

**Rapport de 2020 du groupe de
travail chargé du développement,
de l'utilisation et de la
gouvernance responsables de
l'intelligence artificielle (IA)**

Novembre 2020 - Sommet de Montréal sur le Partenariat
mondial sur l'intelligence artificielle (PMIA)



GPAI

THE GLOBAL PARTNERSHIP
ON ARTIFICIAL INTELLIGENCE

Veillez noter que le présent rapport a été élaboré par des experts du groupe de travail chargé du développement, de l'utilisation et de la gouvernance responsables de l'IA du Partenariat mondial sur l'intelligence artificielle (PMIA). Le rapport reflète les opinions personnelles des experts du PMIA et ne reflète pas nécessairement le point de vue des organisations des experts, du PMIA, de l'OCDE ou de leurs membres respectifs.



Mot de bienvenue des coprésidents	4
Présentation de l'IAR	5
Membres de l'IAR	6
Mandat de l'IAR	7
Méthodes de travail	8
Calendrier du groupe de travail	10
Recommandations préliminaires et résultats pour le Sommet	11
Priorités de l'IAR pour les trois à six prochains mois	12
Vision à long terme	14
Annexe 1	15

Mot de bienvenue des coprésidents



Yoshua Bengio
Fondateur et directeur scientifique de Mila (Institut d'intelligence artificielle du Québec)



Raja Chatila
Directeur du Labex SMART sur les interactions homme-machine, Université de la Sorbonne

Fondé en 2020, le [PMIA](#) a pour mission d'entreprendre et de soutenir des projets d'IA appliquée et d'offrir un mécanisme d'échanges pour l'analyse, la prévoyance et la coordination multidisciplinaires dans le but de faciliter la collaboration et les synergies internationales et de réduire la duplication des travaux dans le domaine de la gouvernance des systèmes d'IA.

Nous coprésidons l'un des quatre groupes de travail d'experts, le groupe de travail chargé du développement, de l'utilisation et de la gouvernance responsables de l'IA ([ci-après l'« IAR »](#)).

L'IAR a pour mandat d'encourager et de contribuer au développement, à l'utilisation et à la gouvernance responsables des systèmes d'IA centrés sur l'humain en cohérence avec les objectifs de développement durable des Nations Unies. Les travaux de l'IAR reposent sur une vision de l'IA qui est juste et respectueuse des droits de la personne et de la démocratie. Son objectif est d'être équitable et inclusif, et de contribuer positivement à l'intérêt général.

Nous reconnaissons que les partenariats seront essentiels pour remplir notre mission et maximiser l'impact de nos travaux. Les partenariats offrent l'occasion de créer en collaboration des solutions aux défis cernés grâce aux spécialistes de domaines donnés et à l'expertise technique et en IA du PMIA. En outre, ils donnent la possibilité de trouver des solutions coordonnées et structurées qui combinent le financement, les connaissances et la mise en œuvre nécessaires à la dimension des défis que nous chercherons à relever avec les comités proposés, par exemple, ceux portant sur les changements climatiques, l'éducation et la découverte de médicaments et la science ouverte. Nous sommes impatients d'explorer plus en détail ces possibilités dans la prochaine phase de nos travaux, et nous reconnaissons également que notre groupe de travail retirera un bénéfice considérable de la collaboration sur les questions transversales avec les collègues des autres groupes de travail.

Le présent rapport fait le point des travaux effectués par l'IAR au cours des six derniers mois ainsi que de ses objectifs pour le prochain semestre et les deux ou trois prochaines années. Bien que préfacé et présenté par nous, le rapport est en réalité le fruit des efforts déployés par les 30 membres de l'IAR ainsi que les neuf membres du comité directeur de l'IAR, à l'égard desquels nous sommes extrêmement reconnaissants.

Comme on le répète souvent, à juste titre, l'IA peut favoriser les progrès de nos sociétés dans de nombreux domaines différents, mais elle pourrait également avoir des conséquences négatives aussi bien sur les personnes que sur les collectivités dans leur ensemble si elle est développée et utilisée de manière imprudente ou sournoise.

Ainsi, nous espérons que l'IAR contribuera à l'élaboration et à la mise en place de mécanismes et d'outils de coordination internationale qui nous aideront, collectivement, à utiliser les systèmes d'IA comme un moyen d'atteindre les objectifs de développement durable des Nations unies; en plus d'aider à la mise en évidence de mécanismes et à la mise au point d'outils de gouvernance qui nous permettront de minimiser les effets négatifs de l'IA.

Présentation de l'IAR

L'IAR est composé de 30 membres. Ses experts internationaux sont issus de différents domaines, ce qui favorise des débats énergiques et l'émergence de points de vue divers. Plus précisément, 14 membres de l'IAR proviennent du monde technique (par exemple, l'apprentissage machine, les technologies de l'information), tandis que 16 d'entre eux sont issus du secteur des sciences sociales et humaines et de domaines comme les communications, l'anthropologie, la littérature, la gestion, l'histoire, la psychologie, la philosophie, les affaires internationales, le développement international, le journalisme, l'économie et les sciences politiques.

40 % des membres sont des femmes, un chiffre que nous nous emploierons à augmenter à l'avenir.

La plupart des membres (63 %) sont issus du secteur universitaire, mais 17 % travaillent dans le secteur privé, 13 % pour le compte d'un organisme à but non lucratif et 7 % dans le secteur public. Il faudra atteindre un meilleur équilibre dans les mois et les années à venir, car nous pensons que la collaboration de *toutes* les parties prenantes sera nécessaire pour veiller à ce que l'IA soit produite et utilisée de manière responsable.

L'IAR représente également une diversité intéressante de pays, bien que davantage de pays et d'organisations internationales devraient être représentés à court et moyen terme, en particulier les pays et entités provenant du Sud.

Au 30 novembre, les membres étaient basés dans 17 pays, c'est-à-dire l'Argentine, l'Australie, le Canada (2 personnes), la France (3), l'Allemagne (2), l'Inde (2), l'Italie, le Japon (2), la Corée, le Mexique (2), les Pays-Bas, la Nouvelle-Zélande (2), Singapour, la Slovénie (2), la Suède, le Royaume-Uni (3) et les États-Unis (3).

Ces membres ont été désignés par les 15 membres fondateurs du PMIA ou recommandés par l'UNESCO. Il est utile de mentionner que les membres désignés par les pays membres du PMIA ou recommandés par les organisations internationales agissent en totale indépendance au sein de l'IAR.

Enfin, sept personnes supplémentaires participent aux activités de l'IAR en qualité d'observateurs. L'un d'eux est un représentant de l'OCDE, partenaire stratégique du PMIA, et un autre est un représentant d'un groupe d'experts chargé de conseiller l'OCDE.

Les experts de l'IAR figurent ci-après.

Membres de l'IAI

Membre de l'IAI (les membres du comité directeur sont mentionnés en caractère gras)

Yoshua Bengio (coprésident), Mila – Institut d'intelligence artificielle du Québec
Raja Chatila (coprésident), Université de la Sorbonne
Carolina Aguerre, Center for Technology and Society (CETyS)
Genevieve Bell, Australian National University
Ivan Bratko, University of Ljubljana
Joanna Bryson, Hertie School
Partha Pratim Chakrabarti, Indian Institute of Technology Kharagpur
Jack Clark, OpenAI
Virginia Dignum, Umeå University
Dyan Gibbens, Trumbull Unmanned
Kate Hannah, Te Pūnaha Matatini, University of Auckland
Toshiya Jitsuzumi, Chuo University
Alistair Knott, University of Otago
Pushmeet Kohli, DeepMind
Marta Kwiatkowska, Oxford University
Christian Lemaître Léon, Metropolitan Autonomous University
Vincent C. Müller, Technical University of Eindhoven
Wanda Muñoz, SEHLAC (Mexique)
Alice H. Oh, KAIST School of Computing
Luka Omladič, Institute of Applied Ethics
Julie Owono, Internet Sans Frontières
Dino Pedreschi, University of Pisa
V K Rajah, Advisory Council on the Ethical Use of Artificial Intelligence and Data (Singapour)
Catherine Régis, Université de Montréal
Francesca Rossi, IBM Research
David Sadek, Thales Group
Rajeev Sangal, International Institute of Information Technology Hyderabad
Matthias Spielkamp, Algorithm Watch
Osamu Sudo, Chuo University
Roger Taylor, Centre for Data Ethics and Innovation

Observateurs

Amir Banifatemi, AI Commons
Vilas Dhar, The Patrick J. McGovern Foundation
Marc-Antoine Dilhac, ALGORA Lab
Adam Murray, Réseau d'experts de l'IA à l'OCDE
Karine Perset, OCDE
Stuart Russell, UC Berkeley
Cédric Wachholz, Programme Innovation et transformation numérique, secteur de la Communication et de l'information (UNESCO)

Mandat de l'IAR

Comme évoqué dans l'avant-propos, le travail de l'IAR repose sur une vision de l'IA centrée sur l'humain et juste, équitable, inclusive et respectueuse des droits de la personne et de la démocratie et qui a pour objectif de contribuer positivement à l'intérêt général.

Le mandat de l'IAR s'aligne étroitement sur cette vision et la mission globale du PMIA, c'est-à-dire que l'IAR s'efforce d'encourager le développement, l'utilisation et la gouvernance responsables des systèmes d'IA centrée sur l'humain et d'y contribuer, en accord avec les objectifs de développement durable des Nations Unies.

Il convient de signaler que l'IAR, comme tous les autres groupes de travail, ne fonctionne pas en vase clos dans le cadre du PMIA. En effet, il entend collaborer avec d'autres groupes de travail chaque fois que cela sera indiqué. L'IAR sera par exemple en contact avec le groupe de travail chargé de la gouvernance des données lorsque leurs projets respectifs revêtent des dimensions communes.

Enfin, il doit cependant être souligné qu'à la lumière du contexte international actuel, le groupe d'intervention du PMIA a invité l'IAR à constituer un sous-groupe *ad hoc* sur l'IA et la réponse à la pandémie afin de soutenir le développement et l'utilisation responsables de solutions basées sur l'IA pour lutter contre la COVID-19 et d'autres pandémies futures; ce sous-groupe relève du groupe de travail chargé du développement, de l'utilisation et de la gouvernance responsables de l'IA. Le sous-groupe sur l'IA et la réponse à la pandémie a été créé en juillet 2020. Il est coprésidé par un membre de l'IAR, Alice Oh, et par Paul Paul Suetens, qui ne fait pas partie de l'IAR



Méthodes de travail

L'IAR a été créé il y a moins de six mois pendant une période, il va sans dire, assez difficile.

Le groupe de travail a tenu cinq réunions depuis le début de ses activités. Ces réunions ont permis de discuter du travail que le groupe aurait à entreprendre et du travail qu'il avait amorcé.

Le 31 juillet 2020, les membres de l'IAR se sont réunis en mode virtuel pour discuter de ce que devrait être le nouveau mandat du groupe de travail et pour lancer des idées sur ce que cette nouvelle entité devrait tenter d'accomplir à court terme.

Le 25 août 2020, les membres ont également élaboré le premier projet qu'ils avaient l'intention d'entreprendre en tant que groupe. Étant donné que l'IAR souhaite cerner les lacunes que le groupe de travail pourrait combler au cours des prochains mois sans dupliquer les différentes initiatives nationales et internationales existantes visant à encourager et à garantir le développement et l'utilisation responsables des systèmes d'IA, et afin d'orienter nos travaux sur l'utilisation faite de l'IA pour atteindre les objectifs de développement durable des Nations unies jusqu'à ce jour, les membres de l'IAR ont décidé de procéder à un examen des initiatives nationales et internationales ayant pour objectif :

- de cataloguer les projets menés par diverses parties prenantes afin de promouvoir la recherche et le développement responsables de systèmes et d'applications bénéfiques d'IA;
- d'analyser les initiatives prometteuses qui peuvent contribuer pour beaucoup au développement et à l'utilisation de systèmes et d'applications bénéfiques d'IA et qui pourraient tirer bénéfice d'une collaboration internationale et intersectorielle;
- de recommander de nouvelles initiatives et la manière de les mettre en pratique et de contribuer à promouvoir le développement, l'utilisation et la gouvernance responsables des systèmes d'IA centrée sur l'humain.

En septembre 2020, l'IAR a lancé un [appel de propositions](#) public afin de recruter un consultant capable de l'aider à mener cette étude avant la fin du mois de novembre, soit à temps pour la tenue du Sommet de décembre. Aussi, un comité directeur composé de neuf volontaires de l'IAR (dont les noms figurent dans l'annexe 1) a été constitué pour évaluer les propositions reçues par le groupe de travail et superviser les travaux réalisés par le cabinet ou le groupe retenu.

L'IAR a reçu neuf propositions avant la date limite du 20 septembre. Le 28 septembre, au terme de deux cycles d'évaluation, le comité directeur a arrêté son choix sur The Future Society (TFS) pour mener le premier projet d'IAR sous la supervision du comité directeur.

TFS a rencontré les membres du comité directeur lors d'une réunion de lancement qui a eu lieu le 2 octobre. Cette réunion a permis de développer une compréhension commune des objectifs de ce projet, de définir le plan de travail final du projet et de planifier de futures réunions.

Le 5 octobre 2020, TFS a présenté la portée et le calendrier du projet aux membres de l'IAR, qui ont eu la possibilité de faire des suggestions. Cette réunion a également donné une occasion aux coprésidents d'explorer la possibilité de mettre sur pied de nouveaux comités de travail internes qui pourraient, à court terme, aborder des questions particulières ayant un grand intérêt.

D'autres réunions ont été organisées par le comité directeur et TFS les 12 et 22 octobre pour discuter des versions préliminaires du rapport et de ses différentes composantes, de même que pour planifier l'intervention de TFS dans le cadre de réunions du groupe élargi.

Le 27 octobre 2020, TFS a présenté une ébauche de son rapport à l'IAR. Dans les jours suivants, TFS a organisé trois séances de travail pour donner à des sous-ensembles de l'IAR la possibilité d'obtenir plus de précisions sur les conclusions et les recommandations préliminaires de TFS et d'en discuter.



Le comité directeur s'est réuni à deux autres occasions, les 2 et 12 novembre, pour débattre des recommandations finales qui pourraient être incluses dans le rapport de TFS.

Le 8 novembre, les coprésidents ont envoyé un message aux membres de l'IAR pour les inviter à manifester leur intérêt à faire partie de l'un des cinq nouveaux comités internes que le groupe de travail pourrait éventuellement mettre sur pied dans les prochaines semaines. À la suite des discussions initiales et des recommandations préliminaires figurant dans le rapport de TFS, on a déterminé que ces comités pourraient travailler sur la découverte de médicaments et la science ouverte, la gouvernance et la transparence des réseaux sociaux, les changements climatiques, l'éducation et les questions portant sur les moyens de gouvernance.

Le rapport quasi final de TFS a été présenté à l'IAR le 16 novembre et déposé le 23 novembre. Ce rapport fera l'objet d'une présentation à l'occasion du Sommet. De plus, les recommandations ont servi à alimenter le rapport de 10 à 15 pages que vous lisez en ce moment. The Steering Committee met two other times, on November 2 and 12, to debate the final recommendations that could be included in TFS report.



Calendrier du groupe de travail

JUILLET

Présentation des coprésidents (le 7).

Première réunion du groupe de travail (le 31): présentations et accord sur le mandat; les vice-présidents ont invité les membres de l'IAR à les aviser s'ils souhaitaient porter certains points à l'attention.

AOÛT

Deuxième réunion du groupe de travail (le 25) – discussions autour du premier livrable qui serait produit par l'IAR.

SEPTEMBRE

Publication d'un [blogue de présentation de l'IAR sur le site Web de l'OCDE](#), incluant un appel de propositions et un mandat.

Premier cycle d'évaluation des propositions reçues par l'IAR (le 24).

Deuxième cycle d'évaluation des propositions reçues par l'IAR et désignation de TFS (le 28).

OCTOBRE

Réunion de lancement du projet avec TFS et le comité directeur (le 2).

Troisième réunion du groupe de travail (le 5) : présentation de TFS, de son mandat et de son plan de travail; discussions concernant la possible création de comités de travail internes.

Réunion du comité directeur et de TFS (le 12).

Réunion de tous les coprésidents afin de comparer les progrès réalisés et de discuter de possibles synergies (le 16).

Réunion du comité directeur et de TFS (le 22).

Quatrième réunion du groupe de travail : (le 27) : présentation de la première ébauche du rapport et des recommandations de TFS.

NOVEMBRE

Séances de travail dirigées par TFS (le 4 et le 5) : les membres de l'IAR ont eu l'occasion de discuter du rapport de TFS au cours de trois séances de travail organisées dans trois fuseaux horaires différents.

Cinquième réunion du groupe de travail (le 16) : présentation de la dernière ébauche du rapport de TFS; discussions sur l'intérêt suscité par la possible création de cinq comités.

Dépôt du rapport de TFS au Centre d'expertise de Montréal en intelligence artificielle (« CEIMIA ») (le 23).

Réunion de tous les coprésidents avant le Sommet (le 27).

DECEMBRE

Présentation des résultats finaux et atelier ouvert sur les prochains projets au Sommet.



Recommandations préliminaires et résultats pour le Sommet

L'examen des initiatives nationales et internationales menées par TFS est le premier livrable commandé par l'IAR. Cet examen contient de précieuses leçons et recommandations qui guideront les travaux que l'IAR réalisera dans les mois et les années à venir.

Élaboré par TFS avec le soutien d'IAR, en octobre et novembre 2020, le rapport présente 30 initiatives prometteuses qui contribuent à encourager le développement ou l'utilisation de l'IA responsable ou à atteindre les objectifs de développement durable des Nations Unies. TFS a alors procédé à l'analyse de ces initiatives pour déterminer leur capacité à aider le PMIA à réaliser ses objectifs; leur efficacité et leur conformité aux Principes sur l'intelligence artificielle de l'OCDE et au programme lié aux objectifs de développement durable des Nations Unies; leur évolutivité dans les diverses régions géographiques et secteurs; et pour savoir si oui ou non elles sont aussi représentatives que possible du point de vue des régions géographiques, des secteurs, des parties prenantes et des groupes cibles. La dernière section du rapport de TFS s'appuie sur les occasions et les lacunes cernées par TFS pour proposer quatre domaines d'intervention future et neuf recommandations qui contribueront à orienter le programme de l'IAR pour l'avenir.

Selon nous, trois de ces domaines et les recommandations qui les accompagnent sont particulièrement pertinents :

1. L'IA a un large éventail d'applicabilité et est donc susceptible d'avoir une influence sur la plupart des problèmes les plus pressants auxquels est confrontée l'humanité; elle peut constituer une force positive afin d'atténuer les changements climatiques ou prévoir la prochaine pandémie, tout comme elle peut exacerber les défis mondiaux comme le montre l'importance prise par la désinformation. L'ampleur des applications potentielles de l'IA responsable pose un défi en matière d'établissement des priorités. **Ainsi, l'IAR devrait mettre sur pied des comités axés sur le traitement des problèmes pressants cernés.**
2. Pour bâtir un écosystème capable de soutenir et de stimuler le changement, il est nécessaire de disposer d'outils et de cadres de gouvernance qui promeuvent la transparence et modifient les incitations et les comportements dans l'ensemble de la société afin de contribuer à l'adoption de pratiques d'IA responsables. Il convient aussi de mettre en place une collaboration et une coopération systématiques dans l'ensemble de l'écosystème ainsi qu'un mécanisme permettant de relier les initiatives transversales au niveau du domaine. Enfin, pour que les gouvernements puissent mettre en place ces outils et ces cadres à grande échelle, il est nécessaire de renforcer les capacités des décideurs politiques ainsi que les boucles de rétroaction entre les gouvernements et les autres acteurs de l'écosystème. **Ainsi, l'IAR devrait mettre sur pied un comité axé sur les enjeux et les moyens de gouvernance.**
3. Il a été difficile pour de nombreuses initiatives de l'écosystème de l'IA responsable de recueillir des données représentatives pour étayer leurs activités. Ce défaut d'inclusion montre que les initiatives, les parties prenantes et les gouvernements n'ont pas la capacité de faire participer un groupe plus large à la transition technologique et, par conséquent, d'élaborer conjointement des solutions innovantes pour gérer les occasions et remédier aux risques. En fin de compte, ce défaut d'inclusion risque de saper l'efficacité de nombreuses initiatives d'IA responsable et de les décrédibiliser, ainsi que leur capacité à prendre de l'ampleur. **Ainsi, l'IAR devrait élaborer et faire connaître les bonnes pratiques en matière de diversité et d'inclusion.**



Priorités de l'IA pour les trois à six prochains mois

Le travail mené par TFS pour l'IA a été extrêmement utile et nous avons maintenant pour objectif d'inventorier les efforts déployés par la communauté internationale qui vont au-delà du travail déjà accompli et catalogué dans le rapport de TFS.

Plus précisément, nous souhaitons mettre l'accent sur les domaines dans lesquels la structure actuelle du marché et les politiques gouvernementales actuelles sont insuffisantes pour atteindre les objectifs de développement durable des Nations Unies ou d'autres objectifs clés.

Nous avons décidé, pour y parvenir, de nous pencher sur quelques thèmes d'action concrets liés à ces objectifs qui intéressent les membres de l'IA et qu'ils sont prêts à explorer dans le cadre des comités de travail internes :

1. Le comité sur la découverte de médicaments et la science ouverte (qui est lié à l'objectif de développement durable 3 : Bonne santé et bien-être) pourrait examiner comment créer un contexte favorable pour que l'IA contribue à la découverte de médicaments de manière ouverte et équitable, en privilégiant les besoins de santé publique à l'échelle internationale, p. ex. pour lutter contre la COVID-19 ou la résistance aux antibiotiques. Il pourrait également analyser comment les efforts déployés en recherche et développement doivent être organisés et quelles devraient être les règles d'engagement pour garantir que l'homologation des médicaments qui en découle soit favorable aux pays pauvres et aux pays donateurs.
2. Le comité sur les changements climatiques (objectif de développement durable 13 : Mesures relatives à la lutte contre les changements climatiques) pourrait élaborer des approches pratiques axées sur la collaboration pour lutter contre les changements climatiques au moyen de l'IA (p. ex., faire en sorte que l'IA rende les énergies renouvelables à zéro carbone aussi productives que les fournisseurs traditionnels d'hydrocarbures) d'une part, et pour rendre l'IA plus respectueuse de l'environnement, d'autre part (p. ex., pour mieux évaluer les répercussions environnementales de l'apprentissage machine).
3. Le comité sur l'IA et l'éducation (objectif de développement durable 4 : Éducation de qualité) pourrait définir des projets de collaboration dont la mise en œuvre contribuerait à : a) optimiser les avantages de l'IA pour la gestion de l'éducation et la prestation de services éducatifs, en donnant les moyens à l'enseignement et aux enseignants d'améliorer l'apprentissage et l'évaluation de l'apprentissage, en offrant des occasions d'apprentissage tout au long de la vie pour tous, etc.; b) aborder des questions transversales comme la promotion de l'utilisation équitable et inclusive de l'IA dans l'éducation, la formation des étudiants afin qu'ils deviennent des producteurs et des utilisateurs responsables de l'IA ou le suivi de l'incidence de l'IA sur l'éducation.
4. Le comité sur la gouvernance et la transparence des réseaux sociaux (objectif de développement durable 16 : Paix, justice et institutions efficaces) pourrait se focaliser sur l'élaboration de principes et d'outils permettant de faire la distinction entre ce qui est acceptable ou non sur le plan de la manipulation dans la publicité et les réseaux sociaux. Par la suite, des programmes conjoints de recherche pourraient être entrepris en vue de développer des outils d'IA pour atteindre des objectifs comme déceler les fausses nouvelles ou les partis pris ou signaler des messages démagogiques.

Nous pensons que, pour promouvoir l'intérêt public dans chacun de ces domaines, il faut prendre en considération les enjeux de gouvernance qui sont fondamentalement liés à la manière dont les progrès proposés pourraient être utilement déployés. Par exemple, les initiatives à financement international sur l'utilisation de l'IA pour la découverte de médicaments contre la résistance aux antibiotiques exigeront que des règles d'engagement soient définies avec les organisations privées ou publiques qui entreprennent la recherche et le développement proprement dit, par exemple en ce qui concerne le partage des données, les normes d'échange d'informations et l'homologation favorables à l'intérêt public (p. ex., protéger la planète entière contre la prochaine pandémie).

C'est la raison pour laquelle il a été décidé de créer un cinquième comité, transversal, qui sera consacré à la question de la gouvernance dans le développement d'une utilisation de l'IA orientée vers le bien public.



Au-delà des enjeux de gouvernance soulevés ci-dessus, le comité sur les enjeux et les moyens de gouvernance pourrait travailler sur les mécanismes de certification, d'évaluation et de vérification utilisés pour évaluer les systèmes d'IA en ce qui concerne la responsabilité et la fiabilité en s'appuyant sur des paramètres tels que la responsabilité, la transparence, la sécurité, la justice, le respect des droits de la personne et la promotion de l'équité.



Vision à long terme

D'autres thèmes transversaux viendront, à plus long terme, apporter de la couleur aux travaux des comités thématiques désignés ci-dessus.

Parmi ces thèmes figurent l'inclusion et la diversité, notamment en ce qui concerne les effets (positifs ou négatifs) que l'IA pourrait avoir sur les personnes et les groupes les plus vulnérables parmi nous.

Et nous le savons, certaines personnes, certains groupes ou certains pays risquent de ne pas bénéficier des avantages rendus possibles par l'IA parce qu'ils n'auront pas accès à certains des outils les plus prometteurs qui seront développés (par exemple, seuls les pays ou les écoles riches pourraient être en mesure d'utiliser les technologies d'IA les plus récentes pour aider les étudiants). Il est manifeste également que par la manière dont ils ont été conçus et développés, certains outils d'IA pourraient conduire à l'exclusion de personnes ou de groupes souvent déjà fragiles, à la reproduction ou même au renforcement de préjugés qui existent déjà dans nos sociétés.

La question de la diversité et de l'inclusion sera l'un des thèmes abordés par l'IAR à moyen et à long terme. Comme le rapport de TFS le suggère, l'IAR pourrait contribuer à façonner et à diffuser les bonnes pratiques en matière de diversité et d'inclusion. Une stratégie pourrait permettre de déterminer dans quelle mesure des segments de la société ou des zones géographiques sont actuellement sous-représentés ou exclus de l'écosystème de l'IA responsable. Il conviendrait également d'encourager le libre accès à l'information et aux infrastructures qui contribuerait à se débarrasser de toute entrave à la communication entre les régions, les groupes sociaux et les disciplines. Il faudra élaborer des processus permettant de garantir le déploiement d'applications qui bénéficieront aux groupes et sociétés marginalisés et de veiller à ce que les développeurs proviennent d'horizons plus divers et respectent des codes de conduite ou visant à élaborer des mécanismes dont la mise en place contribuera à réduire ou à éliminer les répercussions négatives que les décisions algorithmiques peuvent avoir sur certains d'entre nous.

En plus des travaux des cinq comités qui seront lancés très prochainement, l'IAR envisage également d'approfondir des sujets particuliers à l'avenir, soit par l'essaimage des comités actuels ou par la création de nouveaux comités qui se concentrent sur d'autres objectifs de développement durable, comme la paix ou la faim, et qui pourraient être instaurés au milieu de l'année 2021; des mécanismes généraux de financement international et de gouvernance de l'IA bénéficiant d'un financement public pourraient également être mis au point pour des projets pour le bien social.

Les travaux sur les mécanismes et les processus de gouvernance se poursuivront en parallèle. En particulier, l'IAR veillera à coordonner étroitement ses activités avec celles de l'OCDE et du groupe de travail pour une IA digne de confiance du Réseau d'experts de l'IA à l'OCDE (ONE AI). L'OCDE décrit ce groupe de travail comme un groupe qui « aide à identifier les idées prometteuses et les bonnes pratiques afin de mettre en œuvre les cinq Principes de l'OCDE sur l'IA pour des systèmes d'IA digne de confiance : 1) croissance inclusive, développement durable et bien-être; 2) droits de la personne, valeurs démocratiques et justice; 3) transparence et explicabilité; 4) solidité, sécurité et sûreté; et 5) responsabilité. Ces pratiques comprennent des codes de conduites, des lignes directrices, des normes, des certifications, des cadres de gouvernance d'entreprise, des approches de gestion des risques, la recherche technique, des outils logiciels ainsi que des outils de renforcement des compétences et de la sensibilisation. L'objectif est de déterminer les orientations pratiques et les approches procédurales communes pour aider les acteurs et les décideurs en IA à adopter une IA digne de confiance, notamment en soulignant la manière dont les outils et les approches peuvent varier d'un contexte opérationnel à l'autre. »

Enfin, suivant une recommandation formulée par TFS dans son rapport, l'IAR travaillera de concert avec des organisations internationales comme l'OCDE, l'Organisation mondiale de la Santé et l'UNESCO pour s'assurer que le groupe de travail recueille des points de vue représentatifs émanant de groupes marginalisés et provenant du Sud. Le rôle de l'IAR dans ces partenariats consistera à faire participer de manière proactive les groupes traditionnellement marginalisés à ces dialogues, et à soutenir les initiatives qui favorisent l'apprentissage de connaissances de base en matière d'IA afin de donner au public les moyens de participer pleinement.



Annexe 1

Membres du comité directeur du projet

Raja Chatila (coprésident), Université de la Sorbonne
Dyan Gibbens, Trumbull Unmanned
Kate Hannah, Te Pūnaha Matatini, University of Auckland
Toshiya Jitsuzumi, Chuo University
Vincent C. Müller, Technical University of Eindhoven
Wanda Muñoz, SEHLAC (Mexique)
Dino Pedreschi, University of Pisa
Catherine Régis, Université de Montréal
Francesca Rossi, IBM Research

