace », <i>CIGI Papers</i> , n° 178, Centre pour l'innovation dans la gouvernance internationale, www.cigionline.org/sites/default/files/documents/Paper%20No.178.pdf
UNI	UNI Global Union (2017), « Top 10 Principles for Ethical Artificial Intelligence », www.thefutureworldofwork.org/media/35420/uni_ethical_ai.pdf

Initiatives nationales

Tour d'horizon des politiques nationales en matière d'IA

De nombreux pays ont annoncé l'adoption de stratégies et de politiques nationales en matière d'IA, qui visent généralement à leur assurer un rôle de leadership dans le domaine de l'intelligence artificielle. Ces initiatives ont conduit à l'établissement de cibles et d'objectifs qui requièrent une action concertée de l'ensemble des parties prenantes. Les gouvernements servent fréquemment de coordinateurs et de facilitateurs à cet égard. L'Encadré 5.1 décrit les éléments qui apparaissent souvent dans les politiques et les mesures adoptées par les pouvoirs publics en vue de promouvoir la compétitivité nationale dans le domaine de l'IA. En outre, certains pays ont créé une entité publique spécifiquement chargée de s'occuper des questions d'éthique liées à l'IA et aux données, ou confié cette responsabilité à une entité existante.

Encadré 5.1. Comment les pays s'y prennent-ils pour développer un avantage compétitif dans le domaine de l'intelligence artificielle ?

Porter a identifié quatre éléments déterminants dans l'obtention d'un avantage compétitif dans un secteur d'activité particulier : i) les facteurs ; ii) la demande ; iii) les industries amont et apparentées ; et iv) la stratégie, la structure et la rivalité des entreprises. Il reconnaît que les acteurs à l'origine d'un avantage concurrentiel dans un secteur d'activité donné sont les entreprises. Cependant, il souligne le rôle essentiel des pouvoirs publics en ce qu'ils soutiennent ces quatre éléments au niveau des processus de développement industriel national.

- **Facteurs** : ce premier élément dépend de la localisation géographique, de l'existence d'une main-d'œuvre qualifiée, du niveau d'instruction et des capacités de recherche. Pour renforcer leurs capacités de recherche en IA, les pays recourent à différents types de mesures, par exemple : i) l'établissement d'instituts de recherche en IA ; ii) la création de nouveaux diplômes universitaires de troisième cycle et de doctorat dans le domaine de l'IA, et l'adaptation des cursus existants de manière à y inclure un module d'IA, en particulier dans les disciplines scientifiques ; et iii) les mesures visant à attirer des talents nationaux et étrangers, notamment en augmentant le nombre de visas accordés à des spécialistes de l'IA.
- **Demande** : plusieurs pays ont identifié des secteurs stratégiques dans l'optique du développement de l'IA, notamment les transports, la santé et les services publics. Ils mettent alors en place des mesures pour stimuler la demande intérieure de services d'IA dans ces secteurs. Pour ce qui est des services publics, certains pays s'assurent, par le biais des marchés publics, que les systèmes d'IA utilisés respectent des normes particulières, notamment en termes de précision ou de robustesse.
- **Industries amont et apparentées** : pour être concurrentiels, les systèmes d'IA doivent avoir accès aux infrastructures et services numériques, aux données, à des capacités informatiques et au haut débit. Un certain nombre de pays travaillent par conséquent au développement de pôles technologiques autour de l'IA et de structures de soutien pour les petites et moyennes entreprises (PME).
- **Stratégie, structure et rivalité des entreprises** : les pays recourent à diverses approches pour stimuler l'investissement privé et la concurrence dans le domaine de l'IA ; ils s'attachent notamment à : i) établir des feuilles de route pour le développement de l'IA, en vue de promouvoir l'investissement privé ; ii) encourager les entreprises internationales spécialisées dans l'IA à investir sur le territoire national, par exemple

via la création de laboratoires de recherche en IA ; et iii) tester des mesures expérimentales comme les « bacs à sable réglementaires » pour les applications de l'IA, afin d'inciter les entreprises à innover.

En outre, pour assurer une mise en œuvre effective de leurs initiatives nationales dans le domaine de l'IA, de nombreux pays réfléchissent à des mécanismes de gouvernance adéquats, qui garantiraient une approche coordonnée au niveau de l'ensemble de l'administration. La France, par exemple, a mis en place une fonction de coordination de l'IA au sein du cabinet du Premier ministre pour veiller à la mise en œuvre de la stratégie nationale en matière d'IA.

Source : Porter (1990^[4]), « The competitive advantage of nations », <https://hbr.org/1990/03/the-competitive-advantage-of-nations>.

Allemagne

Le gouvernement fédéral allemand a lancé sa stratégie pour l'IA en décembre 2018 (Allemagne, 2018^[5]). L'Allemagne entend devenir un centre majeur de l'IA en favorisant une transposition rapide et systématique des résultats de la recherche en applications, et en faisant de l'« IA made in Germany » une marque d'exportation forte et un gage de qualité reconnu au niveau mondial. Les mesures envisagées pour atteindre cet objectif comprennent : la création de nouveaux centres de recherche, le développement de la coopération franco-allemande dans le domaine de la recherche, le financement de pôles d'activité, et le soutien en faveur des PME. La stratégie nationale aborde également les besoins infrastructurels, l'amélioration de l'accès aux données, le développement des compétences, les normes de sécurité nécessaires pour prévenir l'utilisation abusive de l'IA et les questions éthiques.

En juin 2017, le ministère fédéral des Transports et des Infrastructures numériques a publié des directives éthiques sur les véhicules autonomes. Ces directives, élaborées par la commission d'éthique du ministère, prévoient une série de quinze règles applicables aux décisions programmées qui sont intégrées aux voitures sans conducteur. La commission a examiné de manière approfondie les questions éthiques en jeu, y compris celle des vies à préserver en priorité (le « dilemme du tramway »). Les directives stipulent ainsi que les véhicules autonomes doivent être programmés de manière à considérer toutes les vies humaines sur un pied d'égalité. Lorsqu'un choix s'impose, la décision du véhicule autonome doit porter sur le choc le moins violent, indépendamment de l'âge, de l'origine raciale ou du genre de la personne heurtée. En outre, la commission a clairement énoncé qu'un individu ne doit en aucun cas être contraint de se sacrifier pour sauver d'autres vies (Allemagne, 2017^[6]).

Argentine

Le gouvernement argentin prévoit de lancer en juillet 2019 une Stratégie nationale en matière d'IA, qui s'étendra sur une période de dix ans. Cette décision fait suite à une phase d'évaluation menée en 2018 par le ministère de la Science, de la Technologie et de l'Innovation productive dans le cadre de deux initiatives nationales, baptisées *Agenda Digital Argentina 2030* et *Plan Argentina Innovadora 2030*. La Stratégie nationale pour l'IA porte notamment sur les domaines prioritaires suivants : l'éducation et le développement des compétences, les données, la R-D et l'innovation, l'infrastructure de superinformatique, les mesures visant à faciliter la mobilité professionnelle et le soutien de la coopération public-privé en matière d'utilisation des données. Les services publics et le secteur manufacturier sont également considérés comme des secteurs cibles prioritaires pour le développement de l'IA. La stratégie est organisée autour de plusieurs thèmes transversaux : i) l'investissement, l'éthique et la réglementation ; ii) la communication et la sensibilisation ; et iii) la coopération internationale.

La stratégie, qui fait intervenir sept ministères, prévoit la création d'une plateforme nationale d'innovation en matière d'IA pour la mise en œuvre de projets dans chacun des domaines thématiques. Dans chaque domaine prioritaire, un groupe de pilotage composé d'experts sera chargé de fixer des objectifs et de définir des indicateurs afin de mesurer les progrès accomplis.

Arabie saoudite

L'Arabie saoudite a annoncé en 2016 un plan de réforme économique à l'horizon 2030 (*Vision 2030*). Ce plan vise à stimuler le développement de nouveaux secteurs d'activité et la diversification de l'économie, faciliter l'émergence de modèles économiques mixtes (public-privé) et réduire la dépendance du pays à l'égard des revenus pétroliers. Il envisage la transformation numérique comme un levier essentiel de développement de l'économie reposant à la fois sur l'exploitation des données, l'IA et l'automatisation industrielle. Les secteurs prioritaires, notamment dans la perspective de la création de centres d'innovation, sont : la santé, les services publics, l'énergie durable et l'eau, le secteur manufacturier, ainsi que celui de la mobilité et des transports. Le gouvernement prépare une stratégie nationale en matière d'IA ayant pour but de construire un écosystème innovant et éthique de l'IA en Arabie saoudite d'ici à 2030.

L'Arabie saoudite s'efforce de mettre en place des conditions propices au développement de l'IA, en misant en particulier sur le déploiement du très haut débit et des réseaux 5G, l'accès aux données et la sécurité. Le pays encourage également l'adoption précoce des concepts et méthodes d'IA par le biais de plusieurs projets de villes intelligentes devant servir de catalyseurs à la mise en place de solutions nouvelles. Ces initiatives s'inscrivent dans le sillage du mégaprojet de ville intelligente NEOM, lancé en 2017 avec un investissement important de 1 800 milliards SAR (500 milliards USD). L'Arabie saoudite participe aussi activement aux débats internationaux sur les cadres de gouvernance de l'IA.

Australie

Le gouvernement australien a alloué, au titre du budget 2018-19, plus de 28 millions AUD (21 millions USD) au développement des capacités d'IA et à l'aide au développement responsable de l'IA en Australie. Ces fonds permettront de financer : i) les projets du Centre de recherche coopérative (Co-operative Research Centre) portant spécifiquement sur l'IA (18 millions USD) ; ii) des bourses doctorales dans le domaine de l'IA (1 million USD) ; iii) la création de ressources en ligne pour l'enseignement de l'IA dans les établissements scolaires (1.1 million USD) ; iv) l'établissement d'une feuille de route sur les technologies liées à l'IA en vue d'examiner les incidences de l'IA sur les différents secteurs, les opportunités et les défis pour la main-d'œuvre, et les implications en termes d'éducation et de formation (250 000 USD) ; v) l'élaboration d'un cadre éthique à partir d'études de cas (367 000 USD) ; et vi) l'établissement d'une feuille de route sur les normes dans le domaine de l'intelligence artificielle (72 000 USD, avec un apport équivalent de l'industrie).

Le ministère de l'Industrie, de l'Innovation et de la Science mène également des projets ayant trait à l'IA. Il a en particulier chargé le Conseil australien des sociétés savantes (Australian Council of Learned Academies) d'examiner les opportunités, les risques et les conséquences que pourrait entraîner l'introduction à grande échelle de l'IA en Australie pendant la prochaine décennie. La Commission australienne des droits de l'homme (Australian Human Rights Commission) a pour sa part lancé en juillet 2018 un projet de grande ampleur pour étudier la relation entre droits de l'homme et technologies. La publication d'un document de réflexion et l'organisation d'une conférence internationale sont notamment prévues dans ce contexte ; le rapport final est prévu pour 2019-20.

Brésil

La stratégie de transformation numérique du Brésil *E-Digital*, lancée en mars 2018, a pour but d'harmoniser et de coordonner les différentes initiatives gouvernementales introduites dans le domaine du numérique en vue de progresser dans la réalisation des Objectifs de développement durable au Brésil. Pour ce qui concerne spécifiquement l'IA, elle prévoit des « actions stratégiques » visant à « évaluer l'impact économique et social potentiel de (...) l'intelligence artificielle et des données massives, et proposer des mesures pour en limiter les effets négatifs et optimiser les résultats positifs » (Brésil, 2018^[7]). La stratégie *E-Digital* donne par ailleurs la priorité à l'allocation de ressources à la recherche, au développement et à l'innovation dans le domaine de l'intelligence artificielle, ainsi qu'au renforcement des capacités d'IA. Le Brésil prévoit de lancer en 2019 une stratégie dédiée à l'IA. Il participe activement aux débats internationaux sur les normes techniques et les politiques d'IA.

Entre 2014 et début 2019, le ministère de la Science, de la Technologie, de l'Innovation et des Communications a soutenu au moyen de mesures d'incitation et d'aides financières 16 projets d'IA et 59 startups spécialisées dans l'intelligence artificielle. En outre, 39 initiatives visent à utiliser l'IA pour les fonctions d'administration électronique au niveau fédéral. L'objectif est notamment d'améliorer les procédures administratives et d'évaluation, en particulier dans le domaine de l'état-civil, des services sociaux et de la publication des offres d'emploi. Un nouvel institut de recherche en IA – *Instituto Avançado para Inteligência Artificial* – a été créé en 2019. Il a pour mission de promouvoir les partenariats entre universités et entreprises autour de projets conjoints de R-D et d'innovation en matière d'IA. Il s'occupera spécifiquement de domaines comme l'agriculture, les villes intelligentes, la gouvernance du numérique, l'infrastructure, l'environnement, les ressources naturelles, et la sécurité et la défense.

Canada

Le Canada cherche à se positionner comme un pays leader dans le domaine de l'IA, avec notamment la Stratégie pancanadienne en matière d'intelligence artificielle, lancée en mars 2017 (CIFAR, 2017^[8]). Cette stratégie, qui est placée sous la direction de l'Institut canadien de recherches avancées (CIFAR), un organisme à but non lucratif, a bénéficié d'un financement public de 125 millions CAD (100 millions USD). Ces fonds sont destinés à soutenir, sur une période de cinq ans, des projets visant à : développer le capital humain du Canada, appuyer la recherche en IA au Canada et faire en sorte que la recherche en IA aboutisse à des applications pour les secteurs public et privé. La Stratégie pancanadienne en matière d'intelligence artificielle sert quatre grands objectifs :

1. Accroître le nombre de chercheurs et diplômés qualifiés dans le domaine de l'intelligence artificielle au Canada.
2. Établir des centres d'excellence scientifique interconnectés dans les trois grands instituts canadiens spécialisés dans l'intelligence artificielle, situés à Edmonton (Alberta Machine Intelligence Institute, Amii), Montréal (Institut des Algorithmes d'Apprentissage de Montréal, MILA) et Toronto (Institut Vecteur pour l'intelligence artificielle).
3. Développer un programme mondial sur l'intelligence artificielle dans la société et mener la réflexion, à l'échelle internationale, sur les répercussions économiques, sociales, éthiques, politiques et juridiques des progrès de l'intelligence artificielle.
4. Soutenir une communauté de recherche nationale en intelligence artificielle.

Le gouvernement fédéral, via le Conseil national de recherches Canada (CNRC), prévoit d'investir en tout 50 millions CAD (40 millions USD) sur une période de sept ans dans la recherche sur l'application de l'IA dans un certain nombre de domaines programmatiques clés, notamment l'analytique des données, l'IA et le design, la cybersécurité, les langues autochtones du Canada, le soutien aux « supergrappes » fédérales et aux centres de collaboration avec les universités canadiennes, ainsi que les partenariats stratégiques avec des acteurs internationaux.

Outre ce financement fédéral, le gouvernement du Québec a alloué 100 millions CAD (80 millions USD) au secteur de l'IA à Montréal, et l'Ontario, 50 millions CAD (40 millions USD) à l'Institut Vecteur pour l'intelligence artificielle. En 2016, le Fonds d'excellence en recherche Apogée Canada a alloué 93.6 millions CAD (75 millions USD) à trois universités – l'Université de Montréal, Polytechnique Montréal et HEC Montréal – pour la recherche de pointe sur l'apprentissage profond. Facebook et d'autres entreprises privées dynamiques comme Element AI sont actives au Canada.

Le gouvernement du Québec envisage de mettre sur pied un observatoire mondial sur les impacts sociaux de l'IA et des technologies numériques (Fonds de recherche du Québec, 2018^[9]). Un atelier a été organisé en mars 2018 pour commencer à réfléchir à la mission et à l'organisation de cet observatoire, à son mode de gouvernance et à son financement, ainsi qu'aux modalités de la coopération internationale et aux secteurs et domaines d'intérêt à prendre en compte. Le gouvernement du Québec a alloué une enveloppe de 5 millions CAD (3.7 millions USD) pour soutenir la mise en œuvre de cet observatoire.

Le Canada collabore également avec divers partenaires au niveau international pour mener à bien des initiatives dans le domaine de l'IA. En juillet 2018, par exemple, les gouvernements du Canada et de la France ont annoncé leur intention de travailler ensemble à la création d'un nouveau Groupe international d'experts en intelligence artificielle (G2IA). Ce groupe aura pour mission de soutenir et de guider l'adoption responsable d'une IA centrée sur l'humain et axée sur les droits de l'homme, l'inclusion, la diversité, l'innovation et la croissance économique.

Chine

En mai 2016, le gouvernement chinois a rendu public un plan national pour l'IA d'une durée de trois ans, établi conjointement par la Commission nationale pour le développement et la réforme, le ministère de la Science et de la Technologie, le ministère de l'Industrie et des Technologies de l'information et l'Administration du cyberspace de Chine. L'initiative *Internet Plus*, lancée en 2015 en tant que stratégie nationale pour stimuler la croissance économique au moyen des technologies innovantes liées à l'internet pendant les années 2016-18, englobe l'IA (Jing et Dai, 2018^[10]). Cette initiative porte principalement sur : i) le renforcement des capacités matérielles d'IA ; ii) la consolidation des écosystèmes de plateformes de réseau ; iii) les applications de l'IA dans les domaines socio-économiques importants ; et iv) l'impact de l'IA sur la société. Le gouvernement chinois prévoyait dans ce cadre la création, d'ici à 2018, d'un marché de 15 milliards USD par le biais de la R-D pour le secteur de l'IA en Chine (Chine, 2016^[11]).

Mi-2017, le Conseil des affaires d'État de la Chine a publié des Lignes directrices relatives au Plan de développement de l'IA de nouvelle génération, qui définissent les perspectives à long terme de la Chine en matière d'IA, avec des objectifs industriels échelonnés en trois temps, selon les modalités suivantes : i) parvenir, d'ici à 2020, à une croissance économique tirée par l'IA ; ii) réaliser, d'ici à 2025, des progrès majeurs au niveau des théories fondamentales, ainsi que dans le développement d'une société intelligente ; et iii) faire du

pays, à l'horizon 2030, un centre d'innovation mondial en IA et bâtir un secteur de l'IA de 1 000 milliards CNY (150 milliards USD) (Chine, 2017_[12]). La mise en œuvre du plan semble progresser dans l'ensemble des institutions gouvernementales et la Chine a acquis un certain leadership dans le domaine de l'IA grâce au soutien de l'État et au dynamisme des entreprises privées. Les objectifs fixés par le Conseil des affaires d'État prévoient que les « technologies d'information de nouvelle génération » devront représenter, en tant que secteur stratégique, 15 % du produit intérieur brut d'ici à 2020.

Le 13^e plan quinquennal (2016-20) affirme l'ambition de la Chine de devenir un leader dans le domaine de la science et de la technologie et inclut seize « mégaprojets d'innovation scientifique et technologique à l'horizon 2030 ». L'un d'eux, dénommé *IA 2.0*, a donné un élan à l'action dans le secteur public (Kania, 2018_[13]). Il incite également les entreprises à intensifier leurs activités de R-D concernant le matériel et les logiciels d'IA, en particulier dans les domaines de la reconnaissance visuelle, vocale et biométrique, des interfaces homme-machine et des systèmes de contrôle intelligents.

Le 18 janvier 2018, la Chine a créé un groupe national pour la normalisation de l'IA et un groupe consultatif national d'experts en IA. Au même moment, le comité national de normalisation auprès du ministère de l'Industrie a publié un livre blanc sur la normalisation dans le domaine de l'intelligence artificielle. Ce document a été élaboré avec le soutien de l'Institut national de normalisation électronique (une division du ministère de l'Industrie et des Technologies de l'information) (Chine, 2018_[14]).

Les entreprises privées chinoises avaient commencé à manifester un intérêt pour l'IA avant les mesures d'aide adoptées il y a peu au niveau gouvernemental. Des entreprises chinoises comme Baidu, Alibaba et Tencent ont engagé des efforts et des investissements importants dans le domaine de l'IA. L'industrie chinoise s'intéresse tout particulièrement aux applications et à l'intégration des données, tandis que le gouvernement central focalise son attention sur les algorithmes de base, les données ouvertes et le travail conceptuel. Les municipalités s'intéressent quant à elles à l'utilisation qui peut être faite des applications et des données ouvertes dans le contexte urbain.

Corée

Le gouvernement coréen a publié en mars 2016 une Stratégie pour le développement du secteur des systèmes d'information intelligents. Il a annoncé des investissements publics d'une enveloppe de 1 000 milliards KRW (940 millions USD) d'ici à 2020 dans le domaine de l'IA et des technologies de l'information connexes comme l'internet des objets et l'infonuagique. Le but de la stratégie est de créer un nouvel écosystème à l'échelle du secteur des systèmes d'information intelligents et d'encourager l'investissement privé à hauteur de 2 500 milliards KRW (2.3 milliards USD) d'ici à 2020. La stratégie du gouvernement coréen sert trois objectifs : premièrement, le lancement de projets phares pour le développement de l'IA, consacrés par exemple aux technologies intelligentes dans le domaine langagier, visuel, spatial ou émotionnel ; deuxièmement, le renforcement des compétences de la main-d'œuvre liées à l'IA ; troisièmement, la promotion de l'accès aux données et de leur utilisation par les pouvoirs publics, les entreprises et les instituts de recherche (Corée, 2016_[15]).

En décembre 2016, le gouvernement coréen a publié un « Plan d'action national pour une société de l'intelligence et de l'information à moyen et long terme ». Ce plan décrit les politiques nationales à mettre en place pour se préparer aux changements et aux enjeux induits par la quatrième révolution industrielle. Il jette aussi les bases du développement de technologies informatiques intelligentes de niveau mondial dans la perspective d'une « société intelligente centrée sur l'humain ». Ces technologies pourraient être introduites

dans tous les secteurs d'activité et être utilisées pour moderniser les politiques sociales. Pour mettre en œuvre ce plan d'action, le gouvernement s'appuie sur des bancs d'essai à grande échelle qui aideront au développement de nouveaux produits et services et, en particulier, à l'amélioration des services publics (Corée, 2016^[16]).

En mai 2018, le gouvernement coréen a lancé un plan national d'investissement de 2 200 milliards KRW (2 milliards USD) d'ici à 2022 en vue de renforcer les capacités de R-D dans le domaine de l'intelligence artificielle. Ce plan prévoit la création de six instituts de recherche en IA, le développement des talents au moyen de 4 500 bourses en IA et de cours de formation intensive à court terme, et l'accélération du développement des puces d'IA (Peng, 2018^[17]).

Danemark

Le Danemark a publié en janvier 2018 une Stratégie nationale pour la croissance numérique. Cette stratégie vise à faire du pays un chef de file du numérique, en permettant à l'ensemble des Danois de tirer parti de la transformation. Elle prévoit un certain nombre d'initiatives pour aider à saisir les opportunités qu'offrent l'IA, les données massives et l'internet des objets (IdO) en termes de croissance. Les principales orientations stratégiques définies dans ce cadre sont les suivantes : i) créer une plateforme numérique pour les partenariats public-privé ; ii) aider les PME à avancer sur la voie de la numérisation et à développer des activités commerciales reposant sur l'exploitation des données ; iii) établir des établissements d'enseignement dans le cadre d'un « pacte technologique » visant à favoriser le développement des compétences techniques et numériques ; iv) renforcer la cybersécurité dans les entreprises ; et v) mettre en place une réglementation souple afin de faciliter l'expérimentation et le développement de nouveaux modèles économiques. Le gouvernement danois s'est engagé à consacrer 1 milliard DKK (160 millions USD) jusqu'en 2025 à la mise en œuvre de cette stratégie. Ces fonds ont été répartis comme suit : 75 millions DKK (12 millions USD) en 2018 et 125 millions DKK (20 millions USD) pendant la période 2019-25. La plus grande part de ce budget sera affectée à des initiatives de développement des compétences, puis à la création de la plateforme numérique et à l'aide aux PME (Danemark, 2018^[18]).

Estonie

L'Estonie prépare une nouvelle étape de son système d'administration électronique, qui reposera sur l'intelligence artificielle, afin de réduire les coûts et d'améliorer l'efficacité. Elle teste également des services de santé en ligne et de conscience situationnelle. Le but est d'améliorer le quotidien des citoyens et la vie urbaine, et de promouvoir les valeurs humaines. Sur le plan du contrôle de la mise en œuvre, l'Estonie privilégie des valeurs fondamentales telles que l'éthique, la responsabilité, l'intégrité et la redevabilité, plutôt que de se concentrer sur des technologies en rapide évolution, et la mise en place d'un système de contrôle fondé sur la technologie des « chaînes de blocs » afin d'atténuer les risques en matière d'intégrité et de redevabilité. Un projet pilote est prévu en 2018.

Des véhicules autonomes sont testés sur les axes routiers publics de l'Estonie depuis mars 2017 dans le cadre du projet *StreetLEGAL*. L'Estonie est aussi le premier pays à réfléchir à la possibilité de conférer un statut juridique à certains systèmes reposant sur l'intelligence artificielle, par exemple en accordant à des algorithmes des droits de représentation – et des responsabilités – pour l'achat et la vente de services au nom de leurs propriétaires. En 2016, le gouvernement estonien a créé un groupe de travail pour examiner les problèmes de redevabilité associés aux algorithmes d'apprentissage automatique et les besoins de réglementation de l'IA, en collaboration avec le ministère des Affaires économiques et des Communications et le Bureau du gouvernement (Kaevats, 2017^[19] ; Kaevats, 25 septembre 2017^[20]).

États-Unis

Le président Trump a signé le 11 février 2019 le décret présidentiel n° 13859 sur « le maintien du leadership des États-Unis dans le domaine de l'intelligence artificielle », qui institue une nouvelle stratégie fédérale, l'*American AI Initiative*. Cette stratégie est organisée autour de cinq axes clés : i) l'investissement dans la R-D en IA ; ii) l'élargissement de l'accès aux ressources informatiques aux fins de la R-D en IA ; iii) la définition d'orientations en vue de la réglementation et de l'établissement de normes techniques pour les applications de l'IA ; iv) le développement d'une main-d'œuvre qualifiée en IA ; et v) l'action internationale à l'appui de la recherche et de l'innovation américaines en IA et de l'ouverture de marchés aux entreprises américaines d'IA.

La nouvelle stratégie vient couronner toute une série d'initiatives engagées par l'administration fédérale pour accroître le leadership des États-Unis dans le domaine de l'IA. La Maison Blanche a accueilli en mai 2018 le premier Sommet sur l'intelligence artificielle, qui a réuni des acteurs de l'industrie, des universitaires et des dirigeants gouvernementaux. Les participants à ce sommet ont débattu de l'importance de lever les obstacles à l'innovation en IA dans le pays et de promouvoir la collaboration en matière de R-D dans ce domaine avec les partenaires des États-Unis. Ils ont évoqué la nécessité de sensibiliser le public à l'émergence de l'IA, afin que celui-ci soit à même de mieux comprendre le fonctionnement de ces technologies et les avantages qu'elles offrent au quotidien. Le même mois, l'exécutif américain a publié sous le titre *Artificial intelligence for the American people* (EOP, 2018^[21]) une fiche d'information qui énumère les mesures et politiques adoptées par l'administration actuelle dans le domaine de l'IA, notamment : l'augmentation du financement public de la R-D en IA ; la réforme réglementaire pour faciliter le développement et l'utilisation des drones et des véhicules autonomes ; la priorité donnée à l'enseignement des sciences, des technologies, de l'ingénierie et des mathématiques (STIM), avec un accent particulier sur l'informatique ; et l'amélioration de l'accès aux données fédérales aux fins de la recherche et des applications de l'IA.

Dans les budgets fédéraux de R-D pour les exercices 2019 et 2020, l'intelligence artificielle et l'apprentissage automatique sont définis comme des priorités essentielles. La recherche fondamentale en IA menée à la Fondation nationale des sciences et les activités de R-D appliquée poursuivies au ministère des Transports y sont spécifiquement incluses. Les priorités de recherche englobent également le développement de méthodes analytiques de pointe en matière de santé aux National Institutes of Health et d'une infrastructure informatique pour l'IA au ministère de l'Énergie. Globalement, les investissements du gouvernement fédéral dans la R-D non classifiée liée à l'IA et aux technologies connexes ont augmenté de plus de 40 % depuis 2015.

En septembre 2018, la commission spéciale sur l'intelligence artificielle du Conseil national de la science et de la technologie a commencé à actualiser le Plan stratégique national pour la recherche et le développement en matière d'intelligence artificielle. En effet, depuis la publication de ce plan en 2016, la technologie sous-jacente, les cas d'usage et le déploiement commercial de l'IA ont connu un développement rapide. La commission spéciale sollicite l'avis du public sur les améliorations à apporter au plan, en particulier celui des acteurs directement impliqués dans la réalisation d'activités de R-D en IA ou concernés par les résultats qui en découlent.

Le gouvernement américain accorde aussi la priorité à la formation de la future main-d'œuvre des États-Unis. Le président Trump a signé un décret présidentiel sur l'établissement de programmes d'apprentissage reconnus par les entreprises, qui prévoit la création d'un groupe de travail ministériel sur le développement de l'apprentissage (*Task Force on*

Apprenticeship Expansion). Dans le droit fil des politiques décrites dans la fiche d'information mentionnée plus haut, une note présidentielle affirme également la priorité d'assurer un enseignement de qualité dans les STIM, et tout particulièrement en informatique. Cette note annonce l'affectation de 200 millions USD sous forme de subventions, auxquels est venu s'ajouter un engagement du secteur privé à hauteur de 300 millions USD.

Le Congrès américain a créé en mai 2017 un groupe parlementaire mixte sur l'intelligence artificielle, qui est coprésidé par deux parlementaires, John K. Delaney et Pete Olson (US, 2017^[22]). Ce groupe organise des réunions avec des experts des milieux universitaires, du gouvernement et du secteur privé pour débattre des incidences des technologies liées à l'IA. Le Congrès envisage d'adopter un texte de loi qui établira un comité consultatif fédéral sur l'IA et instituera des normes fédérales de sécurité pour les véhicules autonomes.

Fédération de Russie

Le gouvernement russe a lancé en juillet 2017 une stratégie numérique intitulée « L'économie numérique de la Fédération de Russie ». Cette stratégie donne la priorité à l'exploitation du développement de l'IA, notamment en instaurant des conditions juridiques propices pour faciliter les activités de R-D. Elle prévoit aussi l'adoption de mesures pour inciter les entreprises d'État à participer aux communautés de recherche existant à l'échelon national (centres de compétences). Elle promeut en outre l'établissement de normes nationales pour les technologies d'IA (Russie, 2017^[23]). Avant la publication de cette stratégie numérique, le gouvernement russe avait investi dans divers projets d'IA et mis en place des dispositifs en vue de la création de partenariats public-privé. L'association russe des concepteurs de systèmes d'intelligence artificielle encourage la coopération entre universités et entreprises, afin de faciliter les transferts de technologie vers les entreprises.

Finlande

La Finlande entend développer une société sûre et démocratique s'appuyant sur l'IA, offrir les meilleurs services publics du monde, et faire de l'IA un levier de prospérité, de croissance et de productivité renouvelées pour ses citoyens. La stratégie nationale d'IA publiée en octobre 2017 sous le titre « L'ère de l'intelligence artificielle en Finlande » est une feuille de route décrivant comment mettre à profit à la fois le niveau d'instruction de la population, la transformation numérique déjà bien avancée du pays et les ressources que constituent les données du secteur public. Cette stratégie prévoit aussi le développement de liens internationaux dans le domaine de la recherche et de l'investissement, et l'adoption de mesures pour encourager l'investissement privé. La Finlande espère doubler, d'ici à 2035, la croissance économique nationale grâce à l'IA. Huit mesures clés ont pour but d'améliorer la croissance, la productivité et le bien-être en s'appuyant sur l'IA : i) accroître la compétitivité des entreprises ; ii) favoriser l'utilisation des données dans tous les secteurs ; iii) accélérer et simplifier l'adoption de l'IA ; iv) développer une expertise de haut niveau ; v) prendre des décisions audacieuses, notamment en matière d'investissement ; vi) développer les meilleurs services publics du monde ; vii) établir de nouveaux modèles de coopération ; et viii) faire de la Finlande un pays pionnier à l'ère de l'IA. La stratégie nationale met particulièrement l'accent sur l'utilisation de l'IA pour améliorer les services publics. Le service de l'immigration finlandais utilise par exemple le réseau robotique national de service à la clientèle dénommé Aurora pour fournir des services de communication multilingues (Finlande, 2017^[24]).

Le gouvernement finlandais a également créé en février 2018 un organisme de financement de la recherche en intelligence artificielle et des projets commerciaux faisant appel à l'IA. Cet organisme sera chargé d'allouer 200 millions EUR (235 millions USD) au secteur

privé, dont les PME, sous forme d'incitations et de subventions. Selon les autorités nationales, environ 250 entreprises travaillent au développement de l'IA en Finlande. Dans le secteur de la santé, notamment, le déploiement de l'IA changera la donne pour les patients et les professionnels, ainsi que pour les organisations du secteur et le système de santé, et s'accompagnera de réformes approfondies (Sivonen, 2017^[25]). Par ailleurs, il est prévu d'étendre le rôle du Centre national de recherche technique, financé par l'État, et de l'Agence finlandaise de financement de la technologie et de l'innovation.

France

Le président Emmanuel Macron a annoncé la stratégie française pour l'IA le 29 mars 2018. Cette stratégie prévoit d'allouer au secteur de l'IA un financement public de 1.5 milliard EUR d'ici à 2022 pour aider la France à devenir un leader de la recherche et de l'innovation dans ce domaine. Les mesures envisagées s'appuient en grande partie sur les recommandations formulées dans le rapport du député Cédric Villani (Villani, 2018^[26]). La stratégie appelle à investir dans la recherche et l'enseignement publics, à développer des plateformes de recherche de niveau mondial liées à l'industrie au moyen de partenariats public-privés et d'augmenter l'attractivité de la France pour les chercheurs talentueux en IA expatriés ou étrangers. Pour consolider l'écosystème français de l'IA, la stratégie envisage de « moderniser » les secteurs d'activité existants. Il s'agira ainsi, en introduisant d'abord des applications dans les domaines de la santé, de l'environnement, des transports et de la défense, de renouveler grâce à l'IA l'ensemble des secteurs d'activité. La stratégie propose également d'améliorer en priorité l'accès aux données en créant des « communs de données » entre acteurs privés et publics, de réformer le droit d'auteur de manière à faciliter les pratiques d'exploration de données, et d'ouvrir l'accès aux données du secteur public, notamment les données de santé, à des partenaires privés.

La stratégie française décrit également les mesures initiales à prendre à l'égard des perturbations induites par l'intelligence artificielle, en prenant fermement position sur la question des transferts de données hors d'Europe (Thompson, 2018^[27]). Elle envisage la création d'un pôle central des données, doté d'une équipe d'une trentaine d'experts qui rempliraient un rôle de conseil en IA auprès de l'ensemble des administrations publiques. Sur le plan éthique et philosophique, la stratégie met en avant un principe fondamental, celui de la transparence des algorithmes. Les algorithmes développés par des administrations publiques françaises ou financés sur des fonds publics, par exemple, devront être ouverts. Le respect de la vie privée et des droits de l'homme devra être intégré par défaut, dès la conception des systèmes d'IA. La stratégie envisage aussi le développement de la formation professionnelle dans les métiers menacés par l'IA. Elle appelle à expérimenter des mesures sur le marché du travail et à engager le dialogue sur les moyens de répartir la valeur ajoutée générée par l'IA sur l'ensemble de la chaîne de valeur. En outre, un rapport a été publié fin mars 2018, intitulé *Intelligence artificielle et travail* (Benhamou et Janin, 2018^[28]).

Hongrie

La Hongrie a lancé en octobre 2018 une « alliance pour l'intelligence artificielle », sous la forme d'un partenariat entre administrations publiques, entreprises informatiques de pointe et universités. Cette alliance est chargée de définir une stratégie visant à faire de la Hongrie un pays innovant dans le domaine de l'IA et d'étudier les incidences économiques et sociales de l'IA sur la société. Elle regroupe environ 70 instituts de recherche universitaires, entreprises et organismes publics et est appelée à devenir un forum de coopération transversale dans le domaine de la R-D en IA. L'Université polytechnique et économique de Budapest et l'Université Loránd Eötvös font partie du consortium prévoyant d'investir 20 millions EUR

(23.5 millions USD) dans le projet AI4EU, qui vise à créer une plateforme à la demande pour l'intelligence artificielle en Europe.

Inde

L'Inde a rendu publique sa stratégie d'IA en juin 2018. Cette stratégie énonce un certain nombre de recommandations visant à faire de l'Inde un pays de pointe dans ce domaine grâce à la mobilisation des ressources humaines et à l'impulsion d'une croissance inclusive bénéficiant à l'ensemble de la société. Cette approche inclusive a été baptisée #AIFORALL (« L'IA POUR TOUS »). Les domaines suivants sont désignés comme prioritaires pour l'introduction d'applications reposant sur l'intelligence artificielle : la santé, l'agriculture, l'éducation, les villes intelligentes et les transports. La stratégie souligne que le faible niveau des capacités de recherche et l'absence d'écosystèmes de données en Inde constituent des obstacles à la réalisation du plein potentiel de l'IA. Elle recommande, en conséquence, la création en Inde d'instituts de recherche à deux niveaux (pour la recherche fondamentale et la recherche appliquée), la mise en place de pôles de formation pour la main-d'œuvre actuelle, et le développement d'ensembles de données ciblés et de pépinières de startups. Enfin, elle juge indispensable l'établissement d'un cadre réglementaire sur la protection des données et la cybersécurité (Inde, 2018^[29]).

Italie

L'Italie a publié en mars 2018 un livre blanc sur « L'intelligence artificielle au service des citoyens », qui a été rédigé par un groupe de travail de l'agence italienne pour le numérique. Ce document examine les moyens pour l'administration publique de mettre les technologies d'IA au service des citoyens et des entreprises, et de l'amélioration de l'efficacité des services publics et du niveau de satisfaction des usagers. Il recense les obstacles à surmonter pour introduire l'IA dans les services publics, que ce soit en termes d'éthique, de technologie, de disponibilité des données et d'évaluation d'impact. Le livre blanc formule également des recommandations en vue : de l'établissement d'une plateforme nationale pour les données étiquetées, les algorithmes et les modèles d'apprentissage ; du développement des compétences ; et de la création d'un centre national de compétences et d'un centre transdisciplinaire sur l'IA. Il appelle en outre à élaborer des lignes directrices et des processus pour parvenir à un niveau de contrôle plus élevé et faciliter l'échange de données entre pays européens sur les cyberattaques exploitant les systèmes d'intelligence artificielle ou les visant (Italie, 2018^[30]).

Japon

Le Bureau du Premier ministre japonais a créé en avril 2016 un Conseil de la stratégie des technologies liées à l'IA, afin de promouvoir la R-D et les applications commerciales dans le domaine de l'intelligence artificielle. Ce conseil a publié en mars 2017 une Stratégie pour les technologies d'intelligence artificielle, qui recensait un certain nombre d'enjeux décisifs pour le Japon, en particulier la nécessité d'accroître l'investissement, de faciliter l'accès aux données et leur utilisation et d'augmenter le nombre de chercheurs et d'ingénieurs en IA. Ce document identifiait également les domaines stratégiques où l'IA pourrait avoir d'importantes retombées positives, à savoir : la productivité ; la santé, les soins médicaux et les services sociaux ; la mobilité ; et la sécurité de l'information (Japon, 2017^[31]).

La Stratégie intégrée pour l'innovation du Japon, publiée par le Bureau du Premier ministre en juin 2018, est à l'origine d'un certain nombre d'initiatives publiques dans le domaine de l'IA (Japon, 2018^[32]). Elle prévoit notamment l'organisation de discussions multipartites sur les questions éthiques, juridiques et sociétales liées à l'IA. Ces discussions ont abouti à

la publication par le Bureau du Premier ministre en avril 2019 d'un document intitulé « Principes sociaux pour une intelligence artificielle centrée sur l'humain » (Japon, 2019^[33]).

Lors de la réunion des ministres des TIC du G7 à Takamatsu en avril 2016, le Japon a proposé d'élaborer des principes communs sur la R-D dans le domaine de l'intelligence artificielle. Un groupe d'experts réuni sous le nom de *Conference toward IA Network Society* a produit un « Projet de lignes directrices sur la R-D dans le domaine de l'IA en vue des discussions internationales », qui a été publié par le ministère japonais des Affaires intérieures et des Communications en juillet 2017. Ces lignes directrices ont essentiellement vocation à mettre en balance les avantages et les risques qui sont associés aux réseaux d'IA, tout en garantissant la neutralité technologique et en évitant d'imposer des contraintes excessives aux développeurs de ces réseaux. Elles définissent neuf principes dont doivent tenir compte les chercheurs et développeurs des systèmes d'IA (Japon, 2017^[34]). Le Tableau 5.2 ci-dessous donne un aperçu succinct des dispositions de ces lignes directrices. À l'issue de ses discussions, le groupe d'experts a également publié un « Projet de principes sur l'utilisation de l'IA » en juillet 2018 (Japon, 2018^[35]).

Tableau 5.2. Principes énoncés dans le projet de « Lignes directrices sur la R-D dans le domaine de l'IA en vue des discussions internationales »

Principes	Les développeurs devraient :
I. Collaboration	Veiller à l'interconnectivité et l'interopérabilité des systèmes d'IA.
II. Transparence	Garantir la vérifiabilité des entrées/sorties des systèmes d'IA et l'explicabilité de leurs décisions.
III. Contrôlabilité	Assurer la contrôlabilité des systèmes d'IA.
IV. Sûreté	Veiller, au moyen d'actionneurs ou d'autres dispositifs, à ce que les systèmes d'IA ne puissent porter atteinte à la vie, aux personnes ou aux biens des utilisateurs ou de tiers.
V. Sécurité	Garantir la sécurité des systèmes d'IA.
VI. Respect de la vie privée	Prendre des dispositions pour éviter que les systèmes d'IA ne violent la vie privée des utilisateurs ou de tiers.
VII. Éthique	Respecter la dignité humaine et l'autonomie individuelle dans la R-D sur les systèmes d'IA.
VIII. Assistance aux utilisateurs	Ne pas perdre de vue que les systèmes d'IA sont au service des utilisateurs et donner à ces derniers la possibilité de prendre des décisions appropriées.
IX. Responsabilité	Assumer leurs responsabilités à l'égard des différentes parties prenantes, y compris les utilisateurs des systèmes d'IA.

Source : Japon (2017^[34]), *Draft IA R-D guidelines for international discussions*, www.soumu.go.jp/main_content/000507517.pdf.

Mexique

Le Conseil national de la science et de la technologie du Mexique a créé en 2004 un Centre de recherche sur l'intelligence artificielle, qui dirige le développement des systèmes intelligents.

Un livre blanc a été publié en juin 2018 sous le titre « Vers une stratégie de l'IA au Mexique : exploiter la révolution de l'IA »¹. Ce rapport note que le Mexique occupe la 22^e place parmi les 35 pays de l'OCDE dans le classement établi par Oxford Insight sur la base d'un indice de préparation à l'intelligence artificielle (*AI Readiness Index*). Cet indice composite est en fait la moyenne des notes obtenues par les pays au regard de neuf indicateurs allant du niveau des compétences numériques à celui de l'innovation dans le secteur public. Le Mexique obtient de bonnes notes pour ses politiques en matière de données ouvertes et son infrastructure numérique, mais des notes plus médiocres s'agissant des compétences techniques, de la progression du numérique et de l'innovation dans le secteur public. Le rapport formule à l'intention des pouvoirs publics plusieurs recommandations pour la poursuite du développement et du déploiement de l'IA au

Mexique. Ces recommandations couvrent les cinq domaines suivants : l'administration et les services publics ; la R-D ; les capacités, les compétences et l'enseignement ; les données et l'infrastructure numérique ; et l'éthique et la réglementation (Mexico, 2018_[36]).

Norvège

La Norvège a pris un certain nombre de mesures en faveur de l'IA dans le cadre de son programme numérique national et du plan à long terme pour la recherche et l'enseignement supérieur, notamment :

- La création de plusieurs laboratoires d'intelligence artificielle, comme l'Open AI Lab de l'Université de science et de technologie. L'Open AI Lab, qui est parrainé par plusieurs entreprises, travaille dans les domaines de l'énergie, du transport maritime, de l'aquaculture, des télécommunications, des services bancaires en ligne, de la santé et de la biomédecine, où la Norvège occupe une forte position au niveau international.
- Le lancement d'un programme de réforme des compétences dénommé « Apprendre pour la vie », avec une proposition de budget de 130 millions NOK (16 millions USD) en 2019, afin d'aider la main-d'œuvre à acquérir ou améliorer ses compétences dans des domaines tels que l'intelligence artificielle ou les soins de santé.
- Le développement d'une stratégie d'ouverture des données en vertu de laquelle les administrations publiques sont tenues de permettre l'accès à leurs données dans un format exploitable par la machine via des API, et d'enregistrer dans un catalogue commun les ensembles de données disponibles.
- La mise en place d'une plateforme pour la formulation de lignes directrices et de principes éthiques régissant l'utilisation de l'IA.
- La réforme de la réglementation afin d'autoriser les tests de véhicules autonomes sur le réseau routier, y compris les tests de voitures sans conducteur.

Pays-Bas

Avec l'adoption en 2018 d'une Stratégie nationale de transformation numérique, le gouvernement des Pays-Bas a pris deux engagements : premièrement, celui d'optimiser les opportunités économiques et sociales qu'offre le numérique ; deuxièmement, celui de renforcer les conditions indispensables au développement du numérique comme le développement des compétences, la politique en matière de données, la confiance et la résilience, le respect des droits fondamentaux et l'éthique (en veillant par exemple à ce que les algorithmes n'influent pas sur l'autonomie des individus ni sur l'égalité de traitement), ainsi que la recherche et l'innovation en IA. En octobre 2018, l'AINED (« L'intelligence artificielle pour les Pays-Bas »), un partenariat regroupant des entreprises et des universitaires, a publié un document recensant les objectifs et les actions destinés à nourrir un plan national pour l'IA, à savoir par exemple promouvoir l'accès aux talents et compétences en IA, ainsi qu'aux données publiques à forte valeur ajoutée. L'AINED cherche également à faciliter le développement d'entreprises fondées sur l'IA et à promouvoir l'utilisation à grande échelle de l'IA dans les administrations publiques. Il envisage en outre l'établissement d'un cadre socio-économique et d'un cadre éthique pour l'IA, en encourageant la constitution de partenariats public-privé dans certains secteurs clés et au niveau de certaines chaînes de valeur essentielles, et en faisant des Pays-Bas un centre mondial de la recherche en IA. Le gouvernement néerlandais entend en outre finaliser un plan d'action stratégique pour l'IA à l'échelle de l'ensemble de l'administration avant mi-2019. Ce plan d'action doit tenir

compte du rapport de l'AINED, du plan coordonné de l'UE et des débats menés au sein du Groupe d'experts indépendants de haut niveau sur l'intelligence artificielle (AI HLEG) de la Commission européenne.

République tchèque

Le gouvernement tchèque a appelé en 2018 à la réalisation d'une étude sur la mise en œuvre de l'IA, afin de définir des objectifs stratégiques et de soutenir les négociations aux niveaux européen et international. Une équipe d'universitaires du Centre technologique de l'Académie des sciences, de l'Université technique de Prague et de l'Institut d'administration publique et de droit de l'Académie des sciences a produit un rapport intitulé « Analyse du développement potentiel de l'intelligence artificielle en République tchèque » (OGCR, 2018^[37]). Ce rapport examine : i) l'état actuel de la mise en œuvre de l'IA en République tchèque ; ii) l'impact potentiel de l'IA sur le marché du travail national ; et iii) les aspects éthiques, juridiques et réglementaires du développement de l'IA en République tchèque.

Royaume-Uni

La stratégie numérique du Royaume-Uni (*UK Digital Strategy*), publiée en mars 2017, reconnaît l'importance de l'IA pour la croissance de l'économie numérique britannique (RU, 2017^[38]). Cette stratégie prévoit d'allouer un financement de 17.3 millions GBP (22.3 millions USD) aux universités britanniques pour le développement de l'IA et des technologies robotiques. Le gouvernement a décidé d'augmenter l'investissement dans la R-D en IA de 4.7 milliards GBP (6.6 milliards USD) sur les quatre prochaines années, en partie par l'intermédiaire de l'ISCF (Industrial Strategy Challenge Fund).

En octobre 2017, le gouvernement a publié un examen indépendant du secteur de l'IA au Royaume-Uni. Ce rapport présente le Royaume-Uni comme un centre international d'expertise en IA, en partie grâce aux travaux de pionniers de la recherche en informatique, comme Alan Turing. Le gouvernement britannique estime que l'apport de l'IA à l'économie nationale pourrait atteindre 579.7 milliards GBP (814 milliards USD). Parmi les outils d'IA déjà utilisés au Royaume-Uni, on peut citer le guide personnel de santé (Your.MD), un agent conversationnel (*chatbot*) pour les clients des services bancaires et une plateforme éducative destinée à aider les enseignants à proposer des programmes d'enseignement personnalisés et accompagner les enfants dans leur apprentissage. Le rapport énonce 18 recommandations portant notamment sur l'amélioration de l'accès aux données et le partage des données grâce à la création de fiduciaires données, ou encore l'amélioration de l'offre de compétences en IA en introduisant des masters en intelligence artificielle qui seraient parrainés par des entreprises. Les autres mesures décrites comme prioritaires visent notamment à : optimiser la recherche en IA via la coordination de la demande de capacités informatiques pour ce type de recherche entre les institutions concernées ; soutenir l'adoption de l'IA au moyen d'un conseil national de l'IA ; et élaborer un cadre pour renforcer la transparence des décisions prises par des systèmes d'IA et la redevabilité y afférente (Hall et Pesenti, 2017^[39]).

Le gouvernement britannique a rendu publique sa stratégie industrielle en novembre 2017. Cette stratégie définit l'IA comme l'un des « quatre grands enjeux » pour faire du Royaume-Uni un acteur de premier plan des industries du futur et veiller à ce que le pays tire parti des changements majeurs en cours au niveau mondial (RU, 2017^[40]). En avril 2018, le Royaume-Uni a annoncé le *AI Sector Deal*, un programme d'investissements de 950 millions GBP (1.2 milliard USD) qui vise, en s'appuyant sur les atouts actuels du pays, à maintenir un écosystème de classe mondiale dans le domaine de l'intelligence artificielle. Ce programme comporte trois volets principaux : les talents et les compétences ; la promotion de l'adoption de l'IA ; et les données et l'infrastructure (RU, 2018^[41]).

Le gouvernement a créé un Bureau de l'intelligence artificielle (Office for Artificial Intelligence, ou OAI) chargé de mettre en œuvre le programme d'investissements décrit ci-dessus et de promouvoir plus généralement l'adoption de l'IA. Il a également mis sur pied un Centre pour l'éthique des données et l'innovation. Ce centre vise à renforcer le cadre de gouvernance pour favoriser l'innovation, tout en veillant à préserver la confiance du public. Il fournira au gouvernement une expertise et des avis indépendants sur les mesures nécessaires pour permettre des innovations majeures, respectueuses de l'éthique et sûres dans le domaine des technologies fondées sur les données et l'intelligence artificielle. Il prévoit pour ce faire de lancer une fiducie de données pilote d'ici à fin 2019. Un conseil de l'IA s'appuyant sur l'expertise des entreprises travaille en collaboration étroite avec l'OAI (RU, 2018_[42]).

Singapour

L'Autorité nationale des communications et des médias (Infocomm Media Development Authority, ou IMDA) a publié en mai 2018 un « Cadre d'action pour l'économie numérique ». Ce document trace les grandes lignes de l'action à mener pour faire de Singapour une économie numérique de premier ordre et aborde l'intelligence artificielle comme une technologie de pointe à même de tirer sa transformation numérique (Singapour, 2018_[43]).

La Commission de protection des données à caractère personnel a publié en janvier 2019 un cadre de gouvernance de l'IA pour promouvoir l'adoption responsable de l'IA à Singapour. Ce cadre-type fournit des orientations concrètes pour mettre en pratique les principes éthiques. Il s'appuie sur un document de réflexion, ainsi que sur les discussions engagées au niveau national par la Table ronde des régulateurs, une communauté de pratique regroupant des représentants des instances de réglementation sectorielles et des organismes publics. Le cadre-type, que les entités ou organisations peuvent adopter sur une base volontaire, sert également de point de départ à l'élaboration de cadres sectoriels sur la gouvernance de l'IA.

Un Conseil consultatif sur l'utilisation éthique de l'IA et des données a été créé en juin 2018. Cet organe multipartite conseille le gouvernement de Singapour sur les questions éthiques, juridiques, réglementaires et politiques soulevées par le déploiement commercial de l'IA. Tout en soutenant l'adoption de l'IA dans l'industrie, il veille au respect des principes de responsabilité et de redevabilité dans la mise à disposition des produits et des services reposant sur l'IA.

Un programme de recherche sur la gouvernance de l'intelligence artificielle et l'utilisation des données, d'une durée de cinq ans, a été lancé en septembre 2018 en vue de faire de Singapour un centre de connaissances de premier rang, doté d'une expertise internationale dans le domaine des politiques et réglementations de l'IA. Ce programme est basé dans le Centre pour la gouvernance de l'IA et des données de la Faculté de droit de l'Université de management de Singapour, qui mène des travaux de recherche sur l'intelligence artificielle dans ses relations avec l'industrie, la société et les entreprises.

Suède

Le gouvernement suédois a publié en mai 2018 un rapport intitulé « L'intelligence artificielle dans les entreprises et la société suédoises ». Ce rapport, qui a pour but de stimuler la recherche et l'innovation en IA en Suède, définit six domaines stratégiques prioritaires : i) le développement industriel, notamment dans le secteur manufacturier ; ii) le secteur des voyages et des transports ; iii) les villes intelligentes et durables ; iv) les soins de santé ; v) les services financiers ; et vi) la sécurité, en y incluant la police et les douanes. Il souligne la nécessité d'atteindre une masse critique dans la recherche, l'éducation et l'innovation. Il appelle également à une plus grande coopération, notamment en vue de l'investissement

dans la recherche et l'infrastructure, l'éducation et la formation, la réglementation et la mobilité de la main-d'œuvre (Vinnova, 2018_[44]).

Turquie

Le Conseil de la recherche scientifique et technologique, principal organisme de gestion et de financement de la recherche en Turquie, a financé de nombreux projets de R-D dans le domaine de l'intelligence artificielle. Il prévoit de lancer un appel multilatéral à projets de recherche en IA via le réseau intergouvernemental pour l'innovation EUREKA. Le ministère des Sciences et de la Technologie a établi une feuille de route nationale sur le numérique, dans le cadre de la plateforme pour la transformation numérique de l'industrie turque. Cette feuille de route porte en partie sur les progrès des technologies numériques émergentes comme l'IA.

Initiatives intergouvernementales

G7 et G20

Lors de la réunion des ministres des TIC du G7 à Takamatsu (Japon) en avril 2016, le ministre japonais des Affaires intérieures et des Communications a présenté, pour examen, une série de principes sur la R-D dans le domaine de l'intelligence artificielle (G7, 2016_[45]).

La réunion des ministres des TIC et de l'industrie du G7, qui s'est tenue à Turin en septembre 2017 sous la présidence italienne, a publié une déclaration ministérielle dans laquelle les pays du G7 ont reconnu les avantages potentiels considérables qu'offre l'IA pour la société et l'économie et sont convenus de la nécessité d'une approche centrée sur l'humain en matière d'IA (G7, 2017_[46]).

Les ministres de l'innovation du G7 réunis à Montréal en mars 2018 dans le cadre de la présidence canadienne ont exprimé une vision de l'IA centrée sur l'humain et mis l'accent sur l'interdépendance entre la croissance économique suscitée par l'innovation en IA, l'augmentation de la confiance envers l'IA et de l'adoption de l'AI, et la promotion de l'inclusivité dans le développement et le déploiement de l'IA. Les membres du G7 sont convenus d'agir dans un certain nombre de domaines pertinents, et notamment de s'attacher à :

- Investir dans la R-D fondamentale et la R-D appliquée précoce en vue de produire des innovations en IA, et soutenir l'entrepreneuriat en IA et la préparation de la population active à l'automatisation.
- Continuer d'encourager la recherche, y compris pour relever les défis sociétaux, stimuler la croissance économique et examiner les aspects éthiques de l'IA, ainsi que des questions plus vastes comme celles liées à la prise de décision automatisée.
- Appuyer les efforts voués à la sensibilisation du public pour faire connaître les avantages avérés et potentiels ainsi que les implications plus vastes de l'IA.
- Continuer de promouvoir des démarches neutres sur le plan technologique et appropriées sur les plans technique et éthique.
- Soutenir la libre circulation de l'information grâce à la mise en commun de pratiques exemplaires et de cas d'usage sur la fourniture de données gouvernementales ouvertes, interopérables et accessibles par des moyens sécurisés pour la programmation en IA.
- Diffuser la déclaration du G7 à l'échelle mondiale pour promouvoir le développement de l'IA et la collaboration sur la scène internationale (G7, 2018_[47]).

À Charlevoix en juin 2018, le G7 a publié un communiqué qui promeut une approche de l'intelligence artificielle centrée sur l'humain et l'adoption commerciale de l'IA. Les membres du G7 sont également convenus à cette occasion de continuer à promouvoir des démarches neutres sur le plan technologique et appropriées sur les plans technique et éthique.

Les ministres de l'innovation du G7 ont décidé d'organiser une conférence multipartite sur l'IA au Canada en décembre 2018. Les discussions ont porté en particulier sur les moyens de concrétiser le potentiel transformationnel positif de l'IA pour favoriser une croissance économique inclusive et durable. La France devait également proposer des initiatives concernant l'IA dans le cadre de sa présidence du G7 en 2019.

On notera par ailleurs l'attention que le G20 porte à l'IA, le Japon ayant en particulier proposé de tenir des discussions à ce sujet dans le cadre de sa présidence en 2019 (G20, 2018^[1]). Les ministres de l'économie numérique du G20, réunis à Salta (Argentine) en 2018, avaient encouragé les pays à permettre aux individus et aux entreprises de tirer avantage de la transformation numérique et des technologies émergentes comme les réseaux 5G, l'internet des objets et l'intelligence artificielle. Ils avaient invité le Japon à poursuivre, pendant sa présidence du G20 en 2019, le travail accompli en 2018 au sein du G20 dans plusieurs domaines prioritaires, parmi lesquels l'IA.

OCDE

Principes de l'OCDE sur la confiance dans l'IA et son adoption

En mai 2018, le Comité de la politique de l'économie numérique de l'OCDE a créé le Groupe d'experts sur l'intelligence artificielle à l'OCDE (AIGO) chargé d'élaborer des principes à mettre en œuvre dans le cadre des politiques publiques et de la coopération internationale et ayant vocation à promouvoir la confiance dans l'IA et son adoption. Ces principes ont nourri l'élaboration de la *Recommandation du Conseil de l'OCDE sur l'intelligence artificielle* (OCDE, 2019^[3]), à laquelle 42 pays ont adhéré le 22 mai 2019. Dans le même esprit, la présidence de la Réunion du Conseil au niveau des Ministres (RCM) de 2018 a exhorté « l'OCDE à poursuivre les discussions avec les diverses parties prenantes sur l'élaboration possible de principes devant guider le développement et l'application éthique de l'intelligence artificielle (IA) au profit des personnes ».

Le Groupe d'experts rassemble plus de 50 experts de disciplines et de secteurs différents ; les administrations, les entreprises, la communauté technique, les syndicats et la société civile, ainsi que la Commission européenne et l'UNESCO, y sont représentés. Il a tenu quatre réunions : deux au siège de l'OCDE, à Paris, les 24 et 25 septembre, et le 12 novembre 2018, une au Massachusetts Institute of Technology (MIT) à Cambridge les 16 et 17 janvier 2019, et la dernière à Dubaï, les 8 et 9 février 2019, en marge du World Government Summit. Le groupe d'experts a défini les principes d'une approche responsable à l'appui d'une IA digne de confiance, valant pour l'ensemble des parties prenantes. Parmi ces principes, citons notamment le respect des droits de l'homme, l'équité, la transparence et l'explicabilité, la robustesse, la sûreté et la sécurité, ou encore la responsabilité. Il a également proposé des recommandations spécifiques pour la mise en œuvre de ces principes dans le cadre des politiques nationales. Ces travaux ont étayé l'élaboration de la *Recommandation du Conseil de l'OCDE sur l'intelligence artificielle*, au premier semestre de 2019 (OCDE, 2019^[3]).

Observatoire des politiques relatives à l'IA

L'OCDE prévoit la création en 2019 d'un Observatoire des politiques relatives à l'IA chargé d'examiner les évolutions actuelles et à venir de l'IA et leurs implications en termes de

politiques publiques. Le but est de soutenir la mise en œuvre des principes susmentionnés en travaillant avec un large éventail d'acteurs extérieurs (administrations, entreprises, universitaires, experts techniques et grand public). L'Observatoire a vocation à devenir un centre pluridisciplinaire et multipartite ayant pour mission d'aider à la collecte de données probantes, d'éclairer les débats et de guider les pouvoirs publics aux fins de l'élaboration des politiques. Il offrira en outre aux partenaires extérieurs un point d'accès unique aux activités d'IA et aux conclusions des travaux connexes intéressant l'action publique dans l'ensemble des pays de l'OCDE.

Commission européenne et autres institutions européennes

La Commission européenne a publié en avril 2018 une communication sur « L'intelligence artificielle pour l'Europe », qui énonce trois grandes priorités : premièrement, renforcer la capacité technologique et industrielle de l'Union européenne et intensifier le recours à l'IA dans tous les secteurs de l'économie ; deuxièmement, se préparer aux changements socio-économiques induits par l'IA ; troisièmement, garantir l'existence d'un cadre éthique et juridique approprié. La Commission a également présenté en décembre 2018 un plan coordonné sur le développement de l'IA en Europe. Ce plan a principalement pour but de maximiser l'impact des investissements et de définir collectivement la marche à suivre dans le domaine de l'intelligence artificielle. Il sera appliqué jusqu'en 2027 et contient environ 70 mesures individuelles dans les domaines suivants :

- **Actions stratégiques et coordination** : Inciter les États membres à mettre en place des stratégies nationales d'IA indiquant les niveaux d'investissement et les mesures de mise en œuvre prévus.
- **Maximisation des investissements par le biais de partenariats** : Promouvoir l'investissement dans la recherche et l'innovation stratégiques en IA au moyen de partenariats public-privé, d'un groupe de leaders et d'un fonds spécifique pour soutenir les jeunes pousses et PME innovantes du secteur de l'IA.
- **Du laboratoire au marché** : Renforcer les centres d'excellence dans la recherche en IA et les pôles d'innovation numériques, et mettre en place des installations d'essai et, éventuellement des « bacs à sable réglementaires ».
- **Compétences et apprentissage tout au long de la vie** : Favoriser le talent, les compétences et l'apprentissage tout au long de la vie.
- **Données** : Appeler à la création d'un Espace européen des données pour faciliter l'accès aux données d'intérêt public, et de plateformes de données industrielles pour l'IA, y compris pour les données de santé.
- **Intégration de l'éthique dès la conception et cadre réglementaire** : Souligner la nécessité d'une éthique de l'IA et d'un cadre réglementaire adapté à cette fin (couvrant notamment les questions de sécurité et de responsabilité). Ce cadre éthique s'appuiera sur les lignes directrices en matière d'éthique pour le développement et l'utilisation de l'intelligence artificielle élaborées par un groupe d'experts indépendant (AI HLEG). La Commission européenne s'engage aussi à exiger dans ses politiques d'approvisionnement la prise en compte des considérations éthiques dès le stade de la conception des systèmes d'IA (*ethics by design*).
- **IA dans le secteur public** : Définir des mesures régissant l'adoption de l'IA dans le secteur public, notamment en matière d'achats conjoints et de traduction.

- **Coopération internationale** : Faire valoir l'importance de l'action au niveau international, intégrer l'IA dans les politiques de développement et organiser une réunion mondiale au niveau ministériel en 2019.

Dans le cadre de sa stratégie pour l'IA, la Commission a également créé en juin 2018 le Groupe d'experts indépendants de haut niveau sur l'intelligence artificielle (AI HLEG). Ce groupe d'experts, qui rassemble des représentants des milieux universitaires, de la société civile et des entreprises, s'est vu confier deux missions : premièrement, élaborer à l'intention des développeurs, responsables de déploiement et utilisateurs des systèmes d'IA un projet de « Lignes directrices en matière d'éthique pour une IA digne de confiance » ; deuxièmement, préparer pour la Commission européenne et les États membres des recommandations, destinées à étayer les politiques d'IA et les investissements connexes, sur les évolutions à moyen et long termes de l'IA à même de stimuler la compétitivité mondiale de l'Europe. La Commission a créé parallèlement un forum multipartite, l'Alliance européenne pour l'IA, pour permettre de larges discussions sur les politiques d'IA en Europe. Toute personne peut contribuer, via une plateforme en ligne, aux travaux du groupe d'experts et participer ainsi à l'élaboration des politiques de l'UE.

Le Groupe d'experts indépendants de haut niveau sur l'intelligence artificielle a publié un premier projet de lignes directrices en matière d'éthique pour commentaires en décembre 2018. Ce document établit un cadre, fondé sur la Charte des droits fondamentaux de l'UE, en vue de tendre vers une IA digne de confiance, c'est-à-dire légale, éthique et robuste d'un point de vue sociotechnique. Il énumère une série de principes éthiques applicables à l'IA. Il énonce également un certain nombre de critères clés pour garantir la fiabilité de l'IA, ainsi que les méthodes à utiliser pour les mettre en œuvre. Enfin, ce projet de lignes directrices contient une liste d'évaluation non exhaustive qui, sous forme de questions pratiques au regard de chaque critère, aidera les acteurs concernés à contrôler la bonne application des principes éthiques. À la date de rédaction de ce chapitre, le groupe d'experts travaillait encore à la révision des lignes directrices, sur la base des commentaires recueillis, afin de pouvoir les soumettre officiellement à la Commission européenne le 9 avril 2019. Cette dernière devait alors préciser les étapes à suivre en vue de l'adoption des lignes directrices et d'un cadre éthique mondial pour l'IA. Les recommandations, deuxième volet de travail du groupe d'experts, doivent être prêtes à l'été 2019.

En 2017, l'Assemblée parlementaire du Conseil de l'Europe a publié une Recommandation sur « La convergence technologique, l'intelligence artificielle et les droits de l'homme ». Cette recommandation invite instamment le Comité des Ministres à charger les organes compétents du Conseil de l'Europe d'examiner la manière dont les technologies émergentes comme l'IA remettent en question les droits de l'homme. Elle propose également que des lignes directrices soient élaborées sur des questions comme la transparence, la responsabilité et le profilage. En février 2019, le Comité des Ministres du Conseil de l'Europe a adopté une Déclaration sur les capacités de manipulation des processus algorithmiques. Cette déclaration reconnaît « les dangers qui menacent les sociétés démocratiques » liés à la capacité des « outils d'apprentissage automatique (...) d'influencer les émotions et les pensées » et encourage les États membres à combattre ces dangers. En février 2019, le Conseil de l'Europe a organisé une conférence de haut niveau sur le thème « Maîtriser les règles du jeu – L'impact du développement de l'intelligence artificielle sur les droits de l'homme, la démocratie et l'état de droit ».

En outre, la Commission européenne pour l'efficacité de la justice du Conseil de l'Europe a adopté en décembre 2018 la première Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement. Cette Charte définit cinq

principes devant guider le développement d'outils d'IA dans les systèmes judiciaires européens. En 2019, la Commission des questions juridiques et des droits de l'homme de l'Assemblée parlementaire du Conseil de l'Europe a décidé de créer une sous-commission sur l'intelligence artificielle et les droits de l'homme.

En mai 2017, le Comité économique et social européen (CESE) a adopté un avis sur l'impact sociétal de l'IA. Cet avis appelle les parties prenantes dans l'UE à veiller à ce que le développement, le déploiement et l'utilisation de l'IA soient bénéfiques à la société et contribuent au bien-être social. Le CESE insiste sur la nécessité que les humains conservent le contrôle du moment et des modalités d'utilisation de l'IA dans la vie quotidienne ; il pointe les domaines dans lesquels l'IA soulève des enjeux de société, tels que l'éthique ; la sécurité ; la transparence ; la vie privée ; les normes ; le travail ; l'éducation ; l'égalité d'accès ; la législation et la réglementation ; la gouvernance et la démocratie ; la guerre ; et la superintelligence. Le CESE préconise également l'instauration d'un code de déontologie de l'IA au niveau paneuropéen et l'adaptation des stratégies en matière d'emploi. Enfin, il plaide pour une infrastructure d'IA européenne, composée d'environnements d'apprentissage à code source libre (Muller, 2017^[48]). Le CESE a créé un groupe de travail temporaire sur l'IA pour examiner plus en détail ces questions.

Région nordique-balte

En mai 2018, les ministres des pays nordiques et baltes ont signé une déclaration commune sur « L'intelligence artificielle dans la région nordique-balte ». Les pays de la région comprennent le Danemark, l'Estonie, la Finlande, les Îles Féroé, l'Islande, la Lettonie, la Lituanie, la Norvège, la Suède et les Îles d'Åland. Ils ont décidé de renforcer leur coopération en matière d'IA, tout en maintenant leur position de première région d'Europe en termes de progression du numérique (Nordic, 2018^[49]). La déclaration mentionne sept axes d'intervention essentiels pour le développement et la promotion de l'utilisation de l'IA au service des individus : i) accroître les possibilités de développement des compétences, afin de permettre à un plus grand nombre d'administrations publiques, d'entreprises et d'organisations d'utiliser l'IA ; ii) améliorer l'accès aux données exploitables par l'IA, afin de fournir des services de meilleure qualité aux citoyens et aux entreprises de la région ; iii) élaborer des lignes directrices, normes, valeurs et principes éthiques clairs indiquant quand et comment utiliser les applications d'IA ; iv) veiller à ce que l'infrastructure, le matériel, les logiciels et les données, qui jouent un rôle clé dans l'utilisation de l'IA, reposent sur des normes garantissant l'interopérabilité, le respect de la vie privée, la sécurité, la confiance, la maniabilité et la portabilité ; v) faire en sorte que l'IA figure en bonne place dans les débats menés au niveau européen et les initiatives mises en œuvre dans le cadre du marché unique numérique ; vi) éviter toute réglementation inutile dans ce domaine, qui se développe rapidement ; vii) faciliter la collaboration dans les domaines pertinents de l'action publique via le Conseil nordique des ministres.

Nations Unies

En septembre 2017, l'Institut interrégional de recherche des Nations Unies sur la criminalité et la justice a signé un accord avec les Pays-Bas, le pays hôte, en vue de l'ouverture, dans le cadre du système de l'ONU, d'un Centre pour l'intelligence artificielle et la robotique à La Haye².

L'Union internationale des télécommunications (UIT) a travaillé de concert avec plus de 25 autres organes de l'ONU pour organiser le Sommet mondial « *AI for Good* ». Elle a également noué des partenariats avec des organisations comme la Fondation XPRIZE et

l'ACM (Association for Computing Machinery). Suite au premier sommet de juin 2017, l'UIT a organisé un deuxième sommet, qui s'est tenu à Genève en mai 2018³.

L'UNESCO est à l'initiative d'un dialogue mondial sur l'éthique de l'intelligence artificielle, afin de mener une réflexion sur sa complexité et ses incidences sur la société et l'humanité en général. Elle a organisé en septembre 2018 une table ronde publique avec des experts puis, en mars 2019, une conférence mondiale sur le thème « Principes pour l'IA : vers une approche humaniste ? ». Ces deux événements visaient à sensibiliser le public et encourager la réflexion sur les opportunités et les défis associés à l'IA et aux technologies apparentées. En novembre 2019, la 40^e Conférence générale de l'UNESCO pourrait examiner la possibilité d'élaborer une recommandation sur l'IA en 2020-21, sous réserve de l'approbation du Conseil exécutif de l'UNESCO en avril 2019.

Organisation internationale de normalisation

L'Organisation internationale de normalisation (ISO) et la Commission électrotechnique internationale (IEC) ont créé en 1987 un comité technique commun ISO/IEC JTC 1 chargé de définir des normes pour les applications professionnelles et grand public des technologies de l'information. En octobre 2017, le sous-comité 42 (SC 42) a été créé sous l'égide du comité technique JTC 1 afin d'élaborer des normes en matière d'IA. Ce sous-comité produit des directives à l'intention des comités de l'ISO et de l'IEC qui travaillent sur les applications de l'IA. Il se charge notamment d'établir un cadre et une terminologie communs, d'identifier les approches computationnelles et architectures des systèmes d'IA, et d'évaluer les risques et les menaces qui leur sont associés (Price, 2018_[50]).

Initiatives d'acteurs privés

Un grand nombre de partenariats et d'initiatives ont été lancés par des acteurs non gouvernementaux pour réfléchir aux questions soulevées par l'intelligence artificielle. Bien que beaucoup de ces initiatives présentent un caractère multipartite, la présente section mentionne avant tout celles qui émanent de la communauté technique, du secteur privé, des syndicats ou des milieux universitaires. La liste ci-dessous n'est donc pas exhaustive.

Communauté technique et milieux universitaires

L'IEEE a lancé en avril 2016 une Initiative mondiale sur l'éthique dans la conception des systèmes autonomes et intelligents. Cette initiative a pour but de faire progresser le débat public sur la mise en œuvre des technologies d'IA et de définir des normes éthiques et valeurs prioritaires à cet égard. L'IEEE a également publié en décembre 2017 la deuxième version de ses Principes d'intégration de l'éthique dès la conception (*Ethically Aligned Design*), en sollicitant les commentaires du public. La version finale devrait être publiée en 2019 (Tableau 5.3) (IEEE, 2017_[51]). Par ailleurs, l'IEEE a annoncé en juin 2018 la création avec le Media Lab du MIT d'un « conseil pour une approche élargie de l'intelligence » (*Council for Extended Intelligence*). Cet organe cherche à promouvoir le développement responsable de systèmes intelligents, la reprise du contrôle sur les données à caractère personnel et l'élaboration d'indicateurs de la prospérité économique autres que le produit intérieur brut (Pretz, 2018_[52]).

Les Principes d'Asilomar, qui comprennent 23 principes pour un développement sûr et socialement bénéfique de l'IA à court et plus long termes, sont issus d'une conférence organisée par le Future of Life Institute en janvier 2017. Ils ont été élaborés à partir de débats, de réflexions et de documents émanant de l'IEEE, du monde universitaire et d'organisations à but non lucratif.

Tableau 5.3. Principes généraux de l'IEEE pour une intégration de l'éthique dès la conception (*Ethically Aligned Design*, deuxième version)

Principes	Objectifs
Droits de l'homme	Veiller à ce que les systèmes autonomes et intelligents (SAI) ne portent pas atteinte aux droits fondamentaux reconnus au niveau international
Priorité au bien-être	Donner la priorité aux indicateurs de bien-être dans la conception et l'utilisation des SAI, les indicateurs traditionnels de la prospérité ne prenant pas pleinement en compte les effets des technologies liées aux systèmes d'intelligence artificielle sur le bien-être des individus
Responsabilité/redevabilité	Veiller à ce que les concepteurs et opérateurs soient responsables et redevables du fonctionnement des SAI
Transparence	Assurer la transparence du fonctionnement des SAI
Conscience des risques d'utilisation abusive	Réduire au minimum les risques d'utilisation abusive des SAI

Source : IEEE (2017^[51]), *Ethically Aligned Design (Version 2)*, http://standards.ieee.org/develop/indconn/ec/ead_v2.pdf.

Ces principes sont classés en trois groupes. Le premier, qui porte sur la recherche, appelle à : financer la recherche sur une utilisation bénéfique de l'IA, qui prenne en compte des questions épineuses en matière d'informatique, d'économie, de droit, d'éthique et de sciences sociales ; développer des relations constructives entre scientifiques et décideurs politiques ; et insuffler une culture de la recherche technique fondée sur la coopération, la confiance et la transparence. Le second, qui concerne l'éthique et les valeurs, appelle à prendre en compte un certain nombre d'exigences dans la conception et le fonctionnement des systèmes d'IA : la sûreté et la sécurité ; la transparence et la responsabilité ; le respect de la liberté individuelle et de la vie privée, de la dignité humaine, des droits fondamentaux et de la diversité culturelle ; et l'autonomisation des individus et le partage des avantages. Le troisième recommande notamment, à plus long terme, d'éviter les hypothèses fortes au sujet des capacités maximum des futurs systèmes d'IA et de planifier soigneusement le développement éventuel de l'intelligence artificielle générale (IAG) (FLI, 2017^[53]). Le Tableau 5.4 présente la liste des Principes d'Asilomar les plus importants.

OpenAI, une organisation de recherche en IA à but non lucratif, a été créée fin 2015. Elle emploie 60 chercheurs à plein temps dont la mission est de « développer des systèmes d'IAG sûrs, en veillant à que les avantages qui en résultent soient répartis de la façon la plus large et la plus équitable possible »⁴.

The Future Society a lancé en 2015 une initiative sur l'intelligence artificielle (*AI Initiative*) afin de contribuer à l'élaboration d'un cadre mondial pour les politiques d'IA. L'organisation héberge une plateforme en ligne permettant un débat civique et des échanges pluridisciplinaires. Cette plateforme a pour but d'aider à comprendre la dynamique des technologies d'IA, ainsi que les avantages et les risques qui leur sont associés, en vue d'étayer la formulation de recommandations d'action⁵.

De multiples initiatives universitaires sont également en cours dans tous les pays de l'OCDE et de nombreuses économies partenaires. L'initiative de recherche sur les politiques de l'internet (*Internet Policy Research Initiative*) du MIT, par exemple, vise à jeter un pont entre la communauté technique et celle des décideurs. En 2017, le Centre Berkman Klein de l'Université d'Harvard a lancé une initiative sur l'éthique et la gouvernance de l'intelligence artificielle (*Ethics and Governance of Artificial Intelligence Initiative*). De son côté, le Media Lab du MIT travaille plus spécialement sur des questions comme : les algorithmes et l'équité, les véhicules autonomes et la transparence, et l'explicabilité des systèmes d'IA.

**Tableau 5.4. Principes d'Asilomar sur l'intelligence artificielle
(intitulés des principes essentiels)**

	Recherche	Éthique et valeurs	Perspectives à plus long terme
Titres des principes	<ul style="list-style-type: none"> - Finalité de la recherche - Financement de la recherche - Relations entre scientifiques et décideurs - Culture de la recherche - Refus de la course à tout prix 	<ul style="list-style-type: none"> - Sécurité - Transparence en cas de défaillance - Transparence judiciaire - Responsabilité - Concordance des valeurs - Valeurs humaines - Données personnelles - Liberté et respect de la vie privée - Bénéfice collectif - Partage de la prospérité - Contrôle humain 	<ul style="list-style-type: none"> - Prudence au sujet des capacités futures - Ampleur des transformations à venir - Risques - Récursivité et répliquabilité des systèmes d'IA - Bien commun

Source : FLI (2017^[53]), *Asilomar IA Principles*, <https://futureoflife.org/ai-principles/>.

Initiatives du secteur privé

En septembre 2016, Amazon, DeepMindGoogle, Facebook, IBM et Microsoft ont lancé le Partenariat pour l'intelligence artificielle au service des citoyens et de la société (*Partnership on Artificial Intelligence to Benefit People and Society*, ou PAI). Cette initiative a pour but d'étudier et de promouvoir les bonnes pratiques relatives aux technologies d'IA, d'aider le public à mieux comprendre l'IA et d'offrir une plateforme ouverte de discussion et d'engagement autour de l'IA et de ses implications pour les individus et la société. Depuis sa création, ce partenariat est devenu une communauté pluridisciplinaire de parties prenantes et compte aujourd'hui plus de 80 membres comprenant aussi bien des entreprises de technologie à but lucratif que des représentants de la société civile, des universités et instituts de recherche, ou des startups.

Le Conseil de l'industrie des technologies de l'information (*Information Technology Industry Council*, ou ITI) est une association d'entreprises de technologie basée à Washington, qui compte plus d'une soixantaine de membres. L'ITI a publié en octobre 2017 des principes applicables aux politiques en matière d'IA (*AI Policy Principles*) (Tableau 5.5). Ce document précise les responsabilités du secteur dans certains domaines, invite les gouvernements à soutenir la recherche en IA et appelle au développement des partenariats public-privé (ITI, 2017^[54]). Les entreprises ont aussi pris individuellement certaines initiatives.

Tableau 5.5. Principes de l'ITI applicables aux politiques en matière d'intelligence artificielle

Responsabilité des entreprises : promouvoir le développement et l'utilisation responsables de l'IA	Opportunité pour les gouvernements : investir dans l'écosystème de l'IA et faciliter sa croissance	Opportunité pour les partenariats public-privé : promouvoir l'éducation tout au long de la vie et la diversité
<ul style="list-style-type: none"> - Conception et déploiement responsables - Sûreté et contrôlabilité - Robustesse et représentativité des données - Interprétabilité - Responsabilité des systèmes d'IA autonomes 	<ul style="list-style-type: none"> - Investissement dans la recherche et le développement en matière d'IA - Démarche réglementaire flexible - Promotion de l'innovation et de la sécurité de l'internet - Cybersécurité et protection de la vie privée - Élaboration de normes mondiales et promotion des meilleures pratiques 	<ul style="list-style-type: none"> - Démocratisation de l'accès et égalité des chances - Enseignement des sciences, des technologies, de l'ingénierie et des mathématiques - Préparation et adaptation de la main-d'œuvre - Partenariats public-privé

Source : ITI (2017^[54]), *AI Policy Principles*, <https://www.itic.org/resources/AI-Policy-Principles-FullReport2.pdf>.

Société civile

The Public Voice, un regroupement d'associations créé par l'EPIC (Electronic Privacy Information Center), a publié en octobre 2018 des lignes directrices universelles sur l'intelligence artificielle (*Universal Guidelines on Artificial Intelligence*, UGAI) (The Public Voice, 2018^[55]). Ces lignes directrices attirent l'attention sur les enjeux croissants associés aux systèmes informatiques intelligents et formulent des recommandations concrètes pour améliorer et étayer leur conception. Elles s'attachent essentiellement à promouvoir la transparence et la responsabilité des systèmes d'IA, et à garantir que les humains conservent le contrôle des systèmes qu'ils créent⁶. Les douze principes énumérés dans ces lignes directrices énoncent toute une série de droits et d'obligations, en particulier : le droit à la transparence ; le droit à la détermination humaine ; l'obligation d'identification ; l'obligation d'équité ; l'obligation d'évaluation préalable et de responsabilité ; l'obligation de précision, de fiabilité et de validité ; l'obligation de qualité des données ; l'obligation de sécurité publique ; l'obligation de cybersécurité ; et l'obligation de supprimer un système d'IA dans l'éventualité où ce système ne pourrait plus être contrôlé par un humain. À cela s'ajoutent l'interdiction du profilage secret et l'interdiction de tout système de notation centralisé reposant sur l'intelligence artificielle.

Syndicats

UNI Global Union est un syndicat qui représente plus de 20 millions de travailleurs qualifiés et de travailleurs du secteur des services répartis dans plus de 150 pays. Préparer un avenir garantissant l'autonomie des travailleurs et l'accès à un emploi décent est une priorité phare d'UNI Global Union. C'est pourquoi le syndicat a formulé « **Dix principes majeurs pour une intelligence artificielle éthique** » (*Top 10 Principles for Ethical AI*). **Le but est d'assurer le respect des droits des travailleurs dans les conventions collectives et les accords-cadres mondiaux, ainsi qu'au niveau des organisations plurinationales de syndicats et de délégués syndicaux et des organisations mondiales** (Tableau 5.1) (Colclough, 2018^[56]).

Tableau 5.6. Dix principes majeurs pour une intelligence artificielle éthique (UNI Global Union)

1. Réclamer que les systèmes d'IA soient transparents	Les travailleurs doivent avoir le droit de réclamer la transparence des décisions/résultats des systèmes d'IA et des algorithmes sous-jacents. Ils doivent aussi être consultés sur la mise en œuvre, le développement et le déploiement des systèmes d'IA.
2. Doter les systèmes d'IA d'une « boîte noire éthique »	La « boîte noire éthique » doit contenir non seulement les données pertinentes pour garantir la transparence et la redevabilité d'un système, mais aussi des données et des informations claires sur les considérations éthiques intégrées à ce système.
3. Faire en sorte que l'IA serve les individus et la planète	La mise en place de codes d'éthique pour le développement, l'application et l'utilisation de l'IA est nécessaire pour que tout au long de leur processus opérationnel, les systèmes d'IA restent compatibles avec, et renforcent, les principes de dignité humaine, d'intégrité, de liberté, de respect de la sphère privée et de diversité culturelle et de genre, ainsi que les droits fondamentaux.
4. Adopter une approche donnant les commandes à l'être humain	Un préalable absolu est que le développement de l'IA soit responsable, sûr et utile, que les machines conservent le statut juridique d'outils et que des personnes morales conservent le contrôle sur ces machines et la responsabilité envers elles en tout temps.
5. Veiller à une IA sans distinction de genre ni préjugés	Dans la conception et la maintenance de l'IA, il est vital que le système subisse des contrôles afin d'éliminer les préjugés négatifs ou préjudiciables envers l'être humain, et de veiller à ce que tout préjugé (genre, origine, orientation sexuelle, âge) soit identifié et ne soit pas propagé par le système.
6. Faire partager les avantages des systèmes d'IA	La prospérité économique créée par l'IA devrait être répartie largement et de manière égale, au profit de l'ensemble de l'humanité. Des politiques mondiales autant que nationales visant à surmonter le fossé numérique économique, technologique et social sont donc nécessaires.
7. Assurer une transition juste et garantir le soutien des	Au fur et à mesure du développement des systèmes d'IA et de la création de réalités augmentées, les travailleurs perdront leur emploi et les tâches professionnelles disparaîtront. Il est vital que des politiques soient mises en place pour garantir une transition juste menant à la

droits et libertés fondamentales	réalité numérique, y compris des mesures gouvernementales spécifiques pour aider les travailleurs qui ont perdu leur emploi à en trouver un nouveau.
8. Établir un mécanisme de gouvernance mondial	Il est nécessaire de créer des instances de gouvernance sur le travail décent et l'IA éthique, regroupant toutes les parties prenantes au niveau mondial comme au niveau régional. Ces instances devraient inclure les concepteurs d'IA, les fabricants, les propriétaires, les développeurs, les chercheurs, les employeurs, les juristes, les organisations de la société civile et les syndicats.
9. Interdire l'attribution de responsabilités aux robots	Les robots devraient être conçus et exploités dans la mesure du possible de manière à respecter la législation existante et les droits et libertés fondamentales, y compris le respect de la sphère privée.
10. Interdire la course aux armements d'IA	Les armes létales autonomes, y compris la cyberguerre, devraient être interdites. UNI Global Union appelle à l'adoption d'une convention mondiale sur l'IA éthique pour prévenir et gérer les conséquences négatives indésirables de l'IA, tout en optimisant les avantages de l'IA pour les travailleurs et la société. UNI Global Union souligne que les agents responsables doivent être les êtres humains et les entreprises.

Source : Colclough (2018^[56]), « Ethical artificial intelligence – 10 essential ingredients », <https://www.oecd-forum.org/channels/722-digitalisation/posts/29527-10-principles-for-ethical-artificial-intelligence>.

Références

- Allemagne (2018), *Artificial Intelligence Strategy*, Gouvernement fédéral allemand, [5]
<https://www.ki-strategie-deutschland.de/home.html>.
- Allemagne (2017), *Automated and Connected Driving*, BMVI Ethics Commission, [6]
https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile.
- Benhamou, S. et L. Janin (2018), *Intelligence artificielle et travail*, France Stratégie, [28]
<http://www.strategie.gouv.fr/publications/intelligence-artificielle-travail>.
- Bésil (2018), *Brazilian digital transformation strategy « E-digital »*, Ministério da Ciência, Tecnologia, Inovações e Comunicações, [7]
<http://www.mctic.gov.br/mctic/export/sites/institucional/sessaoPublica/arquivos/digitalstrategy.pdf>.
- Chine (2018), *AI Standardisation White Paper*, Gouvernement de Chine, traduit en anglais par Jeffrey Ding, Chercheur pour le Future of Humanity's Governance of AI Program, [14]
<https://baijia.baidu.com/s?id=1589996219403096393>.
- Chine (2017), *Guideline on Next Generation AI Development Plan*, Gouvernement de Chine, Conseil d'État, [12]
http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm.
- Chine (2016), *Three-Year Action Plan for Promoting Development of a New Generation Artificial Intelligence Industry (2018-2020)*, Ministère chinois de l'Industrie et des Technologies de l'information, [11]
<http://www.miit.gov.cn/n1146290/n1146392/c4808445/content.html>.
- CIFAR (2017), *Stratégie pancanadienne en matière d'intelligence artificielle*, CIFAR, [8]
<https://www.cifar.ca/fr/ia/strategie-pancanadienne-en-matiere-dintelligence-artificielle>.
- Colclough, C. (2018), « Ethical Artificial Intelligence – 10 Essential Ingredients », *A.Ideas Series*, n° 24, Réseau du Forum, Éditions OCDE, Paris, [56]
<https://www.oecd-forum.org/channels/722-digitalisation/posts/29527-10-principles-for-ethical-artificial-intelligence>.
- Corée (2016), *Mid- to Long-term Master Plan in Preparation for the Intelligent Information Society*, Gouvernement de Corée, Exercice interministériel, [16]
http://english.msip.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2017/07/20/Master%20Plan%20for%20the%20intelligent%20information%20society.pdf.
- Corée (2016), « MSIP announces development strategy for the intelligence information industry », *Science, Technology & ICT Newsletter, Ministry of Science and ICT, Government of Korea*, 16, [15]
<https://english.msit.go.kr/english/msipContents/contentsView.do?cateId=msse44&artId=1296203>.

- Danemark (2018), *Strategy for Denmark's Digital Growth*, Gouvernement du Danemark, [18]
<https://em.dk/english/publications/2018/strategy-for-denmarks-digital-growth>.
- EOP (2018), *Artificial Intelligence for the American People*, Executive Office of the President, [21]
 Gouvernement des États-Unis, <https://www.whitehouse.gov/briefings-statements/artificial-intelligence-american-people/>.
- Finlande (2017), *Finland's Age of Artificial Intelligence - Turning Finland into a Leader in the Application of AI*, page web, Ministère finlandais de l'Emploi et de l'Économie, [24]
<https://tem.fi/en/artificial-intelligence-programme>.
- FLI (2017), *Asilomar AI Principles*, Future of Life Institute (FLI), <https://futureoflife.org/ai-principles/>. [53]
- Fonds de recherche du Québec (2018), « Québec lays the groundwork for a world observatory on the social impacts of artificial intelligence and digital technologies », *communiqué de presse*, 29 mars, <https://www.newswire.ca/news-releases/quebec-lays-the-groundwork-for-a-world-observatory-on-the-social-impacts-of-artificial-intelligence-and-digital-technologies-678316673.html>. [9]
- G20 (2018), *Ministerial Declaration – G20 Digital Economy*, Réunion ministérielle du G20 sur l'économie numérique, 24 août 2018, Salta, Argentine, [1]
https://g20.argentina.gob.ar/sites/default/files/digital_economy_-_ministerial_declaration.pdf.
- G7 (2018), *Résumé des présidents : Réunion ministérielle sur le thème « Se préparer aux emplois de l'avenir »*, https://international.gc.ca/world-monde/international_relations-relations_internationales/g7/documents/2018-03-27-chairs_summary-resume_presidents.aspx?lang=fra. [47]
- G7 (2017), *Artificial Intelligence (Annex 2)*, [46]
http://www.g7italy.it/sites/default/files/documents/ANNEX2-Artificial_Intelligence_0.pdf.
- G7 (2016), *Proposal of Discussion toward Formulation of AI R&D Guideline*, Ministère japonais des Affaires intérieures et des Communications, [45]
http://www.soumu.go.jp/joho_kokusai/g7ict/english/index.html.
- Hall, W. et J. Pesenti (2017), *Growing the Artificial Intelligence Industry in the UK*, Wendy Hall and Jérôme Pesenti, [39]
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf.
- IEEE (2017), *Ethically Aligned Design (Version 2)*, IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, [51]
http://standards.ieee.org/develop/indconn/ec/ead_v2.pdf.
- Inde (2018), « National Strategy for Artificial Intelligence #AI for All, Discussion Paper », [29]
Discussion Paper, NITI Aayog,
http://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf.

- Italie (2018), *Artificial Intelligence at the Service of the Citizen*, Agency for Digital Italy, [30]
<https://libro-bianco-ia.readthedocs.io/en/latest/>.
- ITI (2017), *AI Policy Principles*, Information Technology Industry Council, [54]
<https://www.itic.org/resources/AI-Policy-Principles-FullReport2.pdf>.
- Japon (2019), *Social Principles for Human-Centric AI*, Japan Cabinet Office, April, [33]
<https://www8.cao.go.jp/cstp/stmain/aisocialprinciples.pdf>.
- Japon (2018), *Draft AI Utilization Principles*, Ministry of Internal Affairs and Communications, [35]
Japon, http://www.soumu.go.jp/main_content/000581310.pdf.
- Japon (2018), *Integrated Innovation Strategy*, Japan Cabinet Office, June, [32]
https://www8.cao.go.jp/cstp/english/doc/integrated_main.pdf.
- Japon (2017), *Artificial Intelligence Technology Strategy*, Strategic Council for AI Technology, [31]
<http://www.nedo.go.jp/content/100865202.pdf>.
- Japon (2017), *Draft AI R&D Guidelines for International Discussions*, Ministry of Internal [34]
Affairs and Communications, Japan, http://www.soumu.go.jp/main_content/000507517.pdf.
- Jing, M. et S. Dai (2018), *Here's what China is doing to boost its artificial intelligence [10]
capabilities*, 10 May, <https://www.scmp.com/tech/science-research/article/2145568/can-trumps-ai-summit-match-chinas-ambitious-strategic-plan>.
- Kaevats, M. (2017), *Estonia's Ideas on Legalising AI*, exposé présenté à la conférence AI: [19]
Intelligent Machines, Smart Policies, Paris, les 26 et 27 octobre 2017,
<https://prezi.com/yabrlekhmcj4/oecd-6-7min-paris/>.
- Kaevats, M. (25 septembre 2017), *Estonia considers a 'kratt law' to legalise artificial [20]
intelligence (AI)*, E-residency blog, <https://medium.com/e-residency-blog/estonia-starts-public-discussion-legalising-ai-166cb8e34596>.
- Kania, E. (2018), « China's AI agenda advances », *The Diplomat* 14 février, [13]
<https://thediplomat.com/2018/02/chinas-ai-agenda-advances/>.
- Mexico (2018), *Towards an AI strategy in Mexico: Harnessing the AI Revolution*, British [36]
Embassy Mexico City, Oxford Insights, C minds, <http://go.wizeline.com/rs/571-SRN-279/images/Towards-an-AI-strategy-in-Mexico.pdf>.
- Muller, C. (2017), *Avis du CESE: L'intelligence artificielle*, Comité économique et social [48]
européen, Bruxelles, <https://www.eesc.europa.eu/fr/our-work/opinions-information-reports/opinions/lintelligence-artificielle>.
- Nordic (2018), *AI in the Nordic-Baltic Region*, Nordic Council of Ministers, [49]
https://www.regeringen.se/49a602/globalassets/regeringen/dokument/naringsdepartementet/20180514_nmr_deklaration-slutlig-webb.pdf.
- OCDE (2019), *Recommandation du Conseil sur l'intelligence artificielle*, OCDE, Paris, [3]
<https://legalinstruments.oecd.org/api/print?ids=648&lang=fr>.

- OCDE (2019), *Scoping Principles to Foster Trust in and Adoption of AI – Proposal by the Expert Group on Artificial Intelligence at the OECD (AIGO)*, Éditions OCDE, Paris, <http://oe.cd/ai>. [2]
- OGCR (2018), *Analysis of the Development Potential of Artificial Intelligence in the Czech Republic*, Office of the Government of the Czech Republic, <https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-Summary-Report.pdf>. [37]
- Peng, T. (2018), « South Korea aims high on AI, pumps \$2 billion into R&D », *Medium*, 16 May, <https://medium.com/syncedreview/south-korea-aims-high-on-ai-pumps-2-billion-into-r-d-de8e5c0c8ac5>. [17]
- Porter, M. (1990), « The competitive advantage of nations », *Harvard Business Review*, March-April, <https://hbr.org/1990/03/the-competitive-advantage-of-nations>. [4]
- Pretz, K. (2018), « IEEE Standards Association and MIT Media Lab form council on extended intelligence », *IEEE Spectrum*, <http://theinstitute.ieee.org/resources/ieee-news/ieee-standards-association-and-mit-media-lab-form-council-on-extended-intelligence>. [52]
- Price, A. (2018), « First international standards committee for entire AI ecosystem », *IE e-tech*, 03, <https://ieccetech.org/Technical-Committees/2018-03/First-International-Standards-committee-for-entire-AI-ecosystem>. [50]
- RU (2018), *AI Sector Deal*, Department for Business, Energy & Industrial Strategy and Department for Digital, Culture, Media & Sport, Gouvernement du Royaume-Uni, <https://www.gov.uk/government/publications/artificial-intelligence-sector-deal>. [41]
- RU (2018), *Centre for Data Ethics and Innovation Consultation*, Department for Digital, Culture, Media & Sport, Gouvernement du Royaume-Uni, <https://www.gov.uk/government/consultations/consultation-on-the-centre-for-data-ethics-and-innovation/centre-for-data-ethics-and-innovation-consultation>. [42]
- RU (2017), *Industrial Strategy: Building a Britain Fit for the Future*, Gouvernement du Royaume-Uni, <https://www.gov.uk/government/publications/industrial-strategy-building-a-britain-fit-for-the-future>. [40]
- RU (2017), *UK Digital Strategy*, Gouvernement du Royaume-Uni, <https://www.gov.uk/government/publications/uk-digital-strategy>. [38]
- Russie (2017), *Digital Economy of the Russian Federation*, Gouvernement de la Fédération de Russie, <http://pravo.gov.ru>. [23]
- Singapour (2018), *Digital Economy Framework for Action*, Infocomm Media Development Authority, <https://www.imda.gov.sg/-/media/imda/files/sg-digital/sgd-framework-for-action.pdf?la=en>. [43]

- Sivonen, P. (2017), *Ambitious Development Program Enabling Rapid Growth of AI and Platform Economy in Finland*, exposé présenté à la conférence AI Intelligent Machines, Smart Policies, Paris, les 26 et 27 octobre 2017, <http://www.oecd.org/going-digital/ai-intelligent-machines-smart-policies/conference-agenda/ai-intelligent-machines-smart-policies-sivonen.pdf>. [25]
- The Public Voice (2018), *Universal Guidelines for Artificial Intelligence*, The Public Voice Coalition, October, <https://thepublicvoice.org/ai-universal-guidelines/memo/>. [55]
- Thompson, N. (2018), « Emmanuel Macron talks to WIRED about France's AI strategy », *WIRED*, 31 March, <https://www.wired.com/story/emmanuel-macron-talks-to-wired-about-frances-ai-strategy>. [27]
- US (2017), « Delaney launches bipartisan artificial intelligence (AI) caucus for 115th Congress », Congressional Artificial Intelligence Caucus, communiqué de presse, 24 mai, <https://artificialintelligencecaucus-olson.house.gov/media-center/press-releases/delaney-launches-ai-caucus>. [22]
- Villani, C. (2018), *Donner un sens à l'intelligence artificielle - Pour une stratégie nationale et européenne*, AI for Humanity, consulté en décembre 2018, <https://www.aiforhumanity.fr/>. [26]
- Vinnova (2018), *Artificial Intelligence in Swedish Business and Society*, Vinnova, 28 octobre, https://www.vinnova.se/contentassets/29cd313d690e4be3a8d861ad05a4ee48/vr_18_09.pdf. [44]

Notes

¹ Ce rapport a été commandité par l'ambassade britannique au Mexique, financé par le Prosperity Fund du Royaume-Uni et élaboré par Oxford Insights et C Minds, avec la collaboration du gouvernement mexicain et la contribution d'experts de l'ensemble du Mexique.

² Voir www.unicri.it/news/article/2017-09-07_Establishment_of_the_UNICRI.

³ Voir <https://www.itu.int/en/ITU-T/AI/>.

⁴ Voir <https://openai.com/about/#mission>.

⁵ Voir <http://ai-initiative.org/ai-consultation/>.

⁶ Pour plus de détails, voir <https://thepublicvoice.org/ai-universal-guidelines/memo/>.



Extrait de :
Artificial Intelligence in Society

Accéder à cette publication :
<https://doi.org/10.1787/eedfee77-en>

Merci de citer ce chapitre comme suit :

OCDE (2019), « Politiques et initiatives dans le domaine de l'IA », dans *Artificial Intelligence in Society*, Éditions OCDE, Paris.

DOI: <https://doi.org/10.1787/63e0ffac-fr>

Cet ouvrage est publié sous la responsabilité du Secrétaire général de l'OCDE. Les opinions et les arguments exprimés ici ne reflètent pas nécessairement les vues officielles des pays membres de l'OCDE.

Ce document et toute carte qu'il peut comprendre sont sans préjudice du statut de tout territoire, de la souveraineté s'exerçant sur ce dernier, du tracé des frontières et limites internationales, et du nom de tout territoire, ville ou région.

Vous êtes autorisés à copier, télécharger ou imprimer du contenu OCDE pour votre utilisation personnelle. Vous pouvez inclure des extraits des publications, des bases de données et produits multimédia de l'OCDE dans vos documents, présentations, blogs, sites Internet et matériel d'enseignement, sous réserve de faire mention de la source OCDE et du copyright. Les demandes pour usage public ou commercial ou de traduction devront être adressées à rights@oecd.org. Les demandes d'autorisation de photocopier une partie de ce contenu à des fins publiques ou commerciales peuvent être obtenues auprès du Copyright Clearance Center (CCC) info@copyright.com ou du Centre français d'exploitation du droit de copie (CFC) contact@cfcopies.com.