

**Réunion du Conseil au niveau des Ministres, 2-3 mai 2024**

**RECOMMANDATION RÉVISÉE DU CONSEIL SUR L'INTELLIGENCE  
ARTIFICIELLE**

**(Adoptée par le Conseil au niveau des Ministres le 3 mai 2024)**

**JT03542984**

## **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 normes élaborées par l'OCDE dans les domaines de la protection de la vie privée, de la sécurité numérique, de la protection des consommateurs et de la conduite responsable des entreprises ;

**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 de cette technologie tout en limitant les risques qui y sont associés ;

**SOULIGNANT** que certains cadres juridiques, réglementaires et politiques nationaux et internationaux 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, favorise la recherche, préserve les incitations économiques à innover et 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 faire de la fiabilité de l'IA un atout dans le marché mondial.

### **Sur proposition du Comité de la politique du numérique :**

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

- *Système d'IA* : Un système d'intelligence artificielle est un système automatisé qui, pour des objectifs explicites ou implicites, déduit, à partir d'entrées reçues, comment générer des résultats en sortie tels que des prévisions, des contenus, des recommandations ou des décisions qui peuvent influencer sur des environnements physiques ou virtuels. Différents systèmes d'IA présentent des degrés variables d'autonomie et d'adaptabilité après déploiement.
- *Cycle de vie d'un système d'IA* : Le cycle de vie d'un système d'IA comporte généralement plusieurs phases, à savoir : la phase de planification et de conception ; la phase de collecte et de traitement des données ; la phase de construction du ou des modèle(s) et/ou d'adaptation du ou des modèle(s) existant(s) à des tâches spécifiques ; la phase de test, d'évaluation, de vérification et de validation ; la phase de mise à disposition pour utilisation/déploiement ; la phase d'exploitation et de suivi ; et la phase de mise hors service. 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.
- *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.
- *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, y compris pour ce qui est de la gestion des risques.
- *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 devraient 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 appui d'une IA digne de confiance afin de tendre vers des résultats bénéfiques pour les individus et la planète, tels que l'augmentation des capacités humaines et le développement de la créativité, 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 bien-être, le développement durable et la durabilité environnementale.

### **1.2. Respect de l'état de droit, des droits humains et des valeurs démocratiques, y compris de l'équité et de la vie privée**

a) Les acteurs de l'IA devraient respecter l'état de droit, les droits humains, les valeurs démocratiques et les valeurs centrées sur l'humain, tout au long du cycle de vie des systèmes d'IA. Cela comprend la non-discrimination et l'égalité, la liberté, la dignité, l'autonomie des individus, la protection de la vie privée et des données, la diversité, l'équité, la justice sociale, ainsi que les droits des travailleurs reconnus à l'échelle internationale. Cela recouvre aussi la lutte contre la mésinformation et la désinformation amplifiées par l'IA, dans le respect de la liberté d'expression et des autres droits et libertés protégés par le droit international en vigueur.

b) Pour ce faire, les acteurs de l'IA devraient instituer des garanties et des mécanismes, tels que la capacité d'intervention et de surveillance humaines, y compris pour faire face aux risques découlant d'utilisations à des fins autres que celles prévues ou d'une mauvaise utilisation délibérée ou involontaire, d'une manière qui soit adaptée 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, y compris de leurs capacités et de leurs limites,
- ii. d'informer les parties prenantes de leurs interactions avec les systèmes d'IA, y compris dans la sphère professionnelle,
- iii. de donner, lorsque cela est réalisable et utile, des informations claires et facilement compréhensibles sur les sources des données/entrées, les facteurs, les processus et/ou la logique ayant abouti à la prévision, au contenu, à la recommandation ou à la décision, pour permettre aux personnes concernées par un système d'IA d'en appréhender le résultat, et
- iv. de fournir des informations permettant aux personnes qui subissent les effets néfastes d'un système d'IA d'en contester les résultats.

### **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 des risques de sûreté et/ou de sécurité démesurés.

b) Des mécanismes devraient être en place, le cas échéant, afin de garantir que, dans l'éventualité où des systèmes d'IA risqueraient de causer des préjudices injustifiés ou présenteraient un comportement indésirable, ils puissent être neutralisés, réparés et/ou mis hors service en toute sécurité, en tant que de besoin.

c) Des mécanismes devraient également être en place, lorsque cela est techniquement réalisable, pour renforcer l'intégrité de l'information, tout en veillant au respect de la liberté d'expression.

### 1.5. Responsabilité

- a) 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.
- 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 issus desdits systèmes d'IA et le traitement des demandes d'information, compte tenu du contexte et de l'état de l'art.
- 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, et assurer la conduite responsable des entreprises afin de gérer les risques y afférents, notamment, en tant que de besoin, par la coopération entre les différents acteurs de l'IA, les pourvoyeurs de connaissances et de ressources liées à l'IA, les utilisateurs de systèmes d'IA et les autres parties prenantes. Ces risques sont ceux liés aux biais préjudiciables, au non-respect des droits humains tels que la sûreté, la sécurité et la protection de la vie privée, et au non-respect des droits des travailleurs et des droits de propriété intellectuelle.

## Section 2 : Politiques nationales et coopération internationale en 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 et dans la science ouverte, notamment les efforts interdisciplinaires, afin de stimuler l'innovation dans une IA digne de confiance, axée sur les 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 outils en *open source* et 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 qui soit exempt de biais préjudiciables, et renforcer l'interopérabilité et l'utilisation de normes.

### 2.2. Favoriser l'instauration d'un écosystème inclusif propice à l'IA

Les pouvoirs publics devraient favoriser le développement et l'accessibilité d'un écosystème numérique inclusif, dynamique, durable et interopérable en appui d'une IA digne de confiance. Cet écosystème se compose, entre autres, des données, des technologies d'IA, des infrastructures de calcul et de connectivité 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 fiduciaires de données (« data trusts »), pour favoriser le partage des données de façon sûre, équitable, légale et éthique.

### **2.3. Façonner un cadre d'action et de gouvernance interopérable favorable à l'IA**

- a) Les pouvoirs publics devraient promouvoir l'instauration d'un cadre d'action souple qui soutienne la transition du stade de recherche et développement à celui de déploiement et d'exploitation 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 déployés à une plus grande échelle, selon les besoins. Ils devraient en outre adopter des approches axées sur les résultats qui offrent de la souplesse dans la réalisation des objectifs de gouvernance et favoriser la coopération entre les pays et territoires et en leur sein, afin de promouvoir la mise en place de cadres de gouvernance et d'action interopérables, en tant que de besoin.
- 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 efficacement les systèmes d'IA et d'interagir avec eux, 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 de l'IA, en particulier par le biais de programmes de formation tout au long de la vie active, du soutien aux personnes affectées par les suppressions de postes, au moyen notamment de la protection sociale, 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 et des services publics, 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 développement, en association avec les parties prenantes, devraient coopérer activement afin de faire avancer 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 du numérique, par l'intermédiaire du Groupe de travail sur la gouvernance de l'intelligence artificielle, de :

- a) 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) continuer d'élaborer et d'enrichir les orientations pratiques sur la mise en œuvre de la présente Recommandation, afin de répondre aux évolutions à l'œuvre et aux nouvelles priorités d'action ;
- c) 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) faire rapport au Conseil, en concertation avec les autres comités concernés, sur la mise en œuvre, la diffusion et le maintien de la pertinence de la présente Recommandation au plus tard cinq ans après sa révision, puis au moins tous les dix ans.