

Réunion du Conseil au niveau des Ministres, 22-23 mai 2019

RECOMMANDATION DU CONSEIL SUR L'INTELLIGENCE ARTIFICIELLE

(Adoptée par le Conseil au niveau des Ministres, le 22 mai 2019)

JT03447953

LE CONSEIL,

CONSIDÉRANT l'Article 5 b) de la Convention relative à l'Organisation de Coopération et de Développement Économiques en date du 14 décembre 1960 ;

CONSIDÉRANT les Principes directeurs de l'OCDE à l'intention des entreprises multinationales [[OECD/LEGAL/0144](#)] ; Recommandation du Conseil concernant les Lignes directrices régissant la protection de la vie privée et les flux transfrontières de données de caractère personnel [[OECD/LEGAL/0188](#)] ; Recommandation du Conseil relative aux Lignes directrices régissant la politique de cryptographie [[OECD/LEGAL/0289](#)] ; la Recommandation du Conseil relative à un accès élargi et une exploitation plus efficace concernant les informations du secteur public [[OECD/LEGAL/0362](#)] ; la Recommandation du Conseil sur la gestion du risque de sécurité numérique pour la prospérité économique et sociale [[OECD/LEGAL/0415](#)] ; la Recommandation du Conseil sur la protection du consommateur dans le contexte du commerce électronique [[OECD/LEGAL/0422](#)] ; la Déclaration sur l'économie numérique : innovation, croissance et prospérité sociale (Déclaration de Cancún) [[OECD/LEGAL/0426](#)] ; la Déclaration sur le renforcement des PME et de l'entrepreneuriat au service de la productivité et de la croissance inclusive [[OECD/LEGAL/0439](#)] ; ainsi que la Déclaration ministérielle de 2016 pour Mettre en place des marchés du travail plus résilients et inclusifs, adoptée lors de la Réunion ministérielle de l'OCDE sur l'emploi et le travail ;

CONSIDÉRANT les Objectifs de développement durable énoncés dans le Programme de développement durable à l'horizon 2030 adopté par l'Assemblée générale des Nations Unies (A/RES/70/1), ainsi que la Déclaration universelle des droits de l'homme de 1948 ;

CONSIDÉRANT les travaux importants menés sur l'intelligence artificielle (ci-après dénommée « IA ») dans le cadre d'autres instances internationales gouvernementales et non gouvernementales ;

RECONNAISSANT que l'IA a des incidences mondiales généralisées et profondes qui transforment les sociétés, les secteurs économiques et le monde du travail, une tendance qui devrait s'accroître à l'avenir ;

RECONNAISSANT que l'IA promet d'améliorer la prospérité et le bien-être des individus, de contribuer à une activité économique mondiale dynamique et durable, de stimuler l'innovation et la productivité, et d'aider à affronter les grands défis planétaires ;

RECONNAISSANT que, dans le même temps, ces transformations pourraient avoir des effets disparates d'une société et d'une économie à l'autre et en leur sein, notamment en termes de mutations économiques, de concurrence, de transitions sur les marchés du travail, d'inégalités, et de conséquences sur la démocratie et les droits de l'homme, la protection de la vie privée et la confidentialité des données, et la sécurité numérique ;

RECONNAISSANT que la confiance est un déterminant essentiel de la transformation numérique ; que, bien qu'il soit difficile de prévoir la nature des futures applications de l'IA et leurs incidences, la confiance dans la fiabilité des systèmes d'IA est un facteur clé de la diffusion et de l'adoption de l'IA ; et qu'un débat public bien informé de l'ensemble de la société est nécessaire afin de concrétiser le potentiel entier de cette technologie tout en limitant les risques qui y sont associés ;

SOULIGNANT que certains cadres juridiques, réglementaires et politiques nationaux existants s'appliquent d'ores et déjà à l'IA, y compris ceux qui ont trait aux droits de

l'homme, à la protection des consommateurs et des données à caractère personnel, aux droits de propriété intellectuelle, à la conduite responsable des entreprises et à la concurrence, tout en sachant qu'il pourrait être nécessaire d'évaluer la pertinence de certains cadres et de concevoir de nouvelles approches ;

RECONNAISSANT que compte tenu de l'évolution et de la mise en œuvre rapides de l'IA, il est nécessaire de bâtir un cadre d'action stable, qui privilégie une IA digne de confiance centrée sur l'humain, qui favorise la recherche, préserve les incitations économiques à innover et qui s'applique à l'ensemble des parties prenantes selon leurs responsabilités et le contexte ;

CONSIDÉRANT qu'il est essentiel de saisir les opportunités offertes et de relever les défis posés par les applications de l'IA et de permettre aux parties prenantes de s'investir pour favoriser l'adoption d'une IA digne de confiance dans la société et de faire de la fiabilité de l'IA un atout dans le marché mondial.

Sur proposition du Comité de la politique de l'économie numérique :

I. CONVIENT qu'aux fins de la présente Recommandation, les termes techniques suivants s'entendent comme suit :

- *Système d'IA* : Un système d'intelligence artificielle (ou système d'IA) est un système automatisé qui, pour un ensemble donné d'objectifs définis par l'homme, est en mesure d'établir des prévisions, de formuler des recommandations, ou de prendre des décisions influant sur des environnements réels ou virtuels. Les systèmes d'IA sont conçus pour fonctionner à des degrés d'autonomie divers.
- *Cycle de vie d'un système d'IA* : Le cycle de vie d'un système d'IA comporte plusieurs phases : *i)* la phase de « conception, données et modèles », séquence dépendante du contexte comprenant la planification et la conception, la collecte et le traitement des données, ainsi que la construction du modèle ; *ii)* la phase de « vérification et validation » ; *iii)* la phase de « déploiement » ; et *iv)* la phase d'« exploitation et (de) suivi ». Ces phases présentent souvent un caractère itératif et ne suivent pas nécessairement un ordre séquentiel.

La décision de mettre un terme à l'utilisation d'un système d'IA peut intervenir à n'importe quel stade de la phase d'exploitation et de suivi.

- *Connaissances en matière d'IA* : Les connaissances en matière d'IA désignent les compétences et les ressources, à l'instar des données, du code, des algorithmes, des modèles, de la recherche, du savoir-faire, des programmes de formation, de la gouvernance, des processus et des meilleures pratiques nécessaires pour comprendre le cycle de vie des systèmes d'IA et y prendre part.
- *Acteurs de l'IA* : Les acteurs de l'IA sont les parties jouant un rôle actif dans le cycle de vie d'un système d'IA, y compris les organisations et les individus qui déploient ou exploitent l'IA.
- *Parties prenantes* : On entend par parties prenantes l'ensemble des organisations et des individus intervenant dans les systèmes d'IA ou concernés par ces systèmes, que ce soit directement ou indirectement. Les acteurs de l'IA constituent un sous-ensemble des parties prenantes.

Section 1 : Principes d'une approche responsable en appui d'une IA digne de confiance

- II. **RECOMMANDE** que les Membres et non-Membres adhérant à la présente Recommandation (ci-après dénommés les « Adhérents ») promeuvent et mettent en œuvre les Principes d'une approche responsable en appui d'une IA digne de confiance suivants, qui intéressent l'ensemble des parties prenantes.
- III. **APPELLE** tous les acteurs de l'IA à promouvoir et mettre en œuvre, selon leurs rôles respectifs, les Principes suivants pour une approche responsable en appui d'une IA digne de confiance.
- IV. **SOULIGNE** que les principes suivants sont complémentaires et doivent être considérés comme un tout.

1.1. Croissance inclusive, développement durable et bien-être

Les parties prenantes devraient adopter de manière proactive une approche responsable en soutien d'une IA digne de confiance afin de tendre vers des résultats bénéfiques pour les individus et la planète, tels que le renforcement des capacités humaines et le renforcement de la créativité humaine, l'inclusion des populations sous-représentées, la réduction des inégalités économiques, sociales, entre les sexes et autres, et la protection des milieux naturels, favorisant ainsi la croissance inclusive, le développement durable et le bien-être.

1.2. Valeurs centrées sur l'humain et équité

- a) Les acteurs de l'IA devraient respecter l'état de droit, les droits de l'homme et les valeurs démocratiques tout au long du cycle de vie des systèmes d'IA. Ces droits et valeurs comprennent la liberté, la dignité et l'autonomie, la protection de la vie privée et des données, la non-discrimination et l'égalité, la diversité, l'équité, la justice sociale, ainsi que les droits des travailleurs reconnus à l'échelle internationale.
- b) Pour ce faire, les acteurs de l'IA devraient instituer des garanties et des mécanismes, tels que l'attribution de la capacité de décision finale à l'homme, qui soient adaptés au contexte et à l'état de l'art.

1.3. Transparence et explicabilité

Les acteurs de l'IA devraient s'engager à assurer la transparence et une divulgation responsable des informations liées aux systèmes d'IA. À cet effet, ils devraient fournir des informations pertinentes, adaptées au contexte et à l'état de l'art, afin:

- i. de favoriser une compréhension générale des systèmes d'IA,
- ii. d'informer les parties prenantes de leurs interactions avec les systèmes d'IA, y compris dans la sphère professionnelle,
- iii. de permettre aux personnes concernées par un système d'IA d'en appréhender le résultat, et,
- iv. de permettre aux personnes subissant les effets néfastes d'un système d'IA de contester les résultats sur la base d'informations claires et facilement compréhensibles sur les facteurs, et sur la logique ayant servi à la formulation de prévisions, recommandations ou décisions.

1.4. Robustesse, sûreté et sécurité

- a) Les systèmes d'IA devraient être robustes, sûrs et sécurisés tout au long de leur cycle de vie, de sorte que, dans des conditions d'utilisation normales ou prévisibles, ou en cas d'utilisation abusive ou de conditions défavorables, ils soient à même de fonctionner convenablement, et ne fassent pas peser un risque de sécurité démesuré.
- b) Pour ce faire, les acteurs de l'IA devraient veiller à la traçabilité, notamment pour ce qui est des ensembles de données, des processus et des décisions prises au cours du cycle de vie des systèmes d'IA, afin de permettre l'analyse des résultats produits par lesdits systèmes d'IA et le traitement des demandes d'information, compte tenu du contexte et de l'état de l'art de la technologie.
- c) Les acteurs de l'IA devraient, selon leurs rôles respectifs, le contexte et leur capacité à agir, appliquer de manière continue une approche systématique de la gestion du risque, à chaque phase du cycle de vie des systèmes d'IA, afin de gérer les risques y afférents, notamment ceux liés au respect de la vie privée, à la sécurité numérique, à la sûreté et aux biais.

1.5. Responsabilité

Les acteurs de l'IA devraient être responsables du bon fonctionnement des systèmes d'IA et du respect des principes exposés ci-dessus, selon leurs rôles, le contexte et l'état de l'art.

Section 2 : Politiques nationales et coopération internationale à l'appui d'une IA digne de confiance

- V. **RECOMMANDE** que les Adhérents mettent en œuvre les recommandations suivantes, en respectant les principes énoncés à la section 1, dans le cadre de leurs politiques nationales et de la coopération internationale, en prêtant une attention particulière aux petites et moyennes entreprises (PME).

2.1. Investir dans la recherche et le développement en matière d'IA

- a) Les pouvoirs publics devraient envisager des investissements publics à long terme et encourager les investissements privés dans la recherche et le développement, notamment interdisciplinaire, afin de stimuler l'innovation dans une IA digne de confiance, relevant des défis techniques importants, ainsi que sur les implications sociales, juridiques et éthiques et les politiques liées à l'IA.
- b) Les pouvoirs publics devraient par ailleurs envisager des investissements publics et encourager les investissements privés dans des ensembles de données en libre accès qui soient représentatifs et qui garantissent la protection de la vie privée et des données, pour soutenir un environnement de recherche et développement en matière d'IA sans partialité induite et renforcer l'interopérabilité et l'utilisation de normes.

2.2. Favoriser l'instauration d'un écosystème numérique pour l'IA

Les pouvoirs publics devraient favoriser le développement et l'accessibilité d'un écosystème numérique à l'appui d'une IA digne de confiance. Cet écosystème se compose notamment des technologies et infrastructures numériques et des mécanismes de partage des connaissances en matière d'IA, en fonction des besoins. À cet égard, les pouvoirs publics devraient envisager de promouvoir des mécanismes, tels que les fiducies de

données (« data trusts »), pour favoriser le partage des données de façon sûre, équitable, légale et éthique.

2.3 Façonner un cadre d'action favorable à l'IA

- a) Les pouvoirs publics devraient promouvoir l'instauration d'un cadre d'action favorable qui soutienne une transition souple du stade de recherche et développement à celui de déploiement de systèmes d'IA dignes de confiance. À cette fin, ils devraient envisager le recours à l'expérimentation, afin de fournir un environnement contrôlé dans lequel les systèmes d'IA peuvent être testés et monter en puissance, selon les besoins.
- b) Les pouvoirs publics devraient examiner et adapter, selon les besoins, leur cadres politiques et réglementaires et leurs mécanismes d'évaluation applicables aux systèmes d'IA, afin d'encourager l'innovation et la concurrence dans le développement d'une IA digne de confiance.

2.4. Renforcer les capacités humaines et préparer la transformation du marché du travail

- a) Les pouvoirs publics devraient travailler en étroite collaboration avec les parties prenantes en vue de préparer la transformation du monde du travail et de la société. Ils devraient donner aux personnes les moyens d'utiliser et d'interagir efficacement avec les systèmes d'IA au travers de leurs différentes applications, notamment en les dotant des compétences nécessaires.
- b) Les pouvoirs publics devraient prendre des mesures, y compris en recourant au dialogue social, pour assurer une transition équitable des travailleurs au fur et à mesure du déploiement d'IA, notamment par le biais de programmes de formation tout au long de la vie active, du soutien aux personnes affectées par les suppressions de postes et de l'accès aux nouvelles opportunités sur le marché du travail.
- c) Les pouvoirs publics devraient par ailleurs travailler en étroite collaboration avec les parties prenantes pour promouvoir l'utilisation responsable de l'IA au travail, renforcer la sécurité des travailleurs et la qualité des emplois, favoriser l'entrepreneuriat et la productivité, et veiller à ce que les avantages de l'IA soient partagés largement et équitablement.

2.5. Favoriser la coopération internationale au service d'une IA digne de confiance

- a) Les pouvoirs publics, y compris ceux des pays en voie de développement, en association avec les parties prenantes, devraient coopérer activement afin de faire progresser la mise en œuvre de ces principes et de progresser dans une approche responsable en appui d'une IA digne de confiance.
- b) Les pouvoirs publics devraient travailler de concert, au sein de l'OCDE et d'autres instances mondiales et régionales pertinentes, en vue de favoriser le partage des connaissances en matière d'IA, selon les besoins. Ils devraient encourager les initiatives multipartites internationales, intersectorielles et ouvertes afin de bâtir une expertise à long terme en matière d'IA.
- c) Les pouvoirs publics devraient encourager l'élaboration de normes techniques internationales multipartites fondées sur la recherche de consensus, au service d'une IA interopérable et digne de confiance.
- d) Les pouvoirs publics devraient en outre encourager la mise au point et l'utilisation, pour leurs propres besoins, d'indicateurs comparables au plan international, afin de

mesurer la recherche et le développement dans le domaine de l'IA et le déploiement de l'IA, et de constituer la base factuelle nécessaire au suivi des progrès quant à la mise en œuvre des principes exposés dans les présentes.

- VI. INVITE** le Secrétaire général et les Adhérents à diffuser la présente Recommandation.
- VII. INVITE** les non-Adhérents à tenir dûment compte de la présente Recommandation et à y adhérer.
- VIII. CHARGE** le Comité de la politique de l'économie numérique :
- a) de poursuivre ses travaux importants sur l'intelligence artificielle en s'appuyant sur la présente Recommandation et en prenant en considération les travaux menés au sein d'autres instances internationales, et de poursuivre la mise au point du cadre de mesure à l'appui de politiques en matière d'IA fondées sur des données empiriques ;
 - b) d'élaborer et de diffuser, de manière itérative, des orientations pratiques quant à la mise en œuvre de la présente Recommandation, et d'informer le Conseil des progrès réalisés d'ici à fin décembre 2019 ;
 - c) de faire office de forum d'échange d'informations sur les politiques et les activités ayant trait à l'IA, notamment sur les retours d'expériences de mise en œuvre de la présente Recommandation, et de favoriser un dialogue multipartite et interdisciplinaire afin de promouvoir la confiance dans l'IA et son adoption; et
 - d) d'assurer le suivi, en consultation avec les autres Comités concernés, de la mise en œuvre de la présente Recommandation et d'en faire rapport au Conseil au plus tard dans les cinq ans suivant son adoption, puis régulièrement par la suite.