

# Groupe de travail sur l'intelligence artificielle et ses implications dans l'espace de l'information et de la communication

*Description du Groupe de Travail - Juillet 2023*

---

## 1/ Les implications croissantes de l'IA pour l'espace informationnel

Le développement rapide et les avancées en intelligence artificielle (IA) et son application dans tous les domaines de la vie ont des implications importantes pour l'espace mondial de l'information et de la communication. Les applications de l'IA deviennent facilement accessibles au grand public et peuvent être utilisées avec peu ou pas de connaissances en programmation. Les outils d'IA générative tels que ChatGPT, les outils de création d'images ou de vidéos, permettent à quiconque de devenir créateur de contenu. Les applications d'IA prédictive jouent un rôle décisionnel croissant dans la curation, la modération et la vérification des informations.

Cette prédominance de l'IA soulève d'importants défis :

- L'IA générative produit des contenus qui, en plus des préoccupations liées aux droits d'auteur et à la vie privée, soulève des défis quant à l'exactitude de ces contenus. L'IA peut inventer des sources, créer des informations erronées et des deep fakes. Ces informations peuvent être utilisées, involontairement ou intentionnellement, pour diffuser de la désinformation ou de la désinformation, créer le chaos ou tromper. Il devient de plus en plus difficile de distinguer les contenus créés par l'IA de ceux créés par l'homme, réduisant la confiance dans l'espace mondial de l'information.
- Les applications de l'IA prennent des décisions cruciales dans la sphère de l'information et de la communication, car la quantité d'informations disponibles et de contenus créés dépasse les capacités humaines à consommer, trier, modérer et vérifier ces informations. Ces outils souffrent de biais, leurs processus de décision ne sont ni transparents, ni infaillibles. L'influence de ces applications d'IA, et donc des entreprises privées, sur le système d'information et de communication s'accroît.
- Globalement, l'IA crée une plus grande manipulation de l'espace de l'information et de la communication. Elle facilite la création de fausses informations difficiles à reconnaître. Elle permet également de suivre de plus près les préférences et le comportement des citoyens afin de cibler le contenu et les messages. Cela peut avoir des effets néfastes sur le pluralisme de l'information et la diversité des médias.

La régulation de la création, du déploiement et l'utilisation de l'intelligence artificielle occupe une place importante dans l'agenda politique. Les principes de l'OCDE sur l'intelligence artificielle adoptés en 2019, la recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle adoptée en 2021 et les négociations en cours pour une loi de l'UE sur l'intelligence artificielle n'en sont que quelques exemples.

Pourtant, les deux défis les plus urgents quant à l'influence de l'IA sur l'espace informationnel demeurent. Ils concernent :

- Les processus de conception et de déploiement ainsi que les règles des outils d'IA afin de minimiser les risques pour un espace global de l'information pluraliste, libre et fiable, favorable à la démocratie.
- La question de la responsabilité légale pour le contenu créé et diffusé par l'IA ainsi que des décisions prises par celle-ci.

## 2/ Objectifs du groupe de travail

Le Forum sur l'information et la démocratie lance un groupe de travail dédié à **l'intelligence artificielle et à ses implications sur l'espace de l'information et de la communication**, avec l'ambition de proposer des principes et des recommandations pour garantir que l'intelligence artificielle contribue à l'intégrité de l'information plutôt que de contribuer un risque pour l'espace de l'information et de la communication.

Le groupe de travail examinera en particulier les questions suivantes :

**Objectif 1** : Définir des garde-fous pour le développement et le déploiement des applications de l'IA ayant un impact sur l'espace de l'information et de la communication.

Le groupe de travail proposera des gardes de fous et des réglementations à respecter lors de la conception, du développement et du déploiement des applications d'IA afin de garantir que celles-ci respectent les normes de sécurité, les règles et réglementations existantes et de minimiser les risques des applications d'IA pour l'espace global de l'information. Dans la pratique, la majorité des produits mis sur le marché (produits pharmaceutiques, jouets, etc.) doivent respecter des règles strictes en matière de développement et de production avant d'être mis à la disposition des consommateurs. De même, les applications de l'IA devraient être réglementées. Le groupe de travail proposera:

- Des règles et des mécanismes d'application pour garantir que les développeurs d'applications d'IA, notamment d'IA générative, respectent les droits de propriété intellectuelle et les droits à la vie privée.
- La définition d'évaluations des risques, d'audits et de contrôles préalables en matière de droits de l'homme avant le lancement de nouvelles applications d'IA, en mettant l'accent sur leur impact sur l'espace global de l'information.
- Des mesures visant à garantir l'intégrité des ensembles de données utilisés pour développer les applications d'IA et à réduire le niveau de biais inhérent à ces systèmes.
- Des mesures de réduction des risques et des dispositifs de protection pour privilégier l'utilisation éthique des applications d'IA
- Des régimes de transparence pour renforcer la fiabilité des applications d'IA

**Objectif 2 :** Définir des régimes de responsabilité pour les développeurs, les déployeurs et les utilisateurs des applications d'IA en ce qui concerne les informations produites, traitées, modérées et vérifiées par le biais de ces applications d'IA.

Le groupe de travail proposera des régimes de responsabilité à mettre en place pour les développeurs, les déployeurs et les utilisateurs des applications d'IA, qui permettraient de protéger l'espace mondial de l'information de la manipulation, de la désinformation et des "deep fakes". Il fera des recommandations sur :

- Les responsabilités et obligations respectives, c'est-à-dire les régimes de responsabilité, des utilisateurs, développeurs et déployeurs des applications d'IA en ce qui concerne le contenu créé par les applications d'IA, étant donné que la responsabilité est plus diffuse dans la chaîne de production. Ces régimes incluront également la responsabilité en matière de droits et d'obligations de propriété intellectuelle, ainsi que la confidentialité des données. Ils aborderont également les responsabilités éditoriales en ce qui concerne le contenu créé.
- Les régimes de responsabilité pour les développeurs, les déployeurs et les utilisateurs en ce qui concerne les décisions prises par les applications d'IA qui modèrent, organisent et vérifient le contenu. Il s'agira de déterminer dans quelle mesure la prise de décision peut être confinée aux applications d'IA.
- Les droits et devoirs à inclure dans les conditions d'utilisation des applications d'IA. Cette question sera également liée à la construction d'une littératie de l'IA afin de permettre aux utilisateurs des applications d'IA, y compris les journalistes, de mieux comprendre les outils qu'ils utilisent.

**Objectif 3 :** Recommander des options de gouvernance pour le déploiement et le suivi des applications de l'IA ayant un impact sur l'espace de l'information et de la communication.

Bien que plusieurs initiatives visant à réglementer l'IA se développent, il n'existe actuellement aucun cadre de gouvernance clair. Par conséquent, le groupe de travail recommandera des options de gouvernance pour surveiller le déploiement des applications d'IA, le respect des règles de développement et les régimes de responsabilité. Il élaborera des recommandations sur :

- La pertinence des processus de licences garantissant le respect des normes de sécurité numérique, ainsi que leurs conditions et les licences spécifiques.
- Les mécanismes et institutions de contrôle, y compris leurs pouvoirs et leur structure institutionnelle, pour surveiller le respect des réglementations en matière de développement et de régimes de responsabilité.
- Les mécanismes de recours et d'appel potentiels en cas de non-respect des réglementations.

### 3/ Calendrier

	Aôut	Sept	Oct	Nov	Dec	Jan	Feb
<b>Lancement du groupe de travail</b> (comité de pilotage, rapporteurs, appel à contribution)	■	■					
<b>Rassembler les contributions</b> (appel global à contributions,, entretiens avec les comités de pilotage, etc.)		■	■	■	■		
<b>Rédaction des recommandations</b>				■	■	■	
<b>Consultation des parties prenantes</b> (gouvernements, plateformes, organisations médiatiques, etc.)				■	■	■	
<b>Publication des recommandations</b>							■

### 4/ Définitions clés préliminaires

**Système d'IA :** Un système d'IA est un système basé sur une machine qui peut, pour un ensemble donné d'objectifs définis par l'homme, générer des résultats tels que du contenu, des prédictions, des recommandations ou des décisions influençant des environnements réels ou virtuels. Les systèmes d'IA sont conçus pour fonctionner avec différents niveaux d'autonomie (OCDE, 2019 et projet de loi de l'UE).

**IA générative :** modèles de base utilisés dans les systèmes d'IA spécifiquement destinés à générer, avec des niveaux d'autonomie variables, des contenus tels que des textes complexes, des images, des sons ou des vidéos (droit de l'UE).

**Modération :** Lorsque les plateformes en ligne filtrent et contrôlent le contenu généré par les utilisateurs sur la base de règles et de lignes directrices spécifiques à la plateforme afin de déterminer si elles peuvent héberger ou continuer à héberger un contenu spécifique dans le cadre de leurs conditions de service. Ces décisions peuvent inclure le retrait de contenus, de manière permanente, temporaire ou par zone géographique (Rapport de responsabilité).

**Curation:** Décisions relatives au filtrage et au classement des contenus destinés à l'attention humaine (Rapport sur le pluralisme)

**Vérification/détection de l'IA :** Lorsque des outils d'IA sont utilisés pour vérifier les informations et détecter les contenus produits par l'IA.

**Développeur/constructeur :** personne physique ou morale, autorité publique, agence ou autre organisme qui conçoit et développe un système d'IA.

**Déploieur/fournisseur :** personne physique ou morale, autorité publique, agence ou autre organisme qui développe ou fait développer un système d'IA en vue de le mettre sur

le marché ou de le mettre en service sous son propre nom ou sa propre marque, à titre onéreux ou gratuit (projet de loi de l'UE).

**Utilisateur** : toute personne physique ou morale, autorité publique, agence ou autre organisme utilisant un système d'IA sous son autorité (droit de l'UE), y compris à des fins personnelles et non professionnelles (contrairement au droit de l'UE).

Responsabilité : être légalement responsable de quelque chose

**Responsabilité (Liability)**: être légalement responsable de quelque chose

**Obligation (Responsibility)**: remplir un devoir pour lequel on peut être tenu responsable.

**Imputabilité (Accountability)**: le fait d'être responsable de quelque chose et d'en rendre compte au public.

## 5/ Lectures préliminaires (en anglais) :

- [Artificial Intelligence and Democracy](#), January 2022 by Thorsten Thiel, Heinrich Böll Foundation
- [Who Is Liable for A.I. Creations?](#) June 3, 2023, New York Times
- [The impact of AI on content accuracy](#), May 11, 2023 by AIContentfy team
- [Implications of AI-driven tools in the media for freedom of expression](#) March 2019, by the Institute for Information Law
- [Synthetic Lies: Understanding AI-Generated Misinformation and Evaluating Algorithmic and Human Solutions](#), April 2023, Georgia Institute of Technology
- [Governing artificial intelligence in the media and communications sector](#), 21 Feb 2023, Internet Policy Review
- [Artificial Intelligence and Freedom of Expression](#), 2020, OSCE
- [Spotlight on Artificial Intelligence and Freedom of Expression](#), A Policy Manual, 2021, OSCE
- [AI and Global Governance: Modalities, Rationales, Tensions](#), 2023, Veale, Matus & Gorwa
- [AI language models technological, socio-economic and policy considerations](#), 13 Apr 2023, OECD
- [Regulating ChatGPT and other Large Generative AI Models](#), May 12, 2023
- [Toward General Design Principles for Generative AI Applications](#), 13 January 2023, IBM Research AI, USA
- [Artificial Intelligence and Democratic Values Index](#), April 2023, Center for AI and Digital Policy
- [AI model GPT-3 \(dis\)informs us better than humans](#), 28 June 2023, Science Advances
- [Towards Guidelines for Guidelines on the Use of Generative AI in Newsrooms](#), 10 July 2023, Hannes Cools
- [Towards a new regime of civil liability for AI systems: Comment on the European Commission's Proposals](#), October 14, 2022, Alexandre Lodie, Stephanie Celis J. and Theodoros Karathanasis
- [The three challenges of AI regulation](#), Tom Wheeler, Thursday, June 15, 2023
- [Generative AI in Media & Journalism: Think Big, But Read the Small Print First](#), Natali Helberger, 18 July 2023
- [Artificial Intelligence Act How the EU can take on the challenge posed by general-purpose AI systems](#), Maximilian Gahntz and Claire Pershan, Nov. 9, 2022

- [A law for foundation models: the EU AI Act can improve regulation for fairer competition](#), July 20, 2023, Kai Zenner
- [Supporting Open Source and Open Science in the EU AI Act](#), July 2023
- [State-of-the-art Foundation AI Models Should be Accompanied by Detection Mechanisms as a Condition of Public Release](#), July 2023, Global Partnership on Artificial Intelligence'

## 6/ AI Initiatives:

- **EU:** [Directive sur la responsabilité en matière d'IA](#) (FR), [Groupe d'experts de haut niveau sur l'intelligence artificielle](#) (FR), [Loi sur l'IA de l'UE](#) (FR), Code de conduite potentiel sur l'IA générative, pacte potentiel sur l'IA avec les acteurs de l'industrie, mise à jour du code de pratique pour lutter contre la désinformation, Groupe européen d'éthique des sciences et des nouvelles technologies "[La démocratie à l'ère digitale](#)"(EN), [Lignes directrices en matière d'éthique pour une IA digne de confiance](#) (FR)- [Document de synthèse de l'Institut pour le futur de la vie sur la loi sur l'IA](#)(EN)
- **Conseil de l'Europe:** Projet de [convention sur l'intelligence artificielle, les droits de l'homme, la démocratie et l'État de droit](#) (FR), Commission de l'intelligence artificielle - [Déclaration de la société civile](#) (EN)
- **UNESCO:** [Déclaration de Windhoek sur l'intelligence artificielle en Afrique australe Windhoek \(Namibie\)](#) (FR), 2022, [Recommandation sur l'éthique de l'intelligence artificielle](#) (FR), 2021, [Méthode d'évaluation de l'état de préparation : un outil de la Recommandation sur l'éthique de l'intelligence artificielle](#) (FR)
- **UN HCHR:** [Résolution sur les technologies numériques émergentes et les droits de l'homme](#) (FR), Juillet 2023
- **Digital Content Next:**[Principes de développement et de gouvernance de l'IA générative](#) (EN)
- **European University Institute:** [Chaire Intelligence artificielle et démocratie](#) (EN)
- **EU-US Trade and Technology Council:** [Feuille de route sur l'IA](#) (EN), 21 Feb 2023
- **OSCE:** #SAIFE:<https://www.osce.org/fom/ai-free-speech> (EN)
- [Responsible AI Licenses](#)
- **Partnership on AI:** [Projet d'IA générative responsable](#) (EN) , [Pratiques responsables pour les médias synthétiques de la PAI](#) (EN)
- **OECD:** [OECD Observatoire des politiques de l'IA](#) (FR), 2019 [Principes sur l'IA](#) (FR)
- **Globalpolicy.AI:** [coalition/plateforme en ligne comprenant 8 OIG travaillant sur l'intelligence artificielle](#) (FR)
- **US:** [Projet de charte des droits de l'IA](#) (EN), [engagements volontaires des principales entreprises du secteur de l'intelligence artificielle](#) (EN)
- **AU** Stratégie continentale en matière d'intelligence artificielle, annoncée pour 2024
- **Canada:** Prochaine [loi sur l'intelligence artificielle et les données](#) (FR)
- **AI4Media:** [Projet de recherche](#)
- **France:** [Amendement à la loi](#) (FR) sur la sécurisation de l'espace numérique, sur la sanction des deepfakes créés par l'IA générative

- **UK:** [Document d'orientation : Une approche pro-innovation de la réglementation de l'IA présenté au Parlement par le Secrétaire d'État à la science, à l'innovation et à la technologie sur ordre de sa Majesté](#) (EN)
- **Brésil:** [Projet de loi sur l'IA](#) (EN)
- **Japon:** [Lignes directrices en matière de gouvernance pour la mise en œuvre des principes de l'IA](#) (EN)
- **Australie:** [Document de discussion sur le soutien à l'IA responsable](#) (EN)
- **Harvard Law School and Berkman Klein Center:** [Initiative sur l'intelligence artificielle et le droit](#) (EN)
- **[AI Now Institute](#)**